

TESIS DOCTORAL
Departamento de Psicología Evolutiva y de la Educación



**VNiVERSiDAD
D SALAMANCA**

**La interacción entre función
y modalidad del lenguaje.
Consecuencias teóricas y prácticas.**

Héctor García Rodicio

Director
Emilio Sánchez Miguel

Salamanca 2009

TESIS DOCTORAL
Departamento de Psicología Evolutiva y de la Educación



VNiVERSiDAD
D SALAMANCA

**La interacción entre función
y modalidad del lenguaje.
Consecuencias teóricas y prácticas.**

Autor

Héctor García Rodicio

Director

Emilio Sánchez Miguel

THESIS
Department of Developmental and Educational Psychology



VNiVERSiDAD
D SALAMANCA

**The interaction between function
and modality of language.
Theoretical and practical
implications.**

Héctor García Rodicio

Director
Emilio Sánchez Miguel

Salamanca 2009

Tabla de contenidos / Table of contents

Capítulo 0. Panorámica de la tesis, 1

- 0. Introducción, 1
- 1. La pregunta que tratamos de responder, 1
 - 1.1. Las ayudas para detectar y reparar en los CBLE, 2
 - 1.2. ¿Cómo diseñar ayudas para detectar y reparar?, 2
- 2. Un marco de investigación: la aproximación funcional, 3
 - 2.1. Los supuestos, 4
 - 2.1.1. Primero: existen distintos tipos de ayudas dependiendo de su función, 4
 - 2.1.2. Segundo: cada modalidad tiene características específicas, 4
 - 2.1.3. Tercero: existe una interacción función x modalidad, 5
 - 2.2. Predicciones y método para probarlas, 5
- 3. Plan de trabajo, 6
 - 3.1. Primero capítulo. Las funciones del lenguaje, 6
 - 3.2. Segundo capítulo. Las modalidades del lenguaje, 6
 - 3.3. Tercer capítulo. La interacción función x modalidad, 7
 - 3.4. Cuarto capítulo. Los experimentos, 7
 - 3.5. Consecuencias, 7

Capítulo 1. Las funciones del lenguaje: regular y declarar, 9

- 0. Introducción, 9
- 1. El lenguaje es un sistema de comunicación, 11
 - 1.1. Una definición de lenguaje, 11
 - 1.2. La comunicación como propiedad fundamental, 12
- 2. Una propuesta para definir la comunicación humana, 14
 - 2.1. Supuestos, 14
 - 2.1.1. Cuatro supuestos, 14
 - 2.1.2. Dos supuestos adicionales, 17
 - 2.2. Prerrequisitos para que exista comunicación, 18
 - 2.2.1. ¿Qué es entender el lenguaje? Las representaciones mentales, 19
 - 2.2.1.1. Qué nos dicen, 19
 - 2.2.1.2. Qué nos quieren decir, 20
 - 2.2.1.3. Qué pretenden con lo que dicen, 20
 - 2.2.1.4. Qué actitud tienen para con lo que dicen, 21
 - 2.2.1.5. Conclusión: qué es comprender, 21
 - 2.2.2. ¿Qué evidencias hay de que existen esas representaciones?, 21
 - 2.2.2.1. Modelos de situación, 21
 - 2.2.2.2. Implicaturas, 22
 - 2.2.2.3. Intención comunicativa, 23
 - 2.2.2.4. Actitudes, 24
 - 2.2.2.5. Conclusión, 24
 - 2.2.3. ¿Cómo construimos esas representaciones?, 25
 - 2.2.3.1. ¿Qué procesos intervienen?, 25
 - 2.2.3.2. ¿Son automáticos o estratégicos esos procesos?, 27

- 2.2.4. ¿Qué conocimientos, pistas y habilidades son necesarios?, 28
 - 2.2.4.1. Conocimientos, 28
 - 2.2.4.2. Pistas, 28
 - 2.2.4.3. Habilidades, 29
- 2.2.5. Resumen, 30
- 3. Revisando los apoyos a los supuestos de la definición propuesta, 31
 - 3.1. ¿Hay dos o más funciones comunicativas?, 31
 - 3.1.1. Hay muchos actos del habla, 31
 - 3.1.1.1. Los actos ilocutivos, 31
 - 3.1.1.2. Muchos actos ilocutivos: muchas funciones, 32
 - 3.1.1.3. Intensidad en la elocución, 33
 - 3.1.1.4. Conclusión, 33
 - 3.1.2. Las taxonomías de actos ilocutivos incluyen más de dos categorías, 34
 - 3.1.2.1. La clasificación de Searle, 34
 - 3.1.2.2. La clasificación de Dore, 36
 - 3.1.2.3. La clasificación de Stiles, 38
 - 3.1.2.4. Conclusión, 39
 - 3.1.3. Existen regulaciones que declaran, 40
 - 3.1.4. Todo acto comunicativo persigue regular, 41
 - 3.1.5. La realidad psicológica de las funciones, 42
 - 3.1.6. No todo son problemas: los fundamentos de la taxonomía, 44
 - 3.1.6.1. La clasificación de Givón, 45
 - 3.1.6.2. La clasificación de Werner, 47
 - 3.1.6.3. La clasificación de Halliday, 48
 - 3.1.6.4. Conclusión, 49
 - 3.1.7. Balance del primer supuesto, 49
 - 3.2. ¿Cada función requiere procesos distintos?, 51
 - 3.2.1. La regulación aparece antes que la declaración, 51
 - 3.2.2. Sin teoría de la mente no hay declaración: los autistas, 53
 - 3.2.3. Los patrones prosódicos en la comunicación temprana, 55
 - 3.2.4. La prevalencia regulatoria en la ontogenia, 56
 - 3.2.5. La prevalencia regulatoria en la filogenia, 57
 - 3.2.6. Balance del segundo supuesto, 58
 - 3.2.7. Evidencias adicionales posibles, 59
 - 3.2.7.1. Lesiones cerebrales, 59
 - 3.2.7.2. Variables críticas, 60
 - 3.2.7.3. Conclusión, 61
- 4. Resumen, predicciones, consecuencias, 61
 - 4.1. Predicciones y consecuencias, 63

Capítulo 2. Las modalidades del lenguaje: control escrito y expresividad oral, 65

- 0. Introducción, 65
- 1. ¿Cuáles son las características específicas de lo oral y lo escrito?, 66
 - 1.1. Propuestas, 66
 - 1.1.1. ¿Qué decimos nosotros?, 66
 - 1.1.2. ¿Qué se dice en el campo de la adquisición de la lectura?, 67

- 1.1.3. ¿Qué dicen otros psicólogos y lingüistas?, 68
 - 1.1.3.1. Nivel de input, 69
 - 1.1.3.2. Nivel léxico sintáctico, 69
 - 1.1.3.3. Nivel semántico, 71
 - 1.1.3.4. Nivel pragmático, 72
 - 1.1.3.5. Nivel interactivo, 76
 - 1.1.3.6. Nivel de pensamiento, 77
- 1.2. Catálogo de características específicas: resumen de las propuestas, 77
- 2. ¿Son relevantes las características identificadas?, 78
 - 2.1. Nivel de input, 79
 - 2.1.1. Canal sensorial, 79
 - 2.1.2. Control, 80
 - 2.1.3. Código, 82
 - 2.2. Nivel léxico sintáctico, 83
 - 2.2.1. Densidad léxica, 83
 - 2.2.2. Elementos auxiliares, 83
 - 2.2.3. Integración, 86
 - 2.3. Nivel semántico, 87
 - 2.3.1. Carga informativa, 87
 - 2.3.2. Señalando la idea central, 87
 - 2.4. Nivel pragmático, 88
 - 2.4.1. Prosodia, 88
 - 2.4.1.1. La prosodia se produce, 88
 - 2.4.1.2. La prosodia se capta, 89
 - 2.4.1.3. Prosodia en la adquisición del lenguaje, 90
 - 2.4.1.4. Prosodia en la desambigación de oraciones, 90
 - 2.4.1.5. Prosodia en la predicción de fin de enunciado, 91
 - 2.4.1.6. Prosodia encarnando actitudes, 92
 - 2.4.1.7. Resumen, 94
 - 2.4.2. Gestos (manuales), 94
 - 2.4.2.1. Los gestos se producen, 94
 - 2.4.2.2. Los gestos se captan, 96
 - 2.4.2.3. Gestos deícticos, 97
 - 2.4.2.4. Gestos icónicos, 98
 - 2.4.2.5. Gestos metafóricos, 98
 - 2.4.2.6. Gestos pragmáticos, 99
 - 2.4.2.7. Resumen, 99
 - 2.4.3. Eventos perceptibles, 99
 - 2.4.4. Propiedades atribuibles al hablante y la situación, 100
 - 2.5. Nivel interactivo, 102
 - 2.5.1. Interacción abierta, 102
 - 2.5.2. Interacción encubierta, 103
 - 2.6. Impacto en el pensamiento, 105
 - 2.7. Balance, 106
- 3. ¿Son realmente específicas las características identificadas?, 108
 - 3.1. ¿Inherentes? Factores coyunturales, 108

- 3.1.1. Presencia de imágenes, 108
- 3.1.2. Cuando la decodificación se automatiza, 109
- 3.1.3. Cuestión de género: lo expositivo, 110
- 3.1.4. Conversación cara a cara, 111
- 3.1.5. La expresividad y el control en distintas coyunturas, 112
- 3.2. ¿Exclusivos? La expresividad escrita y el control oral, 113
- 3.3. ¿Descartamos definitivamente las otras variables identificadas?, 114
- 4. Resumen, predicciones, consecuencias, 114
 - 4.1. Predicciones y consecuencias, 116

Capítulo 3. La interacción función x modalidad en su contexto: diseñando ayudas para los computer-based learning environments, 118

- 0. Introducción, 118
- 1. Qué son los computer-based learning environments, 119
 - 1.1. Los materiales instructivos, 119
 - 1.2. Las posibilidades de los computer-based learning environments, 119
- 2. Aprender de los CBLE es difícil, 122
 - 2.1. Aprender a partir de CBLEs: construir modelos mentales, 122
 - 2.1.1. Aprender a partir de textos, 122
 - 2.1.1.1. Las representaciones: modelos mentales, 122
 - 2.1.1.2. Los procesos: decodificación, sintaxis e inferencias, 127
 - 2.1.1.3. Los conocimientos, 131
 - 2.1.2. Aprender a partir de palabras e imágenes: aprendizaje multimedia, 131
 - 2.1.2.1. Ventajas de la inclusión de imágenes, 131
 - 2.1.2.2. Desventajas, 133
 - 2.2. Obstáculo 1: tópicos complejos, 133
 - 2.3. Obstáculo 2: habilidades especiales, 135
 - 2.3.1. Integración verbal-pictórica, 135
 - 2.3.2. Navegación en hipertexto, 137
 - 2.4. Obstáculo 3: conocimiento previo específico de dominio, 138
 - 2.5. Obstáculo 4: memoria de trabajo, 140
 - 2.6. Interrelación entre las dificultades: la paradoja, 143
- 3. Ayudas para detectar y reparar, 145
 - 3.1. Aprendizaje estratégico: detectar y reparar, 146
 - 3.1.1. Procesamiento automático, 146
 - 3.1.2. Procesamiento estratégico, 147
 - 3.1.3. Detectar y reparar, 147
 - 3.2. La importancia de detectar y reparar, 149
 - 3.2.1. Estudios de comparación buenos/malos, 149
 - 3.2.2. Estudios de regresión, 151
 - 3.2.3. Estudios de entrenamiento, 153
 - 3.2.4. Resumen, 154
 - 3.3. Un nuevo problema: los aprendices tienen dificultades para detectar/reparar, 154
 - 3.3.1. Los aprendices no detectan, 155
 - 3.3.2. Los aprendices no reparan, 157
 - 3.3.3. Resumen, 159

- 3.4. Ayudando a detectar y reparar, 159
 - 3.4.1. Ayudas para detectar, 160
 - 3.4.2. Ayudas para reparar, 161
 - 3.4.3. Efectividad de las ayudas para detectar y reparar, 162
 - 3.4.3.1. Ayudas para detectar y reparar versus nada, 162
 - 3.4.3.2. Ayudas para detectar y reparar versus una de ellas, 163
 - 3.4.3.3. Ayudas para detectar o reparar versus nada, 164
 - 3.4.3.4. ¿Cómo explicar la falta de efectividad?, 164
 - 3.4.3.5. Conclusión, 165
- 4. Cómo presentar las ayudas para detectar y reparar, 165
 - 4.1. Diseñando ayudas para un CBLE: posibilidades múltiples, 165
 - 4.2. No hay pautas sobre cómo usar la modalidad para las ayudas verbales, 167
 - 4.2.1. El principio de modalidad no es aplicable, 167
 - 4.2.1.1. ¿Qué dice el principio de modalidad?, 167
 - 4.2.1.2. ¿Con qué apoyo cuenta el principio de modalidad?, 169
 - 4.2.1.3. Circunstancias moduladoras, 174
 - 4.2.1.4. Problemas, 174
 - 4.2.1.5. ¿Por qué no nos sirve el principio de modalidad para diseñar ayudas (verbales)?, 175
 - 4.2.2. Los resultados divergentes, 175
 - 4.2.3. Conclusión, 178
 - 4.3. Un marco de investigación: la aproximación funcional, 178
 - 4.3.1. Primer supuesto: hay distintas ayudas dependiendo de la función que cumplen, 178
 - 4.3.1.1. Las ayudas regulatorias, 178
 - 4.3.1.2. Las ayudas explicativas, 178
 - 4.3.1.3. Las ayudas en funcionamiento, 178
 - 4.3.2. Segundo supuesto: cada modalidad, características específicas, 179
 - 4.3.2.1. El control visual, 179
 - 4.3.2.2. La expresividad auditiva, 179
 - 4.3.3. Tercer supuesto: existe una interacción función x modalidad, 180
 - 4.3.3.1. Predicciones, 180
 - 4.4. ¿Encajan los resultados anteriores en la aproximación funcional?, 180
 - 4.5. Panorámica de los experimentos que siguen, 182

(Nota: La parte final se redactó en inglés ya que es un requisito para obtener el Doctorado Europeo)

Chapter 4. Experiments, 184

0. Abstract, 184

1. Introduction, 185

1.1. Aids to monitor and revise misunderstandings in CBLEs, 185

1.2. Designing aids in CBLEs: use of modality when constructing verbal aids, 186

1.3. The functional approach: a research framework, 187

1.3.1. First assumption, 187

1.3.2. Second assumption, 187

1.3.3. Third assumption, 188

- 1.4. Prior mixed results fit into the functional approach, 189
- 1.5. Overview of the study, 190
- 2. Experiment 1: The effect of the visual and the auditory modalities of different verbal aids on learning, 190
 - 2.1. Participants and design, 191
 - 2.2. Materials, 191
 - 2.2.1. Prior knowledge test, 191
 - 2.2.2. General reading comprehension skill test, 192
 - 2.2.3. Computer-based learning environment, 192
 - 2.2.4. Retention test, 194
 - 2.2.5. Transfer test, 195
 - 2.3. Procedure, 195
 - 2.4. Scoring, 196
 - 2.4.1. Prior knowledge test, 196
 - 2.4.2. General reading comprehension skill test, 197
 - 2.4.3. Retention test, 197
 - 2.4.4. Transfer test, 197
 - 2.5. Results, 198
 - 2.5.1. Control variables, 198
 - 2.5.2. Preliminary analyses, 198
 - 2.5.3. Retention test, 199
 - 2.5.4. Transfer test, 199
 - 2.5.5. Time recordings, 199
 - 2.6. Discussion, 200
- 3. Experiment 2: Exploring the impact of the modality of different aids with a pedagogical agent, 202
 - 3.1. Participants and design, 204
 - 3.2. Materials, 205
 - 3.2.1. Computer-based learning environment, 205
 - 3.3. Procedure, 205
 - 3.4. Scoring, 206
 - 3.5. Results, 206
 - 3.5.1. Control variables, 207
 - 3.5.2. Retention test, 207
 - 3.5.3. Transfer test, 207
 - 3.5.4. Time recordings, 207
 - 3.6. Discussion, 207
- 4. Experiment 3: Testing the superiority of the auditory regulatory aids with demanding materials, 210
 - 4.1. Participants and design, 211
 - 4.2. Materials, 211
 - 4.2.1. Prior knowledge test, 211
 - 4.2.2. General reading comprehension skill, 211
 - 4.2.3. Computer-based learning environment, 211
 - 4.2.4. Retention test, 213
 - 4.2.5. Transfer test, 213

- 4.3. Procedure, 214
- 4.4. Scoring, 215
 - 4.4.1. Prior knowledge test, 215
 - 4.4.2. General reading comprehension skill test, 215
 - 4.4.3. Retention test, 215
 - 4.4.4. Transfer test, 216
- 4.5. Results, 216
 - 4.5.1. Control measures, 217
 - 4.5.2. Preliminary analyses, 217
 - 4.5.3. Retention test, 218
 - 4.5.4. Transfer test, 218
- 4.6. Discussion, 218
- 5. Experiment 4: Testing the superiority of the auditory regulatory aids in participants with low self-regulation and literacy, 219
 - 5.1. Participants and design, 221
 - 5.2. Materials, 222
 - 5.2.1. General reading comprehension skill test, 222
 - 5.2.2. Computer-based learning environment, 222
 - 5.3. Procedure, 222
 - 5.4. Scoring, 223
 - 5.5. Results, 223
 - 5.5.1. Control measures, 224
 - 5.5.2. Preliminary analyses, 224
 - 5.5.3. Retention test, 225
 - 5.5.4. Transfer test, 225
 - 5.6. Discussion, 225

Chapter 5. Practical and theoretical implications, 228

- 0. Summary, 228
- 1. Practical implications, 230
- 2. A new look at prior research, 230
- 3. Theoretical implications, 231
 - 3.1. The functions of language, 231
 - 3.2. Differences between reading and listening, 232
- 4. Limitations and future directions, 233

References, 234

Appendix (A), 250

Appendix (B), 253

Capítulo 0. Panorámica de la tesis.

0. Introducción.

Aprendemos. Aprendemos a andar, a inferir qué quieren los demás a partir de sus acciones, que el fuego quema y que un ejército puede más que unos explosivos de fabricación casera. Y aprendemos todo eso sin demasiado esfuerzo y sin que nadie se empeñe en ayudarnos. Basta ponerse en pie y dirigirse al frente, convivir con otros, explorar el entorno u hojear el periódico. Y todo marcha bien. Entonces ¿para qué intentar enseñar? ¿Para qué brindar ayudas a los aprendices? Pues porque algunas veces (la mayoría, en un contexto educativo formal) no es tan fácil aprender: no basta en esos casos con con explorar el entorno u hojear el periódico. Para aprender geología, ingeniería o biología no sirve el camino fácil, eso de dejarse llevar o registrar asociaciones entre estímulos y respuestas. Aprender geología o biología es difícil: requiere conciencia, deliberación y esfuerzo y no siempre uno es capaz de poner en juego todos esos elementos por sí mismo. Por eso tiene sentido enseñar, brindar ayudas. Aquí es donde entramos a escena quienes intentamos desarrollarlas para hacerlas eficaces. De eso trata esta tesis: de diseñar ayudas para facilitar aprendizajes que reclaman esfuerzo y de diseñarlas de la mejor forma posible. Como enseguida entenderemos, el objeto específico de la tesis es explorar cómo deben presentarse ciertas ayudas insertas en ciertos materiales instructivos.

1. La pregunta que tratamos de responder.

Los materiales instructivos son aquellos que tienen por objeto enseñar los conocimientos necesarios para que los aprendices desempeñen su papel en la sociedad (e.g., libro de texto). A nosotros nos ha interesado un tipo particular de material: los *Computer-Based Learning Environments* (o CBLEs, de ahora en adelante). Por CBLE entendemos aquel material instructivo (a) con soporte digital, (b) que presenta información verbal (textos, narraciones) y pictórica (imágenes, animaciones) y (c) que emplea distintas modalidades (auditiva/visual) para ello. Resulta que aprender a partir de un CBLE no siempre es tarea fácil. ¿Por qué? Pues porque es uno de esos aprendizajes que exigen conciencia, deliberación y esfuerzo. Por eso es común que los CBLEs incorporen algún tipo de asistencia o ayuda. La cuestión es que, dada la versatilidad de los CBLEs, estas ayudas pueden diseñarse de maneras

distintas. Una pregunta clave es: ¿debemos presentar estas ayudas en la modalidad visual o en la auditiva? Ésta es la incógnita que tratamos de despejar con esta tesis. Y lo interesante es que resolver la incógnita hará posible, de paso, abordar otras preguntas de mayor envergadura para la Psicología.

1.1. Las ayudas para detectar y reparar en los CBLE.

Decíamos que aprender de los CBLEs puede resultar difícil. Si es así es por varias razones. Primera, los CBLEs versan sobre tópicos complejos, como la geología, la biología o la ingeniería; ello exige a los aprendices incorporar sistemas conceptuales que incluyen un buen número de ideas y relaciones. Segunda, los CBLEs demandan la puesta en funcionamiento de procesos especiales como la integración de información verbal y pictórica o la navegación a través de hipervínculos. Tercera, los aprendices carecen de los conocimientos previos necesarios para interpretar adecuadamente dichos sistemas conceptuales; o, peor aún, el conocimiento que poseen es erróneo, lo que contamina el proceso comprensión. Cuarta, la capacidad de la memoria de trabajo es limitada, lo cual restringe la posibilidad de procesar mucha información de una vez. O sea, los CBLEs demandan procesar un gran volumen de información (e.g., temas complejos, información verbal y pictórica) y el sistema de procesamiento con el que contamos es limitado (e.g., capacidad de la memoria de trabajo, ideas erróneas que consumen recursos). El resultado de esta paradoja es que la construcción de representaciones mentales incompletas y/o incoherentes a partir de los CBLEs es más que probable.

Para evitar tal resultado indeseable, los aprendices deben poner en juego procesos metacognitivos de *detección* y *reparación* de problemas de comprensión. Es decir, deben supervisar sus representaciones mentales en busca de malentendidos y deben generar explicaciones para reconstruirlas cuando es menester. Pero, claro, esto no es tampoco tarea fácil. De hecho, está bien documentado que los aprendices tienen dificultades tanto para supervisar su comprensión a fin de detectar malentendidos como para generar soluciones que permitan repararlos.

Lo que nos corresponde entonces a los psicólogos de la educación es diseñar sistemas de apoyo o ayudas que asistan a los aprendices en la ejecución de los mencionados procesos e incorporarlos a los CBLEs. Así, muchos investigadores han venido proponiendo diversas ayudas para los procesos de detección y reparación. Las evidencias empíricas ponen de manifiesto que estas ayudas son efectivas, en el sentido de que (a) posibilitan a los aprendices ejecutar los procesos de detección y reparación para (b) construir representaciones mentales más completas y coherentes.

1.2. ¿Cómo diseñar ayudas para detectar y reparar?

Decíamos también que los CBLEs son muy versátiles. Lo son porque pueden presentar información verbal y pictórica a través de las modalidades auditiva y visual. Eso significa que las ayudas pueden ser verbales o pictóricas y que pueden presentarse vía visual o auditiva. Así, por ejemplo, una ayuda para detectar un problema de comprensión puede adoptar la forma (1) de un mensaje verbal presentado en la pantalla del ordenador (e.g., “no es correcto

pensar así”), (2) del mismo mensaje verbal presentado en audio, (3) de una flecha o círculo señalando un aspecto que el aprendiz no debe omitir de su representación mental o (4) de un sonido que indica que algo no va bien (e.g., “wreeenk”). De todas estas posibilidades a nosotros nos interesa investigar las dos primeras: es decir, las *ayudas verbales* presentadas bien por vía visual bien por vía auditiva. Y si nos interesan esas dos es porque, como se razonará a continuación, aún no sabemos con certeza qué modalidad es mejor para presentar las ayudas verbales, la visual o la auditiva. La modalidad mejor para las ayudas, claro está, será la que logre hacer más eficaces las ayudas y, por extensión, promueva un mayor aprendizaje.

¿Por qué no sabemos qué modalidad emplear para presentar las ayudas? Pues por dos razones. En primer lugar, el *principio de modalidad* no es aplicable a las ayudas verbales. El principio de modalidad establece que allí donde haya información verbal acompañando a información pictórica, la verbal debe presentarse en modalidad auditiva. Así, por ejemplo, las palabras que acompañan a una animación deben proporcionarse en forma de narraciones (en lugar de texto). Las razones de este modo de proceder son muy simples: (a) si presentamos lo verbal y lo pictórico por vía visual, el sistema de procesamiento visual queda saturado y (b) presentar información verbal y pictórica en diferentes modalidades permite explotar toda la capacidad de la memoria de trabajo (que, según algunos, posee procesadores relativamente independientes para cada modalidad). El principio ha sido ampliamente avalado, como demuestra un reciente meta-análisis. No obstante, las ayudas no (necesariamente) están asociadas a información pictórica, lo cual deja al principio de modalidad fuera de juego en cuanto a diseño de ayudas verbales se refiere.

En segundo lugar, la poca investigación llevada a cabo hasta la fecha sobre el impacto de la modalidad de las ayudas en el aprendizaje arroja resultados divergentes. Que nos conste, sólo se han realizado cuatro estudios, cada cual con unos resultados particulares. Veremos que hay experimentos en los que las ayudas verbales auditivas resultan ser mejores que idénticas ayudas presentadas en modalidad visual; en otros, sin embargo, no habrá diferencias significativas entre una modalidad u otra. De modo que no hay un patrón claro de resultados que represente una base sobre la cual hacer predicciones.

Por estas dos razones el objetivo será (a) proponer y (b) someter a prueba un marco de investigación que determina cómo deben presentarse las ayudas. El marco será útil para reinterpretar los datos dispares de la investigación previa y para formular predicciones precisas de cómo debe rendir cada modalidad bajo ciertas condiciones.

2. Un marco de investigación: la aproximación funcional.

Hemos podido constatar que ni los argumentos teóricos ni los empíricos resuelven la incógnita que planteábamos al comienzo: ¿qué modalidad debemos emplear para las ayudas verbales? Como respuesta a tal estado de cosas, proponemos un marco de investigación: la *aproximación funcional*. El marco está basado en tres supuestos.

2.1. Los supuestos.

2.1.1. Primero: existen distintos tipos de ayuda dependiendo de su función.

Hay al menos dos tipos de ayudas, las cuales difieren en la función lingüística que desempeñan. Las *ayudas regulatorias* cumplen una función manipulativa o regulatoria, esto es, son enunciados mediante los cuales el hablante solicita algo al oyente (e.g., “debes revisar tu comprensión”) y que, por tanto, exigen al oyente sopesar si satisfacer o no lo que se le demanda. Estas ayudas son las que proporcionan apoyo a los aprendices a la hora de supervisar su comprensión para detectar malentendidos: son ayudas al proceso de detección.

Las *ayudas explicativas* desempeñan una función informativa o declarativa, es decir, son enunciados mediante los cuales el hablante afirma algo tratando de modificar las representaciones mentales del oyente (e.g., “el Himalaya se formó por el choque de dos placas continentales”) y que, por tanto, exigen al oyente procesar e integrar información. Estas ayudas son las que brindan apoyo a los aprendices para revisar y reparar sus malentendidos: son ayudas al proceso de reparación.

Veamos en qué consisten estas ayudas con un ejemplo sencillo. Imaginemos que un aprendiz aprende geología a partir de un CBLE que incluye animaciones con sus correspondientes narraciones. En un momento dado, el CBLE plantea una pregunta al aprendiz para la cual éste debe encontrar respuesta entre varias alternativas.

- “Hemos visto que en función del tipo de placas implicadas en un choque, pueden surgir montañas sin volcanes o montañas con volcanes. Pero, ¿sabemos cuáles son exactamente los procesos que hacen que unas veces haya volcanes y otras no?
 - (a) depende de la fuerza del choque;
 - (b) es una cuestión de tiempo, al final terminan surgiendo volcanes;
 - (c) cuando hay una placa oceánica se produce una presión que resquebraja la continental.”
 [el aprendiz selecciona (a)]

- [Ayuda regulatoria] “Si has elegido esta opción es porque piensas que un choque fuerte puede provocar el resquebrajamiento de una de las placas, formando grietas por las que surge magma. Pero no es correcto pensar así: deberías revisar tu comprensión.”
- [Ayuda explicativa] “La clave está en la composición de las placas, que determina su densidad. Si las placas son igual de densas, la presión que ejercen la una sobre la otra es la misma, de modo que sólo cabe esperar que se compriman mutuamente dando lugar a montañas sin volcanes. Si las placas son de distinta densidad, la presión que la más densa ejerce sobre la menos densa es mayor, por eso la primera logra resquebrajar a la segunda formando grietas por las que surge el magma; se forman así montañas con volcanes.”

2.1.2 Segundo: cada modalidad tiene características específicas, las cuales son potencialmente ventajosas.

La modalidad visual proporciona al aprendiz *control*. Esto significa que el aprendiz puede regular el ritmo de lectura, puede releer y puede concentrarse en cada momento en el segmento del texto de su interés. Hay evidencias de que tales posibilidades resultan beneficiosas a la hora de incorporar nuevas ideas y relaciones a nuestro

fondo de conocimiento y de que restringir estas posibilidades perjudica el aprendizaje. De modo que el control es potencialmente ventajoso.

La modalidad auditiva, por su parte, se sirve de la *expresividad*. La expresividad se basa en el uso de la prosodia (i.e., entonación, énfasis, pausas, etc. de la voz) para reflejar la intención comunicativa del hablante (e.g., pedir, afirmar) y su actitud hacia lo que está diciendo (e.g., si lo considera importante o urgente). Hay evidencias de que la gente se apoya en la expresividad para inferir tales cosas y de que esto es útil para cambiar su disposición a realizar ciertas tareas. Así que la expresividad es potencialmente ventajosa.

Estas características se consideran específicas en la medida en que son inherentes y exclusivas de cada modalidad. Inherentes porque no dependen de factores coyunturales sino que están presentes en todas las variantes de la modalidad en cuestión (e.g., sea un post-it, una carta personal o un e-mail, todos son controlables; sea una conversación telefónica, una grabación en audio o un discurso en viva voz, la expresividad sigue ahí). Exclusivas porque no aparecen en la otra modalidad (e.g., ni un post-it ni una carta pueden modular su entonación o volumen; ni una conversación ni un discurso son susceptibles de inspección autoadministrada).

2.1.3. Tercero: existe una interacción entre función y modalidad. Por un lado, al procesar un enunciado que les solicita tomar cierto curso de acción, los aprendices podrían encontrar útil el tener acceso a la intención y actitud del hablante a través de la expresividad. Dicha información les sería de utilidad al reconocer qué se espera de ellos (que revisen su comprensión) y lo importante/urgente de tal cosa; ello, presumiblemente, les conduciría a hacerlo con diligencia. Al fin y al cabo, sabemos que las actitudes de los otros hacia cierta tarea influyen sobre nuestra disposición a realizarla. Por consiguiente, podría haber una relación entre modalidad auditiva y ayudas regulatorias.

Por otro lado, al procesar una elaboración sobre los contenidos a aprender, los aprendices podrían encontrar útil el tener control sobre el input lingüístico (e.g., concentrándose en cierto segmento del texto de especial interés). Tal posibilidad les sería de utilidad para procesar la elaboración con la profundidad que requiere. Recordemos que hay evidencias de que el control facilita la incorporación de nuevas ideas y relaciones. Por consiguiente, podría haber una relación entre la modalidad visual y las ayudas explicativas.

2.2. Predicciones y método para probarlas.

De este tercer supuesto se desprenden dos hipótesis. Una, las ayudas regulatorias auditivas son mejores que las ayudas regulatorias visuales. Dos, las ayudas explicativas visuales son mejores que las ayudas explicativas auditivas. Las denominaremos, respectivamente, *hipótesis de la superioridad de la regulación auditiva* e *hipótesis de la superioridad de la explicación visual*.

Éstas son las predicciones que se pondrán a prueba en la parte empírica de la tesis. En cuatro experimentos los participantes aprendieron geología a partir de un CBLE que incluía animaciones con sus correspondientes narraciones y ayudas regulatorias y explicativas

presentadas bien en modalidad auditiva o bien en modalidad visual. Se trataba de ver si quienes recibían ayudas regulatorias auditivas aprendían más que quienes recibían idénticas ayudas presentadas visualmente y si quienes recibían ayudas explicativas visual aprendían más que aquellos que las recibían en modalidad auditiva.

3. Plan de trabajo.

El objetivo, hemos anunciado ya, es proponer y validar empíricamente un marco de investigación que especifica qué modalidad usar para ciertas ayudas. El marco está basado en tres supuestos, dos de los cuales necesitan ser argumentados y uno sometido a prueba empírica. Más específicamente, el marco asume que existen distintas ayudas dependiendo de la función lingüística que desempeñan: regulatoria o declarativa, o sea, que habrá que argumentar por qué partimos de la base de que existen esas dos funciones. El marco asume, además, que las características específicas de las modalidades auditiva y visual son expresividad y control, respectivamente; de manera que habrá que argumentar por qué consideramos que son éstas y no otras las características específicas de cada modalidad. El último supuesto establece una interacción entre función y modalidad del que se derivan dos predicciones, justo las que serán puestas a prueba en la parte empírica de la tesis. Los capítulos de la tesis se encargarán entonces de, según corresponda, argumentar o probar cada uno de los supuestos.

3.1. Primer capítulo. Las funciones del lenguaje.

En un primer capítulo distinguiremos dos funciones del lenguaje. Una vez argumentado que, por encima de la representación, el lenguaje es un sistema de comunicación, pasaremos a especificar cuáles son sus funciones como sistema de comunicación. Comprobaremos que, aunque hay cierto consenso entre los investigadores sobre algunas cuestiones, no todos están de acuerdo en otras. En concreto, veremos que, si bien se distinguen varias funciones del lenguaje, no todos coinciden en señalar las mismas. Nosotros argumentaremos por qué es apropiado pensar que hay dos funciones básicas, regular y declarar, que implican procesos psicológicos específicos. Pero además será necesario proponer evidencias de que son funciones realmente distintas. Por supuesto, presentaremos evidencias que lo respaldan.

Inevitablemente, una discusión de esas características exigirá revisar el trabajo no sólo de psicólogos sino también de lingüistas y filósofos. Eso es precisamente lo que el lector encontrará en el capítulo 1.

3.2. Segundo capítulo. Las modalidades del lenguaje.

En un segundo capítulo identificaremos las características específicas de las modalidades auditiva (oral) y visual (escrita) del lenguaje. Líneas arriba dijimos que, a nuestro parecer, las características específicas de cada modalidad son la expresividad y el control, respectivamente. Pero, claro, no todos los investigadores coinciden al identificar las características específicas de cada modalidad. Habrá que aportar argumentos que sustenten

nuestra postura. Lo que haremos será comprobar si todas las características señaladas son potencialmente relevantes y, en todo caso, si son realmente específicas. Recordemos que por específicas entendemos aquellas que son inherentes y exclusivas de cada modalidad.

Una vez más, la bibliografía consultada provendrá de orígenes dispares. El tema ha interesado a lingüistas y psicólogos (a su vez, interesados en la comunicación o en el aprendizaje) así que para dar la cobertura que el tema requiere atenderemos a los trabajos de unos y otros.

3.3. Tercer capítulo. La interacción función x modalidad.

Una vez argumentados los dos supuestos volveremos al punto de partida. Es decir, retomaremos la cuestión principal: ¿cómo diseñar ayudas en los CBLE? Profundizaremos en lo que ya se ha esbozado en esta introducción: las ayudas, que sabemos de gran utilidad, pueden presentarse en modalidades distintas pero no hay argumentos teóricos ni empíricos que nos permitan formular predicciones sobre cómo cada modalidad resultará. De modo que explicaremos por qué es preciso incorporar ayudas en los CBLE, revisaremos estudios en los que se constata su efectividad, repasaremos el estado de la cuestión en cuanto al impacto de su modalidad y presentaremos con mayor detalle la aproximación funcional. De ella, como sabemos, se desprenden dos predicciones que habrán de ser probadas empíricamente. Terminaremos el capítulo ofreciendo una panorámica de los experimentos que seguirán a continuación.

3.4. Cuarto capítulo. Los experimentos.

En este capítulo se someterán a prueba las hipótesis de la superioridad de la regulación auditiva y de la superioridad de la explicación visual. Más específicamente, se describirán los experimentos que llevamos a cabo, incluyendo su método y resultados. En los experimentos, estudiantes universitarios y de secundaria aprendieron geología a partir de un CBLE. El CBLE incluyó animaciones con narraciones simultáneas, además de ayudas regulatorias y explicativas. Manipulamos la modalidad de las ayudas regulatorias (auditiva/visual) y de las ayudas explicativas (auditiva/visual) para producir un diseño 2 x 2. Tras consultar el CBLE, los participantes resolvieron pruebas de recuerdo y de transferencia. Se esperaba que aquellos que recibiesen las ayudas configuradas de acuerdo con la aproximación funcional rindiesen en tales pruebas significativamente mejor que sus compañeros.

3.5. Quinto capítulo. Consecuencias.

Discutidos los resultados de los cuatro experimentos que llevamos a cabo, comprobaremos que las dos hipótesis formuladas se vieron confirmadas por los datos. ¿Qué implicaciones tiene un panorama de esas características? Para empezar, arrojaremos luz a la gran incógnita que motivaba estas páginas: qué modalidad emplear para las ayudas en los CBLE. Nuestro marco, una vez respaldado empíricamente, permitirá proponer pautas de diseño de ayudas. En concreto, un diseñador de CBLEs habrá de saber que las ayudas regulatorias encuentran en la modalidad auditiva su soporte idóneo mientras que las ayudas

explicativas deben ser presentadas en la modalidad visual. Lejos de ser triviales, estas dos pautas serán de gran utilidad en el diseño de CBLE porque, recordemos, no contábamos con argumentos teóricos ni tampoco empíricos que guiasen nuestros pasos.

Por otra parte, la distinción entre funciones del lenguaje resultará fortalecida por los datos recabados. Los resultados mostrarán que dependiendo de la función que cumpla la ayuda, cabrá esperar una relación particular con la modalidad. Es decir, la distinción entre funciones es acertada, en la medida en que los resultados serían ininterpretables sin dicha distinción: si no fuesen diferentes las funciones, ¿por qué las mismas variables les afectan de un modo distinto?

Por último, las características identificadas como específicas de cada modalidad resultarán ser críticas. Esto permitirá revisar los planteamientos de otros investigadores, los cuales no consideran estas características como tan cruciales. ¿Cómo explicar que las características que nosotros identificamos sean decisivas si dichas características se excluyen de toda consideración?

Un vistazo a las implicaciones permite apreciar la doble naturaleza de este trabajo. Resulta que un problema práctico, como es cómo configurar ayudas en CBLEs, ha hecho posible rescatar otros problemas de mayor calado. Concretamente, si existen funciones del lenguaje diferentes, cada cual implicando distintos procesos psicológicos, y cuáles son las diferencias críticas entre las modalidades del lenguaje. En esto, creemos nosotros, reside el interés de la tesis: aúna cuestiones teóricas y prácticas.

Capítulo 1.

Las funciones del lenguaje: regular y declarar.

0. Introducción.

La aproximación funcional, que es el marco de investigación que proponemos, se basa en varios supuestos. Uno de ellos es que existen distintos tipos de ayuda dependiendo de la función lingüística que cumple cada cual. Distinguimos así las ayudas regulatorias de las explicativas, las cuales cumplen una función regulatoria y declarativa, respectivamente. De manera que el marco asume que existen funciones del lenguaje distintas y que las funciones son justamente regular y declarar. Ahora la cuestión es, ¿en qué nos basamos para suponer tal cosa? Los fundamentos del mencionado supuesto serán lo que expondremos aquí.

Vamos a ver que el lenguaje es, por encima de otras cosas, un sistema de comunicación. Es cierto que también es un sistema de representación del conocimiento, lo que proporciona ventajas considerables. No obstante, hay indicios que sugieren que la comunicación fue la función primordial del lenguaje, la ventaja que resultó ser adaptativa en algún momento de la evolución. Por eso asumiremos que la función principal del lenguaje es actuar como sistema de comunicación.

Claro que, un sistema de comunicación puede, a su vez, servir a distintas funciones. Ésas son las que nos interesan. Para examinar sus funciones será necesario definir qué es un sistema de comunicación. De modo que expondremos una síntesis de lo que lingüistas, psicólogos, filósofos y científicos computacionales¹ opinan sobre la comunicación. Específicamente, propondremos cuatro proposiciones en las que parece haber (razonable) acuerdo. Primera, la unidad básica de comunicación es el acto comunicativo, que es todo acto (lingüístico o no) que un hablante despliega deliberadamente para decir algo y que un oyente capta y reconoce como tal. Segunda, todo acto tiene un referente y una función. Así, aunque “me gustaría un poco más de cerveza, por favor” y “lo que más me gusta es la cerveza” tienen el mismo referente (gustar, cerveza, más), su función es solicitar y comentar, respectivamente. Tercera, pueden clasificarse los actos comunicativos de acuerdo con la función que cumplen para generar una taxonomía funcional. Cuando te digo “nunca más volveré a beber” hago una promesa, cuando te digo “dame de beber” doy una orden, si digo “me gustaría un poco más de

¹ Con éste término nos referimos a los estudiosos en computational linguistics, computer science y artificial intelligence.

cerveza, por favor” realizo una demanda, si pronuncio “lo que más me gusta es la cerveza” expreso una preferencia, etc. Hay diversas taxonomías que han intentado reducir todas esas funciones a un número finito de funciones básicas. Cuarta, la comunicación exige a los interlocutores poseer ciertos conocimientos y habilidades, sin los cuales la comunicación no tendría lugar.

Además, hay dos proposiciones adicionales en las que el consenso no es tan amplio como el que existe respecto de las cuatro ya listadas. Primera, algunos investigadores proponen que todas las funciones se reducen a dos básicas: regular y declarar. La primera función sirve para modular la acción y la segunda para intercambiar información. Como esos investigadores, nosotros asumiremos que éstas son las funciones básicas. Segunda, se trata de funciones distintas, esto es, cada cual implica procesos psicológicos distintos. Como el consenso no es tan amplio sobre estas dos proposiciones será necesario defenderlas ante algunos problemas.

Respecto de la primera proposición, que las funciones son regular y declarar, habrá que lidiar con una serie de problemas. Primero, como hemos observado con los ejemplos de hace un instante, el lenguaje parece servir para muchas cosas; más de dos, para ser precisos. Habrá que argumentar por qué todas las funciones se pueden reducir a dos. Segundo, otros investigadores proponen que las múltiples funciones pueden reducirse a un número finito de funciones básicas, pero a un número superior a dos. Habrá que argumentar por qué es razonable pensar que son dos y no más las funciones. Tercero, una regulación puede ser también declarativa, lo que pone en cuestión si merece la pena diferenciar funciones. Habrá que argumentar por qué puede seguir manteniéndose que existe una función regulatoria. Cuarto, es posible que todo enunciado sea finalmente regulatorio, es decir, que el fin último de la comunicación sea siempre manipular la conducta de los otros. Habrá que argumentar por qué, incluso asumiendo eso, cabe distinguir una función declarativa. Quinto, hay evidencias de que la gente opera con una dimensión no funcional para diferenciar actos comunicativos, lo que hace dudar de si la distinción funcional tiene realidad psicológica. Habrá que argumentar por qué cabe seguir creyendo en la existencia de funciones.

Respecto de la segunda proposición, que se trata de funciones que implican procesos psicológicos diferentes, será necesario recabar evidencias que así lo demuestren. Nos consta que existe un conjunto suficiente de evidencias, aunque algunas pueden ser cuestionadas. Primero, hay evidencias de que la regulación aparece antes que la declaración en la comunicación infantil. Sin embargo, tal patrón no ha podido ser replicado en un estudio reciente. Segundo, la regulación parece no exigir el dominio de una teoría de la mente, a diferencia de la declaración. No obstante, hay evidencias que ponen en cuestión este supuesto. Tercero, los niños aplican un patrón prosódico distinto para cada función. Aunque sugestivo, tal resultado proviene de un único estudio, lo que hace dudar de su validez. Cuarto, la regulación es predominante en la comunicación infantil. Sobre esto parece no haber duda. Quinto, la regulación es también predominante en el discurso de los primates superiores. Sobre esto tampoco hay objeción. En síntesis, hay por lo menos dos grupos de evidencias bastante sólidos que indican que las funciones son psicológicamente distintas.

En virtud de todo lo expuesto, la conclusión será que es razonable pensar que existen

dos funciones, regular y declarar, cada cual requiriendo procesos distintos. Lo será porque, aunque hay algunos problemas, pueden esgrimirse argumentos para resolverlos y porque, aunque las evidencias no son todo lo robustas que podrían ser, son más que suficientes.

La tabla de contenidos de lo que viene a continuación es ésta. Comenzaremos defendiendo en una primera sección que el lenguaje es un sistema de comunicación (por encima de un sistema de representación). A fin de saber para qué sirve el lenguaje (como sistema de comunicación) intentaremos definir comunicación. Eso será en la segunda sección. De ello surgirán dos supuestos cuyo apoyo será revisado en la tercera sección. Al final haremos un resumen de todo lo expuesto. Lo que habrá de quedarnos claro es que cabe sostener que el lenguaje tiene dos funciones y que cada cual implica procesos distintos.

1. El lenguaje es un sistema de comunicación.

1.1. Una definición de lenguaje.

¿Qué es el lenguaje? O mejor: ¿para qué sirve? No son precisamente preguntas de fácil respuesta. El lenguaje es en realidad varias cosas y parece que toda definición que encontremos excluye alguna de sus facetas. Entonces, ¿qué hacer? Dado que es difícil encontrar una definición exhaustiva, podemos optar por una solución de compromiso. Tomaremos la definición que nos brinda el diccionario de la Real Academia de la Lengua Española. Quizá no sea la mejor opción pero al menos es un punto de partida que nos permitirá ir desgajando las diferentes propiedades del lenguaje. “Lenguaje” según la RAE (23ª edición) es:

- “El conjunto de sonidos articulados con que el hombre manifiesta lo que piensa o siente”

En esta escueta definición cabe distinguir, a nuestro juicio, varios componentes. La primera parte emplea la expresión “conjunto de sonidos articulados”. Dicha expresión parece recoger una primera propiedad del lenguaje: el lenguaje es un *sistema de signos*. Por “sistema” entendemos un grupo de elementos interrelacionados u organizados de una forma particular. En el caso del lenguaje estos elementos son los fonemas, palabras, frases y oraciones. La relación que mantienen todos ellos entre sí está especificada por la gramática: las reglas de la gramática estipulan qué combinaciones entre elementos son posibles y cuáles son sus significados.

Más adelante la definición de la RAE menciona que tal sistema le sirve al hombre para “manifestar lo que piensa o siente”. Un segundo componente del lenguaje podría ser pues la capacidad para hacer manifiesto lo latente, para volver explícito lo implícito: el lenguaje es un *sistema de representación*. Gracias a un sistema de representación, el conocimiento se vuelve explícito, lo cual hace de éste un objeto susceptible de manipulación mental.

Además de lo dicho, de la definición cabe inferir que el hombre manifiesta eso que “piensa o siente” a un interlocutor. O sea, explicitar lo implícito no sólo puede servir para ganar control sobre el conocimiento, sino que explicitar puede servir también para transmitir

el mensaje resultante a otra persona. Si asumimos esta suposición, entonces estamos de acuerdo en que el lenguaje es un *sistema de comunicación* entre personas o interlocutores. Así, los interlocutores comparten pensamientos y sentimientos gracias la lenguaje. Éste sería el tercer componente contenido en la definición: el lenguaje es un sistema de comunicación.

Por supuesto, cabría entresacar más componentes de la lacónica definición de la RAE. Componentes, además, que tal como se expresan en ella parecen tomar partido en algunas discusiones interesantes (e.g., el lenguaje está formado por “sonidos articulados” pero ¿por gestos y expresiones faciales también?; el lenguaje le sirve “al hombre” pero ¿a otros seres vivos (e.g., primates superiores) también?). Ahora bien, para nuestros propósitos es suficiente con lo que ya hemos apuntado: el lenguaje es a la vez un sistema de signos, un sistema de representación y un sistema de comunicación.

1.2. La comunicación como propiedad fundamental.

Hasta aquí hemos visto que el lenguaje es un sistema de signos, un sistema de representación y un sistema de comunicación. Pero ¿son igualmente importantes todas estas propiedades? Podemos empezar descartando la primera. Decir que el lenguaje es un sistema de signos supone concentrarse en las cualidades estructurales del lenguaje mientras que las otras propiedades señaladas hacen referencia a las capacidades que el lenguaje puede tener. Y lo que a nosotros nos interesaba era justamente para qué sirve el lenguaje. (Al fin y al cabo, decir que el “lenguaje” es un “sistema de signos” es como decir que un “martillo” es un “utensilio que consta de una estaca de madera y una pieza metálica en un extremo”; algo mucho menos informativo que “utensilio que se usa para clavar alcayatas y clavos”).

En cuanto a las capacidades señaladas, representación y comunicación, hemos de decantarnos por una de ellas. Para hacerlo nos fijaremos en las ventajas que ofrece cada cual y que pudieron ser claves en algún momento de la evolución humana. La capacidad representativa del lenguaje es de indudable valor (lo que sigue está basado en Donald (1993) y Karmiloff-Smith (1992)). Representar el conocimiento lo hace accesible a la conciencia; eso, a su vez, permite perfeccionar nuestro conocimiento y conectarlo con conocimiento de distintos dominios sin necesidad de obrar directamente sobre el mundo. En una palabra, representar el conocimiento equivale a tener control sobre él. Ello dista dramáticamente de la posición alternativa: cuando no tenemos control sobre el conocimiento. En esos casos somos seres conducidos o controlados por el entorno, igual que los monos. Tal coyuntura se traduce en no poder acceder voluntariamente al conocimiento sino que éste se activa solamente ante las circunstancias precisas. Por supuesto, una posición de no reacción ante el entorno (es decir, la que otorga un sistema de representación del conocimiento) pudo ser crítica en algún momento: para poder perfeccionar lo que uno sabe no es igual poder manipularlo mentalmente al margen de la acción, que depender de los ensayos in situ. De modo que la capacidad representativa del lenguaje es potencialmente útil y quizá es la razón por la cual el lenguaje resultó ser algo adaptativo (tal como sostiene Chomsky, por ejemplo).

Y ¿qué hay de la propiedad comunicativa? La comunicación ofrece ventajas más evidentes si cabe. Como señalan Pinker y Bloom (1990), la comunicación permite (a) obtener información de segunda mano y (b) coordinar la acción de un grupo de individuos

interdependientes. Ello también explicaría por qué el lenguaje supuso una ventaja adaptativa en la evolución. Obtener información de otros es realmente ventajoso: podemos ganar conocimiento útil para la supervivencia sin esfuerzo (el que conlleva aprender a partir de ensayo-error o sofisticados razonamientos) y sin riesgo (el que conlleva cometer un error ante un depredador, por ejemplo) a través de sus palabras. Coordinar la acción, por su parte, es algo imprescindible en un grupo de individuos que cooperan para alcanzar metas complejas e inasequible sin lenguaje. Por lo que respecta a las metas complejas, hay evidencias de que el *Homo erectus* recorrió grandes distancias para cazar, desarrolló utensilios sofisticados y emigró desde África hasta Eurasia (Donald, 1993); empresas, todas ellas, inalcanzables sin la coordinación de los estados mentales de los participantes. En cuanto al papel del lenguaje en esta coordinación, resulta que para sintonizar estados mentales no hay más herramienta que el lenguaje. Werner (1988a) lo plantea así: sin lenguaje no podemos obtener las evidencias necesarias para inferir los estados mentales de los otros y, aunque pudiésemos hacerlo sin las evidencias que el lenguaje aporta, ésta sería una tarea abrumadora (que requeriría complicados y demandantes cómputos). Así pues, la comunicación también pudo ser una capacidad muy valiosa en tanto que posibilitó la transmisión de conocimientos y la organización de la acción conjunta.

Hasta el momento no hemos conseguido más que elogiar las ventajas que pudo brindar cada una de las capacidades del lenguaje, pero decíamos que había que desechar una de las dos. Pinker y Jakendoff (2005) argumentan de la siguiente manera a favor de la comunicación. Si el lenguaje hubiese servido originalmente como sistema de representación, entonces no habría forma de explicar por qué el tracto vocal se transformó para facilitar el habla aún cuando dicha transformación supone un riesgo para la respiración y la deglución; tampoco explicaría por qué el habla sigue un orden lineal. Si la utilidad principal del lenguaje fuese la representación, no necesitaría un sofisticado aparato de modulación, ni seguiría un orden lineal. Nótese que ambas cosas servirían para reducir la ambigüedad y hacer más comprensible el lenguaje al oyente, pero no habría tal necesidad si la representación -y no la comunicación- fuese la función principal. De acuerdo con los investigadores, esto demuestra que, aunque pudiera servir también como sistema de representación del conocimiento, el lenguaje surgió originalmente como sistema de comunicación. Por consiguiente, vamos a considerar la comunicación como la propiedad fundamental del lenguaje (hasta el punto de que consideraremos lenguaje y comunicación sinónimos, salvo que se especifique lo contrario).

En resumidas cuentas, el lenguaje, desde un punto de vista estructural, es un sistema de signos mientras que, atendiendo a sus capacidades, es un sistema de representación y comunicación. De sus capacidades la más importante es la comunicación, en tanto que (a) hay indicios de que fue crítica para la supervivencia en algún momento de la evolución y que (b) el diseño del lenguaje parece responder a tal fin. El lenguaje como sistema de comunicación permite conectar las mentes de dos interlocutores (para obtener información de otros y para coordinar la acción conjunta).

2. Una propuesta para definir la comunicación humana.

2.1. Supuestos.

Una vez argumentado que el lenguaje es -ante todo- un sistema de comunicación queda por saber qué es exactamente comunicarse. Más concretamente, falta por determinar para qué sirve la comunicación, cuáles son sus funciones. Eso es justo lo que vamos a aclarar enseguida proponiendo una definición de comunicación basada en varias proposiciones: en concreto, (a) cuatro supuestos que gozan de (bastante) consenso entre lingüistas, filósofos, psicólogos y científicos computacionales (e.g., Allen & Perrault, 1980; Bach, 1994, 2003; Clark, 1996; Clark & Schaefer, 1989; Dore, 1979; Grosz & Sidner, 1986; Hancher, 1979; Harnish, 2007; Searle, 1969, 1975, 1979; Stiles, 1981; Werner, 1988a, 1988b) y (b) otros tantos supuestos sobre los que no hay el mismo acuerdo. Veamos los supuestos consensuados para después atender a los supuestos no tan consensuados.

2.1.1. Cuatro supuestos. Primero, la comunicación se basa en el uso de *actos comunicativos*. Por acto comunicativo entendemos cualquier acto que despliega un hablante para decir algo a un oyente. Así, actos comunicativos pueden ser cada uno de nuestros comentarios en una conversación:

- (1) Carmen: ¿Podrías pasarme la sal?
- (2) Carol: Ahora mismo.
 Por cierto, la sopa te ha quedado riquísima

Tanto (1) como (2) son actos comunicativos; en este caso, comentarios insertos en una conversación. También el texto que uno escribe (esta tesis, por ejemplo) es un acto comunicativo, en tanto que constituye un acto que se realiza para decir algo a alguien. En general podemos decir que un acto comunicativo es toda contribución a un intercambio comunicativo.

En un sentido más amplio, no sólo los enunciados (verbales) son actos comunicativos (Clark, 1996). Si, a la vista de Carmen, Carol cerrase los ojos y emitiese un “uhmmm” tras haber tomado una cucharada de sopa, también estaría desplegando un acto comunicativo. En tal caso, Carol estaría intentando transmitir vía no verbal su (buenísima) opinión sobre la sopa. Incluso si tras la solicitud de Carmen en (1) Carol, tras haber dado muestras de escuchar perfectamente la pregunta de Carmen, se quedase inmóvil, estaría contribuyendo a la conversación que ambas mantienen. Su negativa a acercar la sal podría significar enfado (“que sepas que estoy enfadada contigo”) porque, pongamos, en un momento antes de (1) las dos interlocutoras habrían estado discutiendo. O sea, los actos comunicativos pueden ser verbales, no verbales (e.g., gestos, expresiones faciales, posición corporal) e incluso acciones. Por eso algunos prefieren hablar de “señales” en lugar de “enunciados” para referirse a los actos que constituyen actos comunicativos (Clark, 1996); nosotros, sin embargo, hablaremos en general de enunciados (aunque con ello nos refferamos a todo el espectro posible de señales).

Hay dos restricciones que acotan la definición de acto comunicativo. Por una parte,

sólo vamos a considerar los actos que se llevan a cabo de forma intencional (tal como hacen Wilson & Sperber (2002)). Quedarán fuera los que se producen de forma accidental. Así, si a partir de mi aparente nerviosismo infieres que lo que estoy diciendo es mentira, siendo mi verdadera intención que me creas, no aceptaremos tal acto (accidental) como comunicativo. Por otra parte, un acto comunicativo no puede considerarse como tal si los interlocutores no lo consideran como tal (Clark & Schaefer, 1989). Cada cual debe cumplir con su parte. De nada sirve que yo señale la sal para que me la acerques si no ves mi dedo señalador. Lo mismo sucede si percibes mi dedo señalador pero no interpretas que estoy pidiéndote algo muy concreto mediante dicho gesto. Así, el hablante debe (a) presentar señales con las que (b) quiere decir algo, mientras que el oyente debe (a) captar las señales y (b) entenderlas como se espera. Sin la participación de ambas partes no cabe hablar de acto comunicativo.

O sea, la comunicación se basa en el uso de actos comunicativos. Un acto comunicativo es (a) todo acto que ejecutamos para decir algo a alguien, (b) que puede estar basado en lenguaje, gestos o acciones, (c) que es deliberadamente producido y (d) que es reconocido como tal por las partes.

El segundo principio es que en un acto comunicativo cabe distinguir al menos dos componentes². Uno de los componentes es *semántico* y, el otro, *pragmático*. El componente semántico es el referente del enunciado o su contenido proposicional (Searle, 1969). En el caso de (1), por ejemplo, el referente sería un escenario hipotético en el que Carol toma la sal y se la acerca a Carmen (o algo por el estilo). Si se prefiere, podría también representarse así (siguiendo el procedimiento de división en *proposiciones atómicas* de van Dijk y Kintsch (1983)): DAR [CAROL, SAL, CARMEN]. En (2) el referente sería quizá la imagen de una sopa y la de cierto comensal satisfecho (o también: ESTAR [SOPA, RICA]). El componente semántico es entonces el referente o estado de cosas descrito por el hablante. Informalmente, puede decirse que el componente semántico se corresponde con *aquello a lo que el hablante se refiere con lo que dice*.

El componente pragmático hace referencia a la intención del hablante: el modo como éste quiere que el enunciado afecte al oyente. Dicho de otro modo, el componente pragmático es la función comunicativa³ que el enunciado cumple. En el caso de (1) la intención o función es pedir: con (1) Carmen está solicitando a Carol que le acerque la sal. Carol, por su parte, habrá de sopesar si de todos los planes de acción posibles en ese momento, el más indicado es el sugerido por Carmen. En la segunda parte de (2) la función es trasladar cierta información al oyente: el hablante hace un comentario sobre la sopa. Al escuchar eso Carmen no deberá elegir entre cursos de acción alternativos, bastará que acepte lo que se le dice y lo integre en su fondo de conocimiento. Así que el componente pragmático es la intención que el hablante tiene con el enunciado o, en una palabra, la función que cumple dicho enunciado. Informalmente, este componente se corresponde con *aquello que el hablante pretende con lo*

2 Como notará el lector, los componentes se corresponden con algunos de los tipos de actos del habla (Austin, 1962; Searle, 1969) que habitualmente se distinguen. De momento, no es necesario presentar la citada distinción.

3 Quizá no todos estén de acuerdo en considerar “intención” y “función” términos sinónimos. Bruner (1975), por ejemplo, plantea que hablar en términos de “intenciones” comunicativas es problemático, entre otras cosas porque, aunque los niños más pequeños se comunican, parece poco apropiado atribuirles intención. Por eso Bruner prefiere hablar de “funciones” comunicativas.

que dice: solicitar o afirmar, como en los ejemplos; pero también rogar, sugerir, cuestionar, etc.

Es de resaltar que existe (cierta) independencia entre los componentes semántico y pragmático. La prueba es que un mismo referente o contenido proposicional (DAR [CAROL, SAL, CARMEN]) puede estar asociado a componentes pragmáticos distintos (Holtgraves & Ashley, 2001; Werner, 1988b):

- (3) Carol dio la sal a Carmen
- (4) Carol, dame la sal

En (3) el hablante expone información: no pretende que el oyente haga nada en concreto, sólo que incorpore esa información que se le transmite. En (4) el objetivo es guiar la conducta del oyente de una determinada manera. Ahora bien, como decíamos, tanto en (3) como en (4) el referente es el mismo.

Tercero, como se hace evidente a la luz de estos ejemplos, existen distintos actos comunicativos en función del tipo de componente pragmático. O mejor dicho, podemos clasificar los actos comunicativos de acuerdo con la función que cumplen. Por ejemplo, no es igual (como en los ejemplos (3) y (4)) comentar que pedir; tampoco es igual amenazar que agradecer, rogar que reflexionar, etc. Dependiendo de su función podemos considerar varias categorías: actos que pretenden dirigir al oyente (pedir, rogar), actos que pretenden fijar un compromiso con uno mismo (amenazar, prometer), actos que pretenden expresar un sentimiento (disculparse, agradecer), actos que afirman algo sobre la realidad (comentar, reflexionar), etc. (e.g., Searle, 1975). Cabe hacer distinciones más sutiles, o sea, considerar subcategorías. Así, algunos criterios adicionales son si el enunciado incluye o no el verbo explícito de la función que cumple (e.g., “te pido que me des la sal” versus “dame la sal”) o si la fuerza de la intención del hablante es mayor o menor (e.g., “¿podrías pasarme la sal?” versus “dame ahora mismo la sal”). Utilizando estos y otros criterios, diversos autores han propuesto taxonomías de funciones comunicativas. Más adelante revisaremos algunas de ellas.

El cuarto principio es que para que un intercambio comunicativo (en la forma descrita arriba) pueda llevarse a cabo son necesarias unas condiciones previas. El hablante debe ser capaz de producir actos potencialmente comunicativos y el oyente debe ser capaz de interpretarlos. Esto último significa que el oyente debe ser capaz de recrear el estado de cosas descrito por el hablante y de reconocer su intención comunicativa. Eso exige la consideración de ciertas pistas, la puesta en marcha de ciertos procesos y la activación de ciertos conocimientos para generar ciertas representaciones mentales. Todo esto será descrito con detenimiento en un apartado posterior, que dedicaremos específicamente a la comprensión del lenguaje.

Sintetizando lo expuesto hasta aquí cabe decir lo siguiente. Primero, la unidad básica de comunicación son los actos comunicativos, que son todos aquellos actos (lingüísticos o no) que se realizan deliberadamente para decir algo y son reconocidos por un oyente. Segundo, es posible distinguir dos componentes en cada acto, uno semántico y otro pragmático (referente

y función, respectivamente). Tercero, hay varios tipos de acto comunicativo de acuerdo con la función que desempeñan. Cuarto, la comunicación no es posible sin que se den unas condiciones previas. Éstos son los principios en los que lingüistas, psicólogos, filósofos y científicos computacionales estarían de acuerdo.

Tabla 1. Los supuestos para una definición de comunicación.

Principio	Descripción
La comunicación se basa en uso de actos comunicativos	Acto comunicativo es todo acto que el hablante despliega para decir algo (verbales, no verbales, acciones) y que es reconocido como tal por las partes
Los actos comunicativos tienen dos componentes: semántico y pragmático	El semántico es el contenido proposicional del enunciado (a qué se refiere con lo que dice) El pragmático es la intención del hablante o función del acto comunicativo (qué pretende con lo que dice)
Hay distintos actos en función de su componente pragmático	De ese modo se distinguen entre actos que hacen afirmaciones sobre la realidad, que comprometen al hablante con lo que ha dicho, que solicitan una acción al oyente, etc. Hay, además, criterios adicionales que permiten subdividir esas categorías
Condiciones previas	Los interlocutores son capaces de producir y comprender enunciados En concreto, comprender requiere poseer ciertas habilidades y conocimientos para generar ciertas representaciones mentales

2.1.2. Dos supuestos adicionales. Nuestra propuesta para definir comunicación, como anunciamos, va algo más allá. Descansa además sobre dos supuestos adicionales. Los vemos a continuación. Supuestos, recordemos, sobre los que el acuerdo no es pleno.

Primero, *hay dos y no más funciones comunicativas que difieren en dos dimensiones:* (a) el objetivo del hablante y (b) la reacción correspondiente del oyente (basado en Givón, 1984; Halliday, 1973; Werner, 1988a, 1988b). A una de las funciones vamos a llamarla *regulatoria*. Los actos regulatorios son aquellos en los que el hablante pretende influir sobre el estado intencional del hablante. Dicho de otro modo, con la regulación el hablante intenta modificar la conducta del interlocutor o, más específicamente, indicar la acción que éste debe acometer y el modo como debe hacerlo. A la otra función la denominaremos *declarativa*. Los actos declarativos son aquellos en los que el hablante pretende influir sobre el estado informacional del hablante, esto es, intenta transmitir información tratando de modificar la visión de cierto aspecto del mundo que quien escucha alberga en su mente. Como receptor de estos actos, al oyente le corresponde llevar a cabo ciertos procesos. En todo caso habrá de captar los componentes semántico y pragmático: para lo primero deberá simular mentalmente la situación o estado de cosas descrito por el hablante y para lo segundo deberá inferir si se le están pidiendo cosas o se le están diciendo cosas. Una vez hecho eso⁴, la tarea que le corresponde es distinta en función del componente pragmático particular: al recibir una regulación el oyente debe sopesar si realizar o no lo que se le pide; al recibir una declaración deberá incorporar la nueva información a su fondo de conocimiento.

⁴ Esta expresión no quiere sugerir que las operaciones que lleva a cabo el oyente se realicen de una forma serial (tal vez lo hagan en paralelo, en cascada...). Si nos expresamos así es con fines expositivos.

Nótese que asumimos que las funciones regulatoria y declarativa son las funciones básicas. Esto supone que todas las que se han venido mencionando (e.g., prometer, expresar, afirmar) pueden reagruparse en torno a nuestra dicotomía. En una sección posterior habremos de defender este supuesto, ya que no es compartido por todos. Al hacerlo nos detendremos en las propuestas de donde nace el supuesto.

Segundo, tanto el hablante como el oyente ponen en funcionamiento *procesos psicológicos distintos dependiendo de la función comunicativa*. Por una parte, el hablante habrá de adecuar sus palabras para un efecto pretendido (regulatorio o declarativo), dados un referente, un oyente y un contexto particulares. Eso exige distintas operaciones mentales en función de si se regula o declara. Por otra parte, el oyente, además de simular mentalmente la situación referida por el hablante (interpretar el componente semántico) e inferir el objetivo que éste persigue (interpretar el componente pragmático): habrá de sopesar si realizar o no lo que se demanda, en el caso de las regulaciones, e incorporar la información nueva a su fondo de conocimiento, en el caso de las declaraciones. Éstos son procesos de naturaleza distinta: no es igual *decidir* si realizar o no lo que se te solicita que *integrar* la información nueva que se te ha trasladado. Más adelante haremos acopio de algunas evidencias que indican que, en efecto, regular y declarar implican procesos distintos.

(Conviene señalar que, si bien este segundo supuesto no es explícito en las propuestas de Givón, Halliday o Werner, puede considerarse implícito en las mismas. Es decir, aunque no lo planteen abiertamente, no costaría trabajo derivarlo de sus afirmaciones. Es Werner el más explícito en este aspecto).

En suma, los dos supuestos para los que el acuerdo es menor son que (a) hay dos y no más funciones comunicativas las cuales difieren en dos dimensiones y que (b) cada función exige tanto a hablante como a oyente realizar procesos específicos; en el caso del oyente éstos son decidir e integrar. En la sección siguiente examinaremos el apoyo con el que cuentan cada uno de estos supuestos. Antes, en el apartado que sigue, repasaremos los prerequisites necesarios para que tenga lugar un proceso comunicativo.

Tabla 2. Supuestos adicionales de la definición de comunicación que proponemos.

Supuestos
Hay dos y no más tipos de funciones comunicativas que difieren en dos dimensiones (intención del hablante y efecto sobre el oyente): Regulación. Manipular conducta del oyente; decidir si satisfacer la demanda Declaración. Transmitir información al oyente; integrar esa información
Producir/comprender cada tipo de comunicación implica procesos específicos; En el caso de la comprensión los procesos específicos son Decidir (en regulación) Integrar (en declaración)

2.2. Prerrequisitos para que exista comunicación.

Estemos o no de acuerdo con la propuesta de definición que acabamos de presentar,

todo modelo de comunicación humana debe dar por sentadas algunas condiciones previas para que la comunicación tenga lugar. Éstas son precisamente las que vamos a revisar a continuación.

Sintéticamente puede decirse que los participantes en la comunicación deben ser capaces de producir y comprender lenguaje. Eso, a su vez, requiere poseer ciertos conocimientos, atender a ciertas pistas y poseer ciertas habilidades para generar ciertas representaciones mentales. Aquí vamos a hablar de los requisitos para comprender lenguaje, sin entrar en lo que supone producirlos (aunque a la luz de las evidencias que revisan Garrod y Pickering (2004) ambos procesos parecen estar estrechamente unidos).

2.2.1. ¿Qué es entender el lenguaje? Las representaciones mentales. Hasta aquí venimos diciendo que entender el lenguaje es entender (a) qué nos dicen (lo semántico) y (b) qué pretenden con lo que nos dicen (lo pragmático). Si se prefiere, venimos asumiendo que comprender es activar mentalmente un referente y una intención comunicativa. Ahora vamos a extender nuestra noción de entendimiento.

Deirdre Wilson (1994) propone que entender es (a) entender qué quiso el hablante decir, (b) qué quiso implicar y (c) cuál era su actitud hacia lo que dijo/implicó. A eso nosotros le agregamos el (d) qué pretendía con lo que dijo. Luego entender o comprender el lenguaje es construir cuatro tipos de representación mental: el significado, las implicaciones, la intención comunicativa y la actitud. Vemos en qué consiste cada una a partir un ejemplo.

- (5) Me encanta que seas tan silenciosa

2.2.1.1. Qué nos dicen. En este ejemplo entender qué nos dicen sería construir una representación del tipo GUSTARME [TÚ, SER, SILENCIOSA] o una escena en la que uno calla y otro disfruta con ese silencio. Esto es, entender qué nos dicen es recrear mentalmente el estado de cosas descrito *literalmente* por el hablante. Si nos hablasen sobre un día en el parque recrearíamos mentalmente todo lo que supone ir de excursión al parque: el aspecto del césped, el mantel y el picnic, el sol acariciando nuestra cara... Si nos hubiesen descrito una anécdota en la playa: tres cuartos de lo mismo (arena, el sonido de las olas, la brisa del mar...). Esta representación, que es una suerte de escenario mental o simulación de la situación referida por el hablante, se conoce con el nombre de *modelo de la situación*⁵ (e.g., Zwaan & Radvansky, 1998). Cabe decir pues que se trata de (a) una representación analógica, en tanto que refleja las condiciones de la situación descrita, (b) multidimensional, dado que incorpora varios tipos de condiciones, y (c) basada en el conocimiento del mundo que poseemos, puesto que para reconstruir mentalmente el estado de cosas descrito debemos aportar más información de la que nos es proporcionada. En cuanto a la naturaleza de estos modelos de situación (¿qué significa analógico?, ¿de qué están hechos los modelos?) anticipamos que es una cuestión objeto de discusión y que en el capítulo 3 plantaremos el debate más extensamente. Por lo que respecta a las dimensiones, de acuerdo con Zwaan y Radvansky, éstas son espacio, tiempo, causalidad, intencionalidad, protagonistas y objetos.

5 Cuando lo que recreamos es una situación cotidiana se utiliza el término modelo de la situación. Más tarde (capítulo 3) emplearemos el término modelo mental para una representación de las mismas características con la excepción de que se refiere a contenidos científicos o alejados de la experiencia.

Así, las simulaciones mentales se desarrollan en un lugar y momento determinados, suceden cosas a causa de ciertos agentes, los personajes llevan a cabo acciones para conseguir cosas y personajes y cosas tienen ciertas características. Obviamente, no siempre se construyen todas las dimensiones o, al menos, no con la misma riqueza. Si decimos que el modelo está basado en el conocimiento del mundo es porque sin éste no podríamos “poner a funcionar” (o “*play the movie*”) la simulación. Donde el hablante dice “ayer fuimos al parque” o “estuve muy a gusto en la playa”, nosotros generamos mentalmente una escena con el sol brillando, la gente riendo, los niños correteando alrededor, etc. (aunque, como advertimos, no siempre se recogen todas las dimensiones). Es decir, como sabemos en qué suele consistir una situación de ese tipo (hemos vivido nuestras experiencias particulares en parques y playas), podemos revivirla recuperando los conocimientos almacenados sobre ella. Ahora bien, necesitamos atender a las palabras del hablante para saber qué conocimientos recuperar. Por eso se dice que el lenguaje actúa como una señal para activar y organizar bloques en un todo coherente (e.g., Gernsbacher, 1997).

2.2.1.2. *Qué nos quieren decir.* La oración ejemplo (5) no sólo describe literalmente un estado de cosas sino que implica algo. En este caso, el hablante quizá quiere implicar justamente lo contrario de lo que dice. Imaginemos que quien habla es un hombre y el objeto del enunciado es una mujer, pareja del primero. Viven juntos y ella sale diariamente pronto a trabajar, mientras él aún duerme. Antes de marcharse, ella necesita buscar su ropa en el armario, para lo cual debe abrir cajones y armarios, colgar y descolgar perchas, abrir y cerrar cremalleras. Vamos, que antes de marcharse ella debe hacer ruido inevitablemente. La cuestión es que el hombre está algo cansado de la poca delicadeza de ella, que parece multiplicar el número de ruidos a propósito (eso le parece a él). En uno de esos momentos en los que tenemos al hombre tratando de dormir y a la mujer rebuscando escandalosamente en los cajones es cuando se produce el enunciado (5). Entonces, aunque lo que el hablante dice es eso, lo que quiere decir o implicar es que “está harto de que ella sea tan escandalosa”. A este fenómeno se refiere Grice (1975) con su noción de implicatura. Luego entender el lenguaje es también entender las implicaturas. Si no las entendemos no cabe decir que se haya entendido el enunciado. Si la mujer se va al trabajo pensando que su pareja ha halagado su habilidad para buscar ropa en silencio, se irá engañada: no es eso lo que se ha querido decir.

Aquí estaría la frontera entre la semántica y la pragmática. La primera se refiere a la representación del significado, que es lo que acabamos de ver. La segunda considera significados pretendidos, intenciones y actitudes, que es lo que veremos a renglón seguido.

2.2.1.3. *Qué pretenden con lo que dicen.* Líneas arriba hemos visto que el lenguaje cumple diferentes funciones. A veces el hablante amenaza al oyente, éste considera medir sus acciones para evitar el cumplimiento de la amenaza; a veces el oyente solicita algo al oyente, éste considera si satisfacer o no la demanda; otras, el hablante sólo hace un comentario, que el oyente asimila sin más. Independientemente de que éstas funciones puedan agruparse en dos, cinco o cualesquiera categorías, lo cierto es que el oyente debe ser capaz de reconocer la intención comunicativa y obrar en consecuencia. En el ejemplo que nos ocupa (5) es muy posible que el hombre no quiera simplemente trasladarle su apreciación a la mujer. Lo que probablemente persigue es modular la conducta de ella para detener el escándalo. (5) es una forma, indirecta en este caso (un acto ilocutivo indirecto, de lo que luego hablaremos, e.g.,

Gibbs (1986), Searle (1975)), de decir “deja de hacer ruido ahora mismo”. Quien lo pronuncia quizá ha pensado que decirlo de forma indirecta es más educado, elegante, inteligible, etc. El caso es que, se exprese directa o indirectamente, el oyente ha de captar que lo que pretenden es guiar su conducta en un sentido concreto. La mujer del ejemplo habrá de sopesar si está dispuesta o no a rebajar el ruido que produce. Luego podrá hacerlo o no, pero debe entender el enunciado como una petición en todo caso. Una vez más, si la mujer se va al trabajo pensando que su pareja hizo un comentario, no podremos afirmar que lo ha entendido: le faltaría por saber qué pretende el hablante con lo que dice.

2.2.1.4. Qué actitud tienen para con lo que dicen. A igualdad de los niveles anteriores, aún cabe una dimensión más de variación. Yo te puedo decir X o Y, directa o indirectamente, para hacer un comentario o para modificar tu conducta, y, además, con convicción, o enfado, o desconfianza, o urgencia, etc. Ésa dimensión es la actitud que uno tiene hacia lo que dice. En el ejemplo (5) no es igual que el hombre lo diga mascullando a que lo diga a pleno pulmón (¡es que me encanta que seas taaaan silenciosa!). En ambos casos la mujer debe entender que se insinúa que está haciendo demasiado ruido y que se le dice con la intención de que deje de hacerlo; pero en el primero su decisión, con toda probabilidad, no estará tan determinada como en el segundo. Si nos dicen “te importa bajar el volumen” nosotros sopesamos si hacerlo o no y, tal vez, decidimos hacer caso omiso. Si nos dicen “¡baja el volumen, por tu vida!” iniciamos el mismo proceso de deliberación (que no tiene por qué ser consciente y complicado sino que puede perfectamente ser automático) y, seguramente, hacemos caso bajando el volumen. Por supuesto, el hablante es consciente de estas probabilidades relativas que dependen de su actitud y, por eso, puede hacer uso de distintos recursos para marcar su actitud y estar más cerca (hacer más probable) del resultado pretendido. Como antes, si la mujer no percibe la actitud no puede decirse que haya entendido: si después del grito encolerizado de su pareja, llega al trabajo comentando que “me dijo si podía hacer menos ruido” es que no ha captado lo que debía (i.e., “me ordenó que bajase el volumen”).

2.2.1.5. Conclusión: qué es comprender. De acuerdo con Wilson (1994) y con lo que acaba de exponerse, entender o comprender el lenguaje es formar cuatro tipos de representación mental. Una sobre el significado, otra sobre lo implicado, otra sobre la intención y otra más sobre la actitud. A partir de ahora cuando hablemos de (a) comprensión del lenguaje, (b) interpretación del lenguaje o (c) comprensión semántica/pragmática del lenguaje nos referimos a estas cuatro comprensiones.

2.2.2. ¿Qué evidencias hay de que existen estas representaciones? A partir del ejemplo (5) hemos ilustrado cuatro tipos de comprensiones (o incomprensiones). Ahora la pregunta es si dicha diferenciación responde a una realidad o es nada más que un artefacto fruto del análisis del psicolingüista.

2.2.2.1. Modelos de situación. Hemos dicho que los oyentes recrean mentalmente la situación descrita por el hablante. Para ello se apoyan en las palabras del hablante y en los conocimientos de los que disponen, los cuales sirven para enriquecer lo contenido en aquéllas.

Una evidencia que suele invocarse para demostrar la existencia de los modelos de

situación es la que sigue. Bransford, Barclay y Franks (1972) realizaron varios experimentos en los que los participantes escuchaban oraciones y más tarde debían reconocerlas entre una lista de candidatos. Había dos clases de oraciones, (6) y (7):

- (6) Tres galápagos descansaban *junto a* un tronco flotante y un pez nadaba bajo ellos
- (7) Tres galápagos descansaban *sobre un* tronco flotante y un pez nadaba bajo ellos

La estructura sintáctica de las oraciones, como se observa, era exactamente la misma. Desde un punto de vista lingüístico, lo único que cambiaba de las oraciones tipo (6) a las tipo (7) era la preposición en cursiva. La clave es que, desde un punto de vista semántico, el modelo de situación suscitado por cada tipo es diferente. En (7) el escenario mental incluye, en un nivel espacial inferior, un pez y, en otro superior, unos galápagos alineados con el pez y un tronco. En (6) el escenario incluye, en orden (espacial) ascendente, un pez, un tronco y unos galápagos. Pero (7) no nos dice exactamente eso (si nos atenemos a las palabras empleadas), es nuestro conocimiento del mundo el que se impone y enriquece la lacónica información ofrecida en la oración.

Los participantes escuchaban un listado de oraciones, algunas de un tipo y otras del otro. Pasados unos minutos los participantes escuchaban un nuevo listado de oraciones. Entre ellas había algunas que habían escuchado anteriormente, la versión contraria de la que habían escuchado u otras completamente distintas. Su tarea era reconocer cuáles pertenecían al listado original. Los resultados mostraron que cuando la presentación original era tipo (7) los participantes reconocían peor. Es decir, las oraciones tipo (7) en efecto inducían la construcción de un pequeño modelo como el descrito antes.

2.2.2.2. *Implicaturas*. Un hablante no sólo dice cosas sino que quiere decir (implica) cosas con lo que dice. Así, cuando el hombre del ejemplo (5) decía “eres silenciosa” quería decir “eres ruidosa”, generando una implicatura en términos de Grice. Se supone que el oyente es capaz de captar estas sutilezas y entender no sólo lo que se dice sino lo que se quiere decir.

Contamos con evidencias de que somos capaces de distinguir esos dos niveles. Lo somos incluso con cuatro años. Cimpian y Markman (2007) implementaron varios experimentos sobre esta cuestión. Reunieron a una muestra de niños de cuatro años y les presentaron enunciados con dos interpretaciones posibles, una general y otra particular. Dilucidar cuál de las dos interpretaciones era la buena requería saber qué quería decir el hablante. A los niños se les presentaba una viñeta con el dibujo de dos osos, por ejemplo. El experimentador les decía entonces a los niños cosas como

- (8) “¿Sabes qué son éstos? Osos, ¿verdad? Déjame contarte algo sobre ellos”
 - (a) pueden oler cosas a larga distancia”
 - (b) están enfermos”

Los enunciados ambiguos son (a) y (b). En (a) la interpretación puede ser tanto genérica como particular, es decir, todos los osos (es una propiedad de la especie) pueden oler cosas que están lejos versus estos dos osos de la viñeta (y no otros) pueden oler cosas que están lejos. En

(b) las interpretaciones posibles son también dos: todos los osos están enfermos versus estos dos. Ahora, en virtud de lo que cualquiera sabe sobre categorías y propiedades, en (a) la interpretación genérica es más plausible y en (b) nos decantamos por la particular. Y ello porque sabemos que las capacidades perceptivas son estables (la mayoría de las veces) mientras que las enfermedades son temporales (la mayoría de las veces).

Tras escuchar la pequeña introducción y uno de los dos enunciados, los niños debían contarle a un peluche “¿quiénes huelen a distancia/están enfermos?”. Su respuesta podía ser genérica o particular. Los resultados mostraron que los niños de cuatro años se decantan sistemáticamente por la interpretación genérica ante enunciados tipo (a) y por la particular ante los de tipo (b). En definitiva, los niños eran capaces de entender qué les quería decir el experimentador.

Una aclaración importante es que no es necesario entender qué nos dicen para después averiguar qué nos quieren decir. O sea, de “qué silenciosa eres” entendemos (cuando es oportuno) “qué ruidosa eres” directamente, sin necesidad de considerar la primera interpretación para después descartarla. Gibbs (1983) presentó peticiones indirectas (e.g., “¿podrías bajar el volumen?” que quiere decir “baja el volumen”) a sus participantes. Tras leerlas, se les presentaban oraciones para que verificasen si era o no una construcción válida en inglés. Algunas no tenían relación con las oraciones originales (las peticiones indirectas), otras eran una paráfrasis de la interpretación literal (“¿eres capaz de bajar el volumen?”) y otras más eran una paráfrasis de la interpretación indirecta (“te digo que bajes el volumen”). Los participantes verificaban más rápidamente las últimas que cualquiera de las otras y no había diferencias entre las dos primeras. Esto significa que, al escuchar un enunciado indirecto, automáticamente entendemos qué se nos quiere decir con él, qué se implica con él; sin embargo, no es necesario considerar qué nos dicen literalmente.

Resultados como éste reclaman que algo que dijimos antes sea puntualizado. Si antes dijimos que comprender es, en realidad, alcanzar cuatro comprensiones, ahora decimos que las dos primeras pueden ser excluyentes. Es decir, en ocasiones entendemos qué nos quieren decir ignorando qué se nos ha dicho literalmente.

2.2.2.3. Intención comunicativa. Algo central en este capítulo es que los oyentes entienden qué pretende un hablante con lo que dice. Eso hace posible reconocer una petición y distinguirla de un comentario, por ejemplo. Lo que nos preguntamos ahora es si, de hecho, los oyentes hacen ese tipo de categorizaciones al procesar el lenguaje.

La respuesta es que sí, a juzgar por datos como los que ofrece Holtgraves (2007; Holtgraves & Ashley, 2001). Por ejemplo, en 2007 este investigador publicó cuatro experimentos explorando el asunto que nos ocupa. En ellos los participantes leían mini-relatos como éste:

- (9) Jenny y Emily han sido amigas desde el colegio. Ahora son compañeras de habitación en una residencia universitaria. Emily es muy olvidadiza. Hoy Jenny está convencida de que Emily ha olvidado (olvidó) que tiene una cita con su dentista.
 - Jenny: “No olvides (olvidé decirte que debías) ir a tu cita de hoy con el dentista”

– Palabra prueba: advertencia

Los segmentos entre paréntesis representan la versión control del relato. En la versión normal (sin paréntesis) los participantes presumiblemente interpretan el comentario de Jenny como un recordatorio o advertencia (que es una regulación). En la control, en cambio, la intervención de Jenny debería entenderse como una afirmación o excusa (que es una declaración). Tras leer uno de los dos tipos de relato, los participantes debían expresar si creían que la palabra prueba se había presentado o no en el texto. Tanto la versión normal como la control contenían casi las mismas palabras (i.e., tienen mismas asociaciones semánticas a priori). Sin embargo, sólo la primera debe interpretarse como una advertencia. Si se interpreta de ese modo entonces la palabra “advertencia” debe rondar la mente del participante (i.e., ganar nivel de activación). Por eso al tener que decidir si dicha palabra estaba o no en el texto se provoca una interferencia (no estaba en el texto pero sí se ha activado en su mente). Los participantes en la versión control no tienen por qué experimentar tal interferencia: simplemente sabrán que la palabra no estaba en el texto. Eso es justamente lo que Holtgraves encontró. A lo largo de cuatro experimentos (y utilizando distintas variables dependientes) quienes leyeron la versión normal activaron la intención del hablante correspondiente. De modo que, sí, reconocemos las intenciones de los hablantes.

2.2.2.4. *Actitudes*. No sólo nos dicen cosas con un propósito sino que lo hacen albergando una actitud particular hacia lo que dicen. En el ejemplo anterior de la pareja (5) ilustrábamos este nivel de comprensión contrastando un grito (“¡no hagas ruido!”) con un tono comedido (“no hagas ruiiido”). Decíamos que si el receptor de este enunciado no era capaz de entender esta diferencia, estrictamente no podría decirse que ha entendido. De nuevo, se trata de ver si la gente reconoce estas actitudes o no.

Brennan y Williams (1995) nos proporcionan pruebas a favor del sí. Los investigadores llevaron a cabo un estudio en dos fases. En la primera unos participantes debían responder a preguntas de cultura general (e.g., “¿Cuál es la capital de Chile?”) mientras eran grabados en audio. Tras responder debían valorar en una escala el grado de éxito que, de acuerdo con sus estimaciones, tendrían en una prueba posterior de reconocimiento (“La capital de Chile es: a. Santiago, b. Valparaíso, c. Valdivia, d. San Fernando”). Los participantes de la segunda fase del estudio escuchaban las grabaciones de las respuestas (sin oír la pregunta) de los primeros y calculaban el grado de seguridad que éstos habían demostrado al responder. Los resultados mostraron una alta correlación entre las estimaciones de los primeros y los segundos. O sea, los que escucharon se basaron en las claves expresivas (i.e., “eer”, “umm”, “...”) de los que respondieron para inferir el grado de seguridad que éstos tenían. Por tanto, hay indicios de que, en efecto, los oyentes se apoyan en ciertas evidencias para reconocer la actitud del hablante para con lo que dice.

2.2.2.5. *Conclusión*. Tomados en conjunto, experimentos como éstos ponen de relieve que los oyentes construyen distintas representaciones mentales a partir de las palabras de un oyente. Hemos comprobado que son capaces de recrear el estado de cosas o situación referida por el hablante, de captar qué quiere implicar un hablante con lo que dice, de reconocer la intención comunicativa de éste y de hacer lo propio con la actitud que alberga hacia lo que dice. De modo que las comprensiones de las que hablábamos (y Wilson (1994) lo suscribía)

son una realidad.

2.2.3. ¿Cómo construimos esas representaciones? Hasta aquí hemos hablado del resultado de la comprensión. Pero falta por determinar cómo los oyentes conforman esas representaciones: qué procesos se ponen en funcionamiento para alcanzar el resultado que ya conocemos. Una vez aclarados los procesos que tienen lugar puede ser interesante plantearse si se producen de forma automática o si, por el contrario, debemos esforzarnos por llevar a cabo complicados cómputos mentales. Acto seguido atendemos a las dos cuestiones.

2.2.3.1. ¿Qué procesos intervienen? Por lo que se refiere a los procesos, para Wilson y Sperber (2002; Wilson, 1994) en la interpretación del lenguaje intervienen de dos clases. Por un lado están los de *decodificación*, que exigen traducir un código con arreglo a un consenso; por otro, los de *inferencia*, que supone optar por el más probable de los significados dadas unas evidencias. Las inferencias, a su vez, necesitan un mecanismo muy particular que las ponga a funcionar.

Aplicar la decodificación es relativamente simple. Hay una señal que se interpreta siempre de cierto modo, dado un acuerdo sobre el código. Por ejemplo, “casa” es un “edificio para habitar” porque así está consensuado. Allí donde aparezca esa secuencia de fonemas (/k/ /a/ /s/ /a/), debemos interpretarla en el sentido apuntado; al menos, quienes compartimos la lengua del castellano. Y, como no hay ambigüedad, lo hacemos sin problema. (Los microprocesos que intervienen en este proceso no serán descritos aquí pero han sido investigados extensamente (e.g., Massaro, 1994)).

La decodificación se extiende también a la interpretación de señales no verbales. Por ejemplo, una sonrisa o el pulgar de la mano extendido se interpretan (al menos dentro de cierta comunidad lingüística) siempre del mismo modo: alegría y acuerdo, respectivamente.

Si bien, no todas las señales no verbales pueden decodificarse. Por ejemplo, si me llevo el dedo índice repetidamente contra mi estómago mientras te digo que “me encuentro mal”, el gesto puede interpretarse como “siento pinchazos en el estómago” pero no con arreglo a un código. En esos casos, computo todas las evidencias disponibles para llegar a la conclusión de que sientes pinchazos ahí.

Pero incluso las señales verbales necesitan algo más que decodificación para ser interpretadas. La palabra “galápagos” tiene cierto significado consensuado, también “tronco” o “pez” pero, como apreciamos con el ejemplo de Bransford et al. (1972), una misma oración admite interpretaciones distintas. O, si se prefiere, una oración deja cierto margen a la interpretación. Si eso ocurre con una oración o con un gesto, imagínense qué sucede cuando consideramos todo un discurso completo con gestos y demás eventos integrados.

Así, la comprensión del lenguaje requiere inevitablemente inferencias. En el ejemplo (5) el hombre decía “me encanta que seas tan silenciosa” para decir “deja de hacer ruido”. Para entender que “silenciosa” quiere decir “ruidosa” deben considerarse las evidencias disponibles. En aquél caso, el escándalo de cajones, perchas y cremalleras sería una evidencia. También el hecho de que los dos vivan juntos y que los dos sepan lo que ocurre día a día

mientras ella se prepara para ir al trabajo. Todas esas son evidencias pertinentes para transformar “silenciosa” en “ruidosa”. Y lo mismo con esta otra transformación: el hombre no dice “te pido que dejes de hacer ruido” sino “estás siendo silenciosa (ruidosa)”. La mujer habrá de considerar lo mucho que le gusta dormir al hombre (algo que sabe porque vive con él y lo conoce bien) y lo mucho que le molesta el ruido (algo que el hombre deja entrever con su tono de voz). Ésas son las evidencias que la mujer habría de contemplar para reconocer la intención comunicativa.

Las evidencias, como puede deducirse de lo que acabamos de decir, provienen tanto de las circunstancias precisas que rodean la comunicación como de la memoria a largo plazo. Así, al hablar de evidencias nos referimos a (a) pistas contextuales y a (b) conocimientos.

Ahora la pregunta es cómo de todas las evidencias posibles, los oyentes seleccionan las oportunas para terminar entendiendo al hablante. En un contexto dado hay cientos de pistas disponibles y en nuestra memoria a largo plazo almacenamos millones de conocimientos. La mujer del ejemplo (5) podría, sobre la base de las evidencias *posibles* haber llegado a una interpretación muy distinta. Por ejemplo, puede observar claramente que el hombre está medio dormido aún (pista contextual), lo que le resta credibilidad al hombre; o puede saber que el hombre tiene problemas de oído (conocimiento almacenado), lo que le impide percatarse del escándalo. Ambas evidencias podrían llevar a la mujer a interpretar que el hombre hace un comentario (y no una petición) sobre su habilidad para no hacer ruido (en lugar de lo contrario). Supongamos que todas las evidencias que hemos mencionado hasta ahora están disponibles. ¿Por qué la mujer considera las pertinentes para interpretar correctamente el enunciado? El mecanismo que proponen Wilson y Sperber es la *relevancia*.

La relevancia es una (a) tendencia natural (resultado de la evolución) a favorecer los inputs del entorno que (b) más efecto cognitivo tienen y (c) menos esfuerzo requieren para ser procesados. Así que nuestro cerebro está *programado* para seleccionar las evidencias oportunas de entre todas las posibles. En el ejemplo (5), el hecho de que el hombre esté medio dormido sólo se considerará si realmente el hombre está así (si su grado de somnolencia es suficiente como para impedirle saber lo que dice) y si da muestras obvias de ello (si no hay que fijarse mucho para advertir que se encuentra en tal estado). Es decir, se tomará en cuenta el atontamiento del hombre si tiene efecto cognitivo y si no precisa gran esfuerzo de procesamiento.

Es importante considerar que estos cálculos no son exclusivos del oyente, también el hablante debe seleccionar las evidencias relevantes. A la hora de producir el enunciado, el hablante debe estimar qué combinación de evidencias del entorno y la memoria más las que él va a aportar conducirán al oyente a la interpretación correcta. O sea, dadas unas circunstancias cuánta evidencia debo aportar para que el otro no me malinterprete. En el ejemplo que venimos desgranando, el hombre puede decir “qué silenciosa eres” para significar “qué ruidosa eres” si asume que la mujer tiene presente en suficiente medida que él odia el ruido y que le despierten temprano. Si no juzga como disponibles esas evidencias él tendrá que aportar más y decir, por ejemplo, “haces demasiado ruido, cosa que no soporto”. En consecuencia, hablante y oyente deben compartir (mentalmente) el set de evidencias pertinentes: deben conocerlas y saber que el otro las conoce (el *common ground*; Clark, 1996).

Por tanto, el mecanismo de relevancia opera también en el hablante.

Un planteamiento similar al de Wilson y Sperber es el de Pickering y Garrod (2004). Ellos proponen el mecanismo de *alineamiento* (*alignment*) para explicar cómo, de todas las evidencias que uno puede considerar para alcanzar una interpretación, elegimos las oportunas.

En resumen, comprender el lenguaje requiere poner en marcha procesos de decodificación y de inferencia. Los primeros sirven cuando hay un consenso que especifica qué significado le corresponde a qué señal. Esto sirve para algunas señales lingüísticas y algunas no verbales. Las inferencias se producen cuando uno extrae un significado de un set de evidencias disponible. Las evidencias, hemos dicho, proceden tanto del entorno como de la memoria. Algunos planteamientos permiten explicar cómo de todas las evidencias posibles uno selecciona las pertinentes para alcanzar la interpretación correcta.

2.2.3.2. *¿Son automáticos o estratégicos esos procesos?* Una pregunta que puede surgir al leer los párrafos precedentes es cómo exactamente se llevan a cabo esos procesos. Una posibilidad es que se realicen sin que nos demos cuenta, sin que nos lo propongamos y sin que nos cueste esfuerzo; esto es, una posibilidad es que sean automáticos. Otra es que requieran conciencia, deliberación y esfuerzo. ¿Cuál es la opción correcta?

La respuesta está a caballo entre las dos posibles: tanto la decodificación como la inferencia pueden operar, dependiendo de las circunstancias, de forma automática o estratégica. No hay más que pensar en cómo habitualmente tiene lugar la comunicación para percatarse de que, por lo general, todo fluye automáticamente. Es decir, ni nos damos cuenta de lo que hacemos (e.g., que hemos considerado la evidencia X para concluir Y), ni nos lo proponemos (sencillamente sucede), ni nos cuesta trabajo. Si conscientemente sopesásemos en cada ocasión decenas de interpretaciones alternativas, podría producirse un estallido computacional. Si siempre hubiésemos de invertir muchos recursos para comunicarnos podría producirse una sobrecarga. Por eso lo común es que los procesos tengan lugar de manera automática. De hecho, los mecanismos de inferencia propuestos por Wilson y Sperber (2002) o Pickering y Garrod (2004) son automáticos (si bien Pickering y Garrod admiten abiertamente que, bajo determinadas circunstancias, puede operar de forma estratégica).

No obstante, cabe también pensar en situaciones en las que dejarse llevar por mecanismos automáticos no basta. ¿Qué pasa si tenemos una discusión con nuestra pareja a cuenta de los ruidos y las horas de sueño y termina diciéndonos: “me voy a casa de mi madre, no me llames”? Ese “no me llames” ¿significa que no le llame esta tarde, en unos días o que hemos terminado y no se me ocurra llamarla jamás? Es probable que un dilema como éste nos asaltase. Es un dilema porque las evidencias más accesibles no agotan la interpretación. En tal caso dedicaríamos consciente y deliberadamente mucho esfuerzo para entender qué nos quisieron decir. Y así reuniríamos muchas más evidencias de las fácilmente accesibles y las combinaríamos para inferir significados alternativos que someteríamos a examen. A veces ocurre así.

Una advertencia de interés para lo que estamos discutiendo es que, para determinar si cierto proceso se da una manera o de la otra, no toda evidencia sirve. En los estudios en los

que se mide la activación on-line de ciertas representaciones mentales puede, con cierta seguridad (al menos si empleamos SOAs), afirmarse que un proceso tiene lugar de manera automática. Pero los estudios que toman medidas off-line pueden resultar algo artificiales. Así, en los experimentos de Holtgraves (2007), que antes comentamos, los procesos tuvieron lugar de forma automática (en el experimento 2 comparó los resultados con SOA de 250 milisegundos y con SOA de 2000 milisegundos encontrando efecto de facilitación sólo con el primero). Pero en los experimentos de Cimpian y Markman (2007) no podemos afirmar algo así. Ellos pedían a los participantes que seleccionasen una entre dos opciones tomándose su tiempo para compararlas. Es posible, aunque no se puede decir taxativamente, que métodos como éste induzcan un procesamiento estratégico. De modo que, aunque sabemos que es posible construir las representaciones a las que antes hemos hecho referencia mediante los procesos de decodificación e inferencia, no estamos en condiciones de determinar si éstos se ejecutaron automática o estratégicamente.

Nos hemos preguntado si la decodificación y la inferencia son procesos automáticos o controlados. Hemos optado por la cautela afirmando que, dependiendo de las circunstancias, pueden operar de un modo u otro.

2.2.4. ¿Qué conocimientos, pistas y habilidades son necesarios? En el subapartado anterior hemos repasado una propuesta que especifica los procesos necesarios para interpretar el lenguaje, decodificación e inferencia. Para llevar a cabo ambos son necesarios, a su vez, ciertos conocimientos, pistas y habilidades. Los describimos a continuación.

2.2.4.1. Conocimientos. Hay ciertos conocimientos que son necesarios para decodificar el lenguaje. Como es de suponer, si uno no sabe que la secuencia de fonemas /r/ /o/ /s/ /a/ designa a una “flor de colores vivos” o que /p/ /a/ /r/ /k/ /e/ significa “espacio público para pasear” no podrá entender qué significa “te llevaste la rosa del parque”. Lo mismo ocurre si uno no sabe que “él” puede referirse a un personaje anteriormente citado: “John es inglés; él es muy gracioso”; aquí no podría entender que “él” sustituye a “John”. Y, si consideramos el tipo de gestos que antes mencionamos, pasa algo idéntico. Debo tener en mi memoria almacenada la asociación entre un pulgar extendido y el significado “Ok”. Estos conocimientos tienen relación con el código: para decodificar necesito conocer cómo se combinan los elementos para producir qué efectos.

Otros conocimientos son necesarios para realizar inferencias. Unos proceden de nuestra experiencia con el mundo. En el ejemplo de Bransford et al. es mi conocimiento del mundo el que me dice que no puede haber un pez sobre un tronco y bajo una tortuga. Es también mi conocimiento del mundo del que deduzco que no es creíble lo que dice un tipo que está medio dormido: sé que en estado de sueño no tenemos control sobre lo que decimos. Otros tantos proceden de nuestra experiencia particular con la persona con la que mantenemos la comunicación. Así, debo saber que mi pareja no tolera el ruido por la mañana para entender que me pide que baje el volumen cuando dice escuetamente “¡ese volumen!”.

2.2.4.2. Pistas. Para interpretar el lenguaje correctamente es preciso considerar ciertas pistas contextuales. Antes ilustramos esto al contrastar un murmullo con un grito. No era igual, decíamos, si nos dicen “baja el volumen” que si nos dicen “¡baja el volumen, por tu

vida!”. Y otro ejemplo. Si, pongamos, en la situación de pareja que venimos desgranando aparece un invitado quizá un murmullo tendría el mismo efecto que un grito. ¿Por qué? Dadas las circunstancias un grito no es factible, porque haría quedar en mal lugar a quien lo pronuncia. Como esa circunstancia es evidente para los dos interlocutores, ambos pueden entender que un murmullo es evidencia suficiente para producir el mismo efecto. Y cabría imaginar muchas otras pistas contextuales posibles y su funcionamiento en otras situaciones hipotéticas.

La disponibilidad de pistas contextuales será materia de discusión en el capítulo 2. En él discutiremos sobre las diferencias entre la modalidad oral y la escrita del lenguaje. El formato prototípico de la modalidad oral es la conversación cara a cara, en donde todas las pistas contextuales son accesibles para los dos interlocutores. Eso dista enormemente de la disponibilidad de pistas con las que cuenta el lector de un texto.

2.2.4.3. *Habilidades*. Decíamos que tanto hablante como oyente deben manejar en sintonía un set de evidencias pertinentes. En otras palabras, deben considerar un conjunto común de pistas y conocimientos para poder formular un mensaje comprensible y para poder comprenderlo. Eso, a su vez, demanda ponerse en el lugar del otro: la teoría de la mente (e.g., Baron-Cohen, Leslie & Frith, 1985).

Pongámonos en el lugar del hombre del ejemplo (5), quien viene acompañándonos desde hace un rato. Recordemos, estamos aún durmiendo en la cama mientras nuestra pareja arma un escándalo considerable. Pretendemos que deje de hacerlo así que le decimos “qué silenciosa eres”. Pretendemos que deje de hacerlo porque consideramos el ruido generado por sus movimientos un auténtico escándalo y, sin embargo, decimos eso y no otra cosa. Desde luego, si aún queriendo lo que queremos utilizamos ese enunciado y no otro (e.g., “¡deja de hacer ruido de una maldita vez!”) es porque estamos seguros de que va a funcionar. Y si estamos seguros porque disponemos de ciertas evidencias que, además, están disponibles para el otro, según estimamos. Las evidencias pueden ser, como decíamos antes, que no toleramos el ruido por la mañana o, pongamos, que nuestra pareja es especialmente descuidada a la hora de no armar revuelo. A la luz de estas evidencias, algo muy probable es que pidamos al otro que rebaje el ruido. Sin las evidencias la cosa no está tan clara. Luego para estar seguros de que vamos a ser comprendidos en el sentido pretendido debemos asumir que el otro es capaz de considerar determinadas evidencias. Es decir, debemos saber (estimar, inferir) qué tiene el otro en la cabeza.

Y, colocándonos ahora en el otro bando (en el de la mujer del ejemplo), para entender el enunciado se precisa también de cierta lectura mental. Si atiendo exclusivamente a las palabras del enunciado (“qué silenciosa eres”) no habría de intentar ser más silencioso: no hay razón para ello porque, según me dicen, ya lo soy. Si atiendo a las evidencias que el otro está también considerando es cuando la cosa comienza a aclararse. Resulta que el otro desea dormir, porque es algo que le encanta; y resulta que no le está gustando nada nuestro comportamiento, porque ha gritado violentamente al dirigirse a nosotros. Luego para interpretar lo que nos dicen debemos ponernos en el lugar del otro también.

La habilidad para inferir los estados mentales de otro se ha dado en llamar teoría de la

mente. Como decimos, una habilidad necesaria para comunicarse es disponer de una teoría de la mente (Sperber y Wilson (2002), de hecho, proponen que la comunicación precisa un submódulo de teoría de la mente). Lo cierto es que hay un buen número de evidencias que ponen de relieve las incapacidades comunicativas de quienes no disponen de teoría de la mente: las personas con autismo.

Por ejemplo, Happé (1993) reunió una muestra de niños la mitad con autismo y la otra mitad normales. Les leyó a todos pequeños textos como éste: “Javier era un chico muy listo y astuto. Era realmente un... ”; junto con un listado de palabras. El listado contenía adjetivos (“zorro”, “ladrillo”, “amable”) que los niños debían utilizar para completar los textos. Los textos eran de dos clases: unos admitían un sinónimo y otros una metáfora (e.g., “Javier era un zorro”). En un experimento la investigadora encontró que los niños con autismo seleccionaban adjetivos apropiados en los textos-sinónimo tan bien como los niños normales; en los textos-metáfora los niños autistas rindieron significativamente peor. En otros dos experimentos ocurrió lo mismo pero con ironía y metáfora. Esto indica que los niños con autismo, quienes no poseen teoría de la mente, tienen dificultades para extraer implicaturas: no entienden bien qué se les quiere decir.

Paul y Cohen (1985) plantearon a un grupo de niños autistas y a otro de niños con retraso mental la siguiente tarea. En una condición los experimentadores interactuaban con los niños y, en un momento dado, les pedían abiertamente llevar a cabo determinada acción (e.g., “Ahora te voy a pedir que colorea varios círculos”). En otra condición la petición se insertaba en medio de la conversación adoptando una forma indirecta (e.g., “¡vaya! aún faltan por colorear unos cuantos círculos”). Se esperaba que los niños entendiesen la instrucción -en cualquiera de sus formatos- y obrasen de acuerdo con ella. Los resultados mostraron que en la condición de petición directa no había grandes diferencias entre niños con autismo y con retraso. Sin embargo, en la condición de petición indirecta hubo diferencias dramáticas: los niños con autismo rindieron mucho peor. Esto sugiere que, aunque los niños con autismo pueden entender el lenguaje literal (tal como indica su rendimiento en la condición abierta), no aciertan a entrever la intención comunicativa de los demás cuando esta está especificada sutilmente.

En suma, comunicarse exige considerar un conjunto de evidencias común entre los interlocutores. Para inferir cuáles son las que maneja nuestro interlocutor debemos poder leer su mente en algún grado. Esa habilidad es la teoría de la mente. Prueba de su importante papel es que quienes no disponen de ella tienen dificultades para comunicarse.

2.2.5. Resumen. Un requisito para la comunicación es que ambos interlocutores sean capaces de producir y comprender lenguaje. Comprender el lenguaje supone conformar una serie de representaciones mentales. Primero, comprender es entender lo que nos dicen, para lo que configuramos un modelo de la situación referida por el hablante. Segundo, comprender es entender lo que nos quieren decir, para lo cual entresacamos las implicaturas de lo que se ha dicho. Tercero, comprender es entender qué pretenden con lo que nos dicen, lo que exige reconocer la intención comunicativa del otro. Cuarto, comprender es también captar la actitud del hablante hacia lo que nos dice, lo que pasa por determinar si lo considera urgente o no, si está convencido de lo que dice o no, etc. Contamos con evidencias empíricas de que tales

representaciones se construyen.

Para construir esas representaciones debemos poner en marcha distintos procesos mentales. Unos son de decodificación, porque extraen los significados de ciertas señales de acuerdo con un código. Otros son de inferencia, porque exigen considerar varias evidencias para derivar un significado. Existe un mecanismo de relevancia que selecciona las evidencias oportunas para derivar el significado adecuado. En cuanto a si los procesos operan de forma automática o estratégica razonamos que, aunque en general lo hacen automáticamente, pueden hacerlo estratégicamente bajo ciertas condiciones.

Comprender supone manejar las evidencias oportunas, hemos dicho. Esas evidencias son conocimientos del mundo y de la persona con la que interactuamos, que están almacenados en la memoria. También son pistas contextuales que están disponibles en la situación que los interlocutores comparten. Como las evidencias que hablante y oyente manejan deben estar en sintonía, otro requisito es tener la capacidad para leer la mente del otro. Por eso quienes no han dominado la teoría de la mente experimentan problemas al comunicarse.

3. Revisando los apoyos a los supuestos de la definición propuesta.

3.1. ¿Hay dos o más funciones comunicativas?

Lo cierto es que antes de aceptar que la comunicación sirve a dos y no más funciones (transmitir información y regular la acción) hemos de afrontar toda una serie de problemas. Primero, la teoría de los actos del habla permite diferenciar muchas más de dos funciones. Segundo, incluso quienes han intentado reducir el número de actos del habla posibles, incluyen en sus taxonomías más de dos categorías. Tercero, cabe identificar regulaciones que son también declarativas. Cuarto, podría argumentarse que en realidad todo acto comunicativo persigue manipular la conducta del oyente. Quinto, una clasificación de funciones podría responder a un proceso de análisis científico pero no a una realidad psicológica. A continuación los revisamos uno a uno y proponemos soluciones para todos ellos. Al final, nos haremos eco de las propuestas en las que nos hemos basado para asumir que hay dos funciones.

3.1.1. Hay muchos actos del habla. Un primer problema es que cabe enumerar no dos sino muchos actos del habla ilocutivos posibles. Antes de proseguir, claro está, se hace necesario explicar qué se entiende por acto del habla; es lo que vamos a hacer ahora.

3.1.1.1. Los actos ilocutivos. En 1962 John Austin se percató de que algunas oraciones no pueden considerarse verdaderas ni falsas. “Te apuesto seis euros a que mañana llueve”, “te pido que vayas allí” o “te recomiendo que no fumes tanto” son ejemplos de esta clase de oraciones. Si nuestro interlocutor no acepta la apuesta, eso no significa que estemos mintiendo; que la acepte tampoco la convierte en verdadera. Lo mismo ocurre con la recomendación de no fumar o la petición de que se desplace a un lugar concreto. Austin distinguió entonces entre oraciones constatativas (*constantives*) y realizativas (*performatives*).

Las primeras son afirmaciones sobre la realidad que pueden ser consideradas verdaderas o falsas (e.g., “J. L. Rodríguez Zapatero es el presidente del gobierno”). Las segundas, por el contrario, son justamente las oraciones que no pueden juzgarse como verdaderas/falsas. Si Austin las llamó realizativas es porque cuando uno pronuncia ese tipo de oraciones lleva a cabo una acción que tiene valor social. Al decir “te pido que vayas allí” le pedimos al oyente que se dirija a determinado lugar; al decir “te recomiendo que no fumes tanto” le hacemos una recomendación, le aconsejamos; al pronunciar “yo os declaro marido y mujer”, el funcionario correspondiente hace efectiva la unión entre los contrayentes. Luego no sólo se *dicen* cosas sino que se *hacen* cosas con las palabras: se pide, se recomienda, se insta una institución; al menos eso ocurre con las oraciones realizativas. Aunque muy sugestivo, este análisis inicial de Austin tenía algunos problemas. Uno es que lleve o no el verbo explícito de la acción, todo enunciado es un acto (“no fumes tanto” es también una recomendación). Dos, la capacidad de realizar cosas con las palabras también se extiende a las oraciones llamadas constatativas: decir que Zapatero es el presidente es también hacer algo, en concreto, hacer una afirmación. En consecuencia, diferenciar entre constatativas y realizativas no parecía demasiado útil.

Más adelante, Austin diferenció tres tipos de actos que se ponen en marcha al hablar y los llamó *speech acts* o *actos del habla*. Esto supuso un avance con respecto a la propuesta anterior. Según Austin, al hablar producimos, para empezar, un *acto locutivo*, que es el acto de decir. Esto se refiere al mismo acto de pronunciar palabras, a decir algo (ya sea “no fumes tanto” o “está lloviendo”). Al hablar producimos también un *acto ilocutivo* (o *fuerza ilocutiva*)⁶, que es ese acto que tiene valor social. Afirmar, pedir, amenazar, prometer, recomendar, declarar marido y mujer, etc. son todos actos con valor social. Por último, existen también *actos perlocutivos*, que son las consecuencias que los ilocutivos tienen en la audiencia. Si al decir “ve allí” la persona que escucha decide desplazarse hasta el lugar indicado, entonces el acto perlocutivo es la acción de desplazamiento. Si decide no moverse de donde se encuentra, también eso se considera un acto perlocutivo. Todo efecto sobre la audiencia lo es⁷. En síntesis, al hablar hacemos cosas como afirmar, recomendar o pedir; estas cosas o acciones son actos ilocutivos, que se diferencian del propio acto de pronunciar palabras (o acto locutivo) y de los efectos que nuestros enunciados puedan tener en la audiencia (o actos perlocutivos). Acciones, palabras y efectos son los tres tipos de acto del habla.

3.1.1.2. *Muchos actos ilocutivos: muchas funciones*. Si cada vez que hablamos llevamos a cabo un acto con valor social y existen muchos actos de ese tipo, ¿cuántos actos ilocutivos cabe distinguir? De momento hemos enumerado unos cuantos (pedir, recomendar, instaurar, afirmar) y pueden encontrarse muchos más: sentenciar, revelar, confesar, criticar, interpretar, reflexionar, aconsejar, sugerir, amenazar, conminar, etc. Lo cierto es que la lista

6 Estos niveles se corresponden con los componentes semántico y pragmático de los que hablamos hace un momento. Así, a un concepto podríamos llamarlo acto locutivo o componente semántico y, al otro, fuerza ilocutiva o componente pragmático.

7 Uno de nuestros criterios para diferenciar funciones era el efecto sobre el estado cognitivo del oyente (intencional o informacional, según el caso) y ahora decimos que, según Austin, todo efecto sobre la audiencia es un acto perlocutivo. Por tanto, uno de nuestros criterios es el acto perlocutivo correspondiente a cada tipo de acto ilocutivo. Es así, pero es preciso matizar. Nos referimos al efecto directo (“me están pidiendo la sal, ¿la paso o no?”) y no a efectos colaterales (que me pidan la sal me enoja, divierte, incomoda, etc.).

puede extenderse casi ilimitadamente. Entonces, ¿por qué seguir manteniendo que hay dos y no más funciones? Pues porque los criterios que antes propusimos permiten discriminar dos grandes grupos en este listado. Recordemos, las funciones comunicativas difieren en el objetivo que persigue el hablante y el modo como le corresponde reaccionar al oyente. El objetivo de un consejo, una amenaza o una sugerencia es siempre⁸ el mismo: guiar la conducta del interlocutor en un sentido concreto (e.g., para que deje de fumar). En todos los casos, el oyente deberá decidir si merece la pena o no seguir nuestra petición. De igual modo, el objetivo de una confesión, una reflexión o una sentencia es siempre trasladar cierta información al interlocutor. Éste habrá de procesar e integrar esa información.

3.1.1.3. Intensidad en la ilocución. Un problema adicional es que pueden distinguirse gradientes dentro de cada categoría. Así, alguien podría pensar que no es igual solicitar que ordenar (e.g., “pásame la sal” versus “dame la sal inmediatamente”) o que afirmar y sugerir son cosas diferentes (e.g., “está lloviendo” versus “parece que llueve”). De nuevo, esto obligaría a aumentar el número de categorías. Searle (1969, 1975), quien profundizó en el trabajo iniciado por Austin, propone varias dimensiones para distinguir actos ilocutivos (después veremos la taxonomía resultante). De hecho, dos de ellas permiten diferenciar gradientes, tal como acabamos de hacer nosotros. Una es el grado de fuerza ilocutiva y la es otra la fuerza de las condiciones de sinceridad. Respecto a la primera, se refiere -como su nombre indica- a la cantidad de fuerza ilocutiva que acompaña a un acto. Así, ordenar es más fuerte que solicitar. La segunda dimensión se refiere al grado de convicción que el hablante tiene en lo que dice. De acuerdo con ella, afirmar demuestra más convicción que sugerir. Fiarnos de estas dimensiones nos llevaría a considerar muchas más categorías de las inicialmente propuestas (regulación/declaración x grado de fuerza ilocutiva x fuerza de las condiciones de sinceridad).

No obstante, hay dos objeciones que alegar. Lo primero es que se trata en realidad de una dimensión única, no dos (de manera que el número de posibles actos se reduce). Como puede apreciarse, la dimensión fuerza de las condiciones de sinceridad sirve sólo para distinguir matices en los actos de corte declarativo. Uno puede creer más o menos en lo que dice, no en lo que pide. Para diferenciar intensidades en los actos regulatorios, la otra dimensión (grado de fuerza ilocutiva) es la que hay que invocar: se puede pedir con más o menos fuerza. Lo segundo es que sea más o menos intensa la fuerza ilocutiva/condición de sinceridad, los criterios originales (i.e., objetivo del hablante/respuesta del oyente) quedan intactos. Se trata pues de dimensiones independientes. Aunque puntualmente puede interesarnos (de hecho, a nosotros nos interesará y ya nos a interesado en el apartado 2.2) diferenciar intensidades, la cuestión es que eso no interfiere en la distinción fundamental regulatorio/declarativo.

3.1.1.4. Conclusión. En suma, aunque es posible enumerar muchos actos ilocutivos, siguen pudiéndose emplear los criterios anteriormente propuestos para distinguir dos grandes tipos entre todos ellos. Aunque pueden diferir en su intensidad, los actos regulatorios y declarativos son las categorías básicas.

⁸ Dentro de un momento veremos que quizá es algo aventurado ser taxativo pero, por ahora, valga la expresión “siempre”.

3.1.2. Las taxonomías de actos ilocutivos disponibles incluyen más de dos categorías. Enseguida revisamos las taxonomías funcionales de varios autores. Tales taxonomías intentan reducir las funciones del lenguaje a un número finito, con arreglo a ciertos criterios.

3.1.2.1. La clasificación de Searle. John Searle, al que hemos citado hace un instante, desarrolló el trabajo de Austin sobre los actos del habla (1969, 1975, 1979). Específicamente, concentró sus esfuerzos en el concepto de acto o fuerza ilocutiva. Una de las cosas que hizo fue proponer una clasificación de actos ilocutivos. Searle diferenció entre cinco actos comunicativos en función de su *punto ilocutivo*. Con punto ilocutivo Searle se refería a la intención *específica* del hablante: si la fuerza ilocutiva se refiere a la intención del hablante, el punto ilocutivo designa la intención raíz del acto comunicativo. Searle propuso también otras dimensiones (e.g., condiciones preparatorias, condiciones de sinceridad) pero resultan poco informativas o redundantes (e.g., el punto ilocutivo “querer que el oyente ejecute cierta acción” tiene la condición de sinceridad de que “el hablante quiere que el oyente ejecute cierta acción”). Veamos cuáles son las funciones comunicativas de Searle, según el criterio más importante, el de punto ilocutivo.

Tabla 3. Las funciones comunicativas según su punto ilocutivo, por Searle.

Punto ilocutivo	Descripción	Ejemplo
<i>Afirmativos</i>	el hablante espera que el oyente incorpore cierta información sobre el mundo	“está lloviendo”
<i>Comprometedores</i>	el hablante se compromete a acometer cierta acción	“mañana me levantaré temprano”
<i>Directivos</i>	el hablante quiere que el oyente ejecute cierta acción	“pásame la sal”
<i>Declarativos</i>	el hablante espera que el estado de cosas que describe se haga efectivo	“el culpable cumplirá dos años y un día de cárcel”
<i>Expresivos</i>	el hablante expresa cierto estado mental	“tengo miedo”

Esto es lo que Searle propuso. Son cinco categorías, no dos, como las nuestras. De manera que la suya no es precisamente un calco de nuestra taxonomía. ¿Pone esto en cuestión nuestra taxonomía? El problema de la clasificación de este autor, a nuestro modo de ver, es que las categorías presentan cierto solapamiento entre sí. En primer lugar, los actos afirmativos transmiten información sobre un estado de cosas en el mundo y se espera, al pronunciarlos, que el oyente se haga cargo de ella; los expresivos, por su parte, transmiten información sobre un estado de cosas en el mundo psicológico del hablante. Sea una situación en el mundo o en el fuero interno del hablante, en ambos casos se expone una información al oyente. Y, por supuesto, al oyente le correspondería hacer cosas semejantes en cada caso. Eso los hace más similares de lo que inicialmente pudiera parecer. Segundo, la diferencia entre los comprometedores y los directivos es también difusa. Con los primeros el oyente espera de sí mismo la ejecución de determinada acción. Con los segundos el agente objeto de esa expectativa es el oyente: deseamos que realice cierta acción. Pero el caso es que en ambos hay un potencial agente del que se espera lleve a cabo una acción. Finalmente, los declarativos

persiguen hacer efectivo un estado de cosas descrito. Aunque no se dirigen expresamente a un agente(s), se espera con los actos declarativos que quien corresponda realice las acciones oportunas para hacer realidad el estado de cosas enunciado. O sea, una vez más se quiere de alguien que haga algo. Y, de nuevo también, a los receptores de enunciados como éstos les correspondería hacer cosas análogas.

A tenor de estos problemas, parece que necesitamos una barrera más nítida entre unas categorías y otras. Afortunadamente, Searle siguió abundando en el asunto. En 1979 propuso nuevos criterios para distinguir los diferentes actos comunicativos, refinando su categorización anterior. De todos los criterios, Searle subrayó la importancia de la dirección de encaje (*direction of fit*). Dirección de encaje se refiere a si las palabras representan o se adecúan al estado de cosas en el mundo (*words-to-world*) o si, por el contrario, es el mundo el que debe cambiar para ajustarse a lo dicho de palabra (*world-to-words*). Con arreglo a esta nueva distinción podemos reordenar y reducir los cinco tipos de acto antes apuntados:

Tabla 4. Las funciones comunicativas según su dirección de encaje, por Searle.

Punto ilocutivo	Dirección de encaje
<i>Afirmativos</i>	Words-to-world porque las palabras reflejan el mundo
<i>Comprometedores</i>	World-to-words porque las acciones futuras del hablante habrán de satisfacer lo que ha sido plasmado verbalmente
<i>Directivos</i>	World-to-words porque las acciones del oyente habrán de adecuarse a la instrucción del hablante
<i>Declarativos</i>	Dirección doble: words-to-world porque se describe un estado de cosas que, aunque no efectivo todavía, lo será en algún momento y también world-to-words porque el mero hecho de describir el nuevo estado de cosas lo instaaura
<i>Expresivos</i>	Dirección nula: porque no se predica sobre el mundo

Searle consiguió así reducir las cinco categorías iniciales. Con el criterio de dirección de encaje se distinguen cuatro y no cinco categorías (*words-to-world*, *world-to-words*, *doble*, *nula*). Claro que, además de *words-to-world* y *world-to-words*, han aparecido sorpresivamente una categoría nula y otra doble. Parece pues que seguimos teniendo que resolver la discrepancia (dos versus cuatro categorías). Una posibilidad es, por un lado, pensar -como hicimos antes- que dar cuenta del mundo no difiere demasiado de hacer eso mismo con el mundo interno del hablante. Puedo decir que “hace un tiempo de perros” o que “no soporto los días lluviosos”, en todo caso no espero más que trasladarte una descripción de un estado de cosas. Luego las palabras se adecuan al mundo (*words-to-world*) -y no hay dirección nula, como plantea Searle. Por otro lado, puede pensarse que la declaración (en el sentido de Searle), más que predicar sobre una situación que se hará efectiva en el futuro, busca lograr que tal cosa tenga lugar: que la realidad hipotética descrita se convierta en la realidad actual. O sea, creemos que la dirección *world-to-words* pesa más que la dirección *words-to-world* en este caso.

Con estos ajustes en la dimensión dirección de encaje, las cinco categorías originales se reducen a dos: los actos afirmativos, comprometedores y directivos quedan igual; los

declarativos y expresivos quedan de la siguiente forma.

- Declarativos, world-to-words porque los agentes correspondientes habrán de adecuar el estado de cosas al mundo descrito de palabra
- Expresivos, words-to-world porque las palabras han de reflejar la situación del mundo interno del hablante

Tras los ajustes es posible distinguir entre un tipo de comunicación que persigue dar cuenta del mundo y trasladar esa información a otra persona (words-to-world) y otro que busca adecuar el mundo al estado de cosas descrito de palabra, requiriendo de ciertos agentes que modifiquen su conducta para hacer esto posible (world-to-words). Retomando nuestra distinción apreciamos que la dirección words-to-world puede corresponder con la función declarativa y la dirección world-to-words con la regulatoria. Ahora sí, cabría decir que la distinción que nosotros trazamos encuentra cierta confirmación en la planteada por Searle (la que distingue direcciones de encaje, con los ajustes pertinentes).

Quizá pueda parecer un encaje de bolillos pero la nuestra no es la única revisión de la taxonomía de Searle. Por ejemplo, Hancher (1979) encontró también algunas carencias en la taxonomía de Searle; claro que su revisión sirvió paradójicamente para ampliar -y no reducir- la mencionada clasificación. Werner (1988b), cuya propuesta estudiaremos luego con más detenimiento, también criticó abiertamente la taxonomía de Searle alegando que las dimensiones que éste proponía eran poco informativas y redundantes. Nosotros ya hemos podido constatar ambas limitaciones. Que son poco informativas lo sabemos porque, como se recordará, el grado de fuerza ilocutiva distinguía el grado con el que una persona enuncia un acto directivo (hay cierta diferencia entre pedir y ordenar); aunque útil para determinados propósitos, esta distinción no es fundamental. Que son redundantes lo hemos sugerido hace un momento: la condición de sinceridad no añade mucho al punto ilocutivo (e.g., el punto ilocutivo “querer que el oyente ejecute cierta acción” tiene la condición de sinceridad de que “el hablante quiere que el oyente ejecute cierta acción”). Werner hará una crítica algo más afinada, como comprobaremos.

Para concluir con Searle, la dirección de encaje (una vez aplicados los ajustes) refuerza la distinción regulatorio/declarativo porque aporta un criterio o dimensión más en la que las funciones difieren. Ésta puede agregarse a las que ya teníamos: objetivo del hablante (modificar conducta del otro o transmitir información, respectivamente), reacción que le corresponde al oyente (sopesar si realizar o no la instrucción o integrar la información, respectivamente) y dirección de encaje (world-to-words o words-to-world, respectivamente).

3.1.2.2. La clasificación de Dore. Otra taxonomía es la propuesta por Dore (1979). A partir de la noción de acto ilocutivo, el investigador configuró una clasificación para categorizar las emisiones de los niños en sus primeros años de vida. Niños cuyas *holofrases* (o enunciados de una sola palabra con significado más amplio, e.g., “tavez” para decir “me gusta, hazlo otra vez”) Dore se encargó de estudiar. En función de la intención comunicativa (es decir, de su fuerza ilocutiva) de las emisiones, Dore distinguió seis tipos de actos comunicativos:

Tabla 5. Las funciones comunicativas según su fuerza ilocutiva, por Dore.

Funciones	Descripción
<i>Peticiones</i>	Solicitan información o una acción determinada
<i>Regulatorias</i>	Controlan el flujo conversacional y el contacto interpersonal
<i>Afirmaciones</i>	Informan sobre hechos
<i>Expresivas,</i>	Manifiestan actitudes
<i>Responsivas</i>	Proporcionan información solicitada
<i>Realizativas</i>	Establecen hechos al ser pronunciadas

El lector avezado habrá notado una clara correspondencia entre la taxonomía de Searle y la que plantea Dore. Las peticiones son para Dore lo que las directivas para Searle; lo mismo con las afirmaciones de aquél y las afirmaciones de éste, con las expresivas de ambos y con las realizativas de Dore y las declarativas de Searle. Sólo faltan por ubicar las regulatorias y responsivas de Dore; lo demás es una cuestión de etiquetas.

Una vez más, debemos revisar la propuesta del autor para resolver la discrepancia (en esta ocasión, seis contra dos categorías). Sin forzar las cosas demasiado, podemos considerar las regulatorias como una forma de control de la conducta del otro. En consecuencia, la diferencia entre las peticiones y las regulaciones de Dore se desdibuja. ¿Dónde está el matiz que las diferencia? Pues para diferenciarlos es necesario apelar a un nuevo criterio o dimensión: el estatus conversacional. La diferencia entre las peticiones y las regulaciones de Dore estriba en que con unas se guía la conducta del oyente en un sentido general mientras que con las otras se controla la conducta del interlocutor que está directamente ligada a la conversación, respectivamente. Así, aunque en los dos casos se demanda algo del oyente, en “dame el muñeco de Nemo” se demanda una acción en relación con un objeto (“cógelo y dámelo”) y en “hola” se demanda una acción en relación con el intercambio comunicativo (“vamos a hablar”). Esta misma dimensión, estatus conversacional, permite interpretar por qué se incluyó la categoría responsiva. Al responder una pregunta lo que hacemos es exponer cierta información; información, eso sí, que nos ha sido solicitada. O sea, la diferencia entre una afirmación como “tengo tres años” y una respuesta (a la pregunta “¿cuántos años tienes?”) como “tengo tres años” radica únicamente en su estatus conversacional pero, como puede apreciarse, no en otro aspecto. Así que un primer problema en la taxonomía de Dore es que, aunque se sostiene que el criterio o dimensión que diferencia unas funciones comunicativas de otras es el tipo de fuerza ilocutiva, sin previo aviso aparece un nuevo criterio -que nosotros hemos- llamado estatus conversacional.

Un segundo problema de Dore es que, tal como sucedía con Searle, hay cierto solapamiento entre las categorías. Aunque las afirmaciones se refieren a hechos y las expresiones a actitudes, no hay razón para negar que ambas sean manifestaciones de cierta información (externa o interna al hablante). En los dos casos, también, el oyente habrá de incorporar esa información que se le ofrece. Y si decíamos que las afirmaciones sólo difieren de las responsivas en su estatus conversacional pero no en su fuerza ilocutiva, entonces afirmaciones y responsivas son lo mismo (y, por extensión, también expresiones y

responsivas). Algo parecido ocurre con las peticiones y las regulaciones, como hemos dicho hace un instante. Sea pedir que comience un intercambio comunicativo o que nos acerquen un muñeco, esperamos del oyente que modifique su conducta para ajustarse a nuestras instrucciones. En ambos casos, de nuevo, la tarea del oyente es equivalente: sopesar si cumplir o no con lo que pedimos. Por lo que respecta a las realizativas, diremos lo mismo que con las declarativas de Searle: es verdad que predicen sobre la realidad (una realidad que se hará efectiva en el futuro) pero lo que hacen básicamente es demandar de los agentes correspondientes que el estado de cosas descrito se materialice -para lo cual los agentes habrán de modificar su conducta.

El cuadro resultante tras la revisión es el mismo que salió del repaso a Searle. Ahora, una vez aclarada la confusión de criterios y el solapamiento entre categorías, tenemos dos categorías básicas. Éstas se corresponden con las que nosotros proponíamos: una función comunicativa que pretende trasladar información al oyente que éste debe procesar e integrar y otra función que pretende modificar la conducta del otro en cierto sentido y ante la cual el oyente deben sopesar si hacer lo que se le pide o no.

3.1.2.3. *La clasificación de Stiles.* Stiles (1981) plantea una taxonomía basada en tres dimensiones que da lugar a ocho categorías de funciones comunicativas. Las dimensiones son (a) fuente de experiencia, (b) marco de referencia y (c) foco. La primera dimensión se refiere a si el enunciado guarda relación con las ideas y experiencias del hablante o las del oyente. Por ejemplo, expresar un sentimiento (“tengo mucho calor”) implica revelar una parte de la experiencia del hablante. Al hacer una pregunta (“¿tú tienes calor?”) la referencia radica en la experiencia del otro (su sensación de calor). La dimensión marco de referencia se corresponde con la perspectiva de la emisión. Una posibilidad es que el enunciado se construya desde la perspectiva del hablante (“estoy disgustado”) y, la otra, que la perspectiva sea compartida por hablante y oyente (“lo pasamos bien anoche”). Por último, la dimensión foco distingue dos modalidades también. En una el foco es el hablante porque se presume cuál es el marco de referencia del otro; en la otra el foco es el oyente porque se asume conocer el marco de referencia del otro. Así, si digo “odio el reggaetón” no asumo nada sobre el marco de referencia del otro; si digo “te gusta el reggaetón” o “debes bailar reggaetón”, en cambio, estoy asumiendo cómo es tu marco de referencia. El cruzamiento total de las tres dimensiones da como resultado ocho categorías, como anunciamos. Éstas son las que aparecen en la tabla 6.

Tabla 6. Las funciones comunicativas según Stiles.

<i>Fuente de experiencia</i>	<i>Marco de referencia</i>	<i>Foco</i>	<i>Funciones resultantes</i>
Habla nte	Habla nte	Habla nte	Revelación
	Habla nte	Oyente	Recomendación
	Oyente	Habla nte	Alecciónamiento
	Oyente	Oyente	Confirmación
Oyente	Habla nte	Habla nte	Interrogación
	Habla nte	Oyente	Interpretación
	Oyente	Habla nte	Reconocimiento
	Oyente	Oyente	Reflexión

Un primer examen de la propuesta de Stiles nos hace percatarnos enseguida de su sistematicidad. Sistematicidad, por cierto, que brillaba por su ausencia en las clasificaciones de Searle y Dore. Aquí hay tres dimensiones con dos niveles cada una que se aplican sistemáticamente para crear ocho categorías (en apariencia) mutuamente excluyentes.

Ahora bien, también hay algunos problemas que merece la pena destacar. El primero y más importante es que las dimensiones distinguen aspectos del contenido no de la función del enunciado. Es decir, rescatando la nomenclatura que empleamos más arriba, las dimensiones diferencian tipos de componentes semánticos pero no de componentes pragmáticos. Por ejemplo, revelación e interrogación difieren únicamente en la dimensión fuente de la experiencia (las dimensiones otras permanecen constantes), lo mismo que confirmación y reflexión. Al confirmar (“sí, asistiré a la cena de Navidad”⁹) expresamos una actitud, algo muy distinto de interrogar (“¿vendrás a la cena de esta noche?”); sin embargo, reflexionar (“las cenas de Navidad te parecen todas igual de aburridas”) y revelar (“esta noche asistirá quien-tu-ya-sabes”) sí son parecidos. La dimensión fuente de la experiencia es entonces insensible a diferentes intenciones comunicativas (exponer información versus demandar información): a igualdad de otras dimensiones, permite que intenciones diferentes coexistan. Con la dimensión marco de referencia pasa tres cuartos de lo mismo. Revelación y aleccionamiento difieren únicamente en esta dimensión; lo mismo para recomendar y confirmar. Al igual que antes, a igualdad de otras dimensiones, revelación y aleccionamiento (“es la primera vez que asisto a una cena de estas características”) sí se parecen; cosa que no sucede con recomendar (“no deberías venir a la cena esta noche”) y confirmar. Si quisiéramos, el mismo problema podría identificarse con la dimensión foco. De modo que la clave del asunto, las funciones comunicativas, es ortogonal a las distinciones de Stiles; siendo más importantes precisiones del tipo: lo que digo me afecta (“no me gustan las cenas de Navidad”) o te afecta a ti (“no te gustan las cenas de Navidad”). Stiles, en una palabra, analiza el contenido (referente) del enunciado pero no su función.

Un segundo problema es que las dimensiones no están suficientemente definidas. Eso hace peligrar la aplicación supuestamente sistemática de los criterios. Por ejemplo, para Stiles la fuente de la experiencia en el reconocimiento (“ajá”, “oh” o “bueno”) es el otro mientras que para la revelación es el propio hablante. Pero en los dos casos el hablante manifiesta una actitud. ¿Acaso no es tan suya la actitud revelada como la reacción expresada mediante el reconocimiento?

El balance de la clasificación de Stiles es el siguiente. El investigador se esfuerza por explicitar de antemano unos criterios y de aplicarlos sistemáticamente, lo cual es un avance con respecto a otras propuestas. El problema es que sus dimensiones tienen más que ver con el contenido o referente de los enunciados que con la función o intención comunicativa que puedan abrigar; además, miradas de cerca las dimensiones parecen no ser tan nítidas como se anunciaba, lo que impide que sean aplicadas rigurosamente.

3.1.2.4. Conclusión. En este subapartado hemos revisado taxonomías que incluyen

9 Todos los ejemplos están inspirados en los aportados por Stiles.

más de dos funciones comunicativas, lo cual difiere notablemente de nuestra dicotomía regulatorio/declarativo. ¿Resultado? Resulta que las categorías de Searle presentaban cierto solapamiento, las de Dore también y las de Stiles tenían que ver con contenido pero no con función: por todo eso creemos que nuestra dicotomía sigue siendo legítima (en tanto que tales problemas no está presentes en ella).

3.1.3. Existen regulaciones que declaran. Un problema algo más serio para nuestro supuesto de que existen dos funciones comunicativas se nos presenta cuando una petición afirma algo. Si, pongamos, estamos estrenando nuestro nuevo coche descapotable, disfrutando de un plácido paseo con un buen amigo, en un momento dado se pone a llover pero estamos tan concentrados en la suavidad del motor de nuestro flamante cabriolet que no colocamos la capota y, entonces, nuestro amigo nos espeta: “oye, está lloviendo a cántaros”; ¿qué pretende con eso nuestro amigo? Seguramente intenta hacernos reaccionar para que coloquemos la capota antes de que quedemos empapados¹⁰. Eso sí, al hacerlo también nos traslada una percepción (“está lloviendo”); percepción que el hablante expresa, además, con cierta precisión (“a cántaros”). Así que, además de pedirnos que coloquemos de una santa vez la capota, nos comunica que, a su entender, llueve con gran intensidad. Del mismo modo, si decimos “hace mucho frío” con la intención de que quien está junto a nosotros ponga en marcha el sistema de calefacción de casa, también le informamos de paso de que la temperatura en la casa es, a nuestro juicio, baja. Estos ejemplos ponen en evidencia que la regulación puede contener declaraciones.

Asumir que una declaración puede confundirse en una regulación implica rechazar de pleno nuestra distinción. Si las dos categorías se solapan, ¿para qué distinguirlas? Lo que podemos hacer es aplicar nuestros criterios a estos dos ejemplos y ver qué ocurre. Dado el contexto que ya describimos, parece que la intención con “está lloviendo a cántaros” es hacer reaccionar al oyente para que ponga en marcha determinada conducta. De modo que el primer criterio, objetivo del hablante, parece identificarse sin demasiada dificultad. Claro que, puede alegarse también que, si ha elegido decir “está lloviendo a cántaros” en lugar de “pon la capota”, será por algo. Quizá porque su intención es doble: comentar y solicitar. Y ¿qué le corresponde entonces al oyente? Debería entender que se trata de una petición (lo que le demandará tomar una decisión: hago o no lo que me solicitan) y, para saber qué es exactamente lo que se le pide (“pon la capota”), debería procesar la instrucción y recrear mentalmente lo que ésta dice (e.g., la imagen de uno mismo colocando la capota del coche). Pero hay dos problemas. Uno, en realidad la demanda o instrucción no se ofrece explícitamente sino que hay que inferirla: nadie dijo “pon la capota”. Dos, resulta que no sólo es que no haya instrucción que entender sino que se nos reclama indirectamente procesar cierta información que habremos de integrar en nuestro fondo de conocimiento: LLOVER [CIELO]; MODO: A-CÁNTAROS. Si este análisis es correcto, entonces, puesto que la intención del hablante es manipular la conducta de su interlocutor y hacer un comentario, la reacción correspondiente del oyente es también mixta: debe sopesar si cumplir o no la demanda y debe integrar información nueva. La única solución es esgrimir que el volumen de información nueva que uno ha de incorporar es tan pequeño que esta tarea resulta insignificante. Sobre todo teniendo en cuenta que lo más urgente es atender la demanda:

10 Como sabemos, este fenómeno ha dado en llamarse acto del habla indirecto (Gibbs, 1986; Searle, 1975): un enunciado que cumple varias funciones (i.e., ejecutan múltiples actos ilocutivos).

“¿pones o no la capota de una vez?”. Es decir, para el oyente la regulación prima sobre la declaración.

Se desprende del razonamiento anterior la siguiente regla. Dada una intención del hablante mixta (e.g., pedir+comentar), si podemos distinguir una tarea del oyente principal de otra secundaria, entonces optaremos por la principal. En los ejemplos, es verdad que además de sopesar si seguir o no la instrucción (poner capota, encender calefacción), el oyente debe integrar una pequeña pieza de información (llueve a cántaros, la temperatura es muy baja). No obstante, (a) el volumen de información a incorporar es mínimo y (b) la tarea de deliberación es acuciante. Si aceptamos la regla, entonces podemos encajar que existan regulaciones que declaren pero sin dejar de ser regulaciones.

3.1.4. Todo acto comunicativo persigue regular. Otra posible crítica a la dicotomía regulación/declaración es que, a la postre, todo acto comunicativo pretende manipular la conducta de los demás. Imaginemos que nos encontramos una vecina en el portal y nos dice algo como esto:

- (10) “El ascensor se ha estropeado de nuevo. La verdad es que no parece un aparato de gran calidad: lo pusimos hace tan sólo un mes y el técnico ha tenido que reajustarlo ya tres veces. Para mí es especialmente molesto porque justo ayer me torcí un tobillo y tengo molestias al andar. Y es más molesto aún cuando tengo que subir la compra a casa. Además, hoy he comprado patatas, cebollas, peras y manzanas... En total unos seis kilos... Confieso que no sé ni cómo voy a hacerlo. Bueno, sólo es hasta el primer piso...”

A estas alturas todos hemos entendido que lo que quiere esta mujer es pedirnos que carguemos con su compra y le ayudemos a subir por la escalera de paso. En este sentido, su mensaje podría resumirse a “coge mi compra y ayúdame a subir al primero”, lo cual es, a todas luces, una regulación. Pero el caso es que nos ha dado mucha información entre tanto. Información que hemos de procesar e incorporar a nuestro banco de conocimientos... Y no es un ejemplo anecdótico. La vecina nos habla sobre los problemas técnicos del ascensor y sus lesiones para que le ayudemos, un amigo nos comenta sus problemas económicos para que le prestemos dinero, un compañero de trabajo recién llegado nos habla de su dilatada experiencia y amplio reconocimiento para que al dirigirnos a él lo hagamos con respeto, y así sucesivamente. En una palabra, al margen de su apariencia, todo son regulaciones.

Para resolver este problema puede ser de ayuda la regla que aplicamos antes. Comprobémoslo. Aquí hay el objetivo *de fondo* del hablante es fijo: pretende que llevemos a cabo determinada acción. Ahora bien, ¿cuál es la tarea del oyente? Una vez procesado todo el discurso correspondiente, el oyente habrá de reparar en que le están pidiendo algo y, por tanto, sopesar si satisfacer o no la demanda. Pero fijémonos bien: hemos dicho “una vez procesado todo el discurso” podremos hacer todo eso. Esto significa que, aunque haya una tarea final, hay otra que demanda nuestra intervención inmediata. Hay un (a) considerable volumen de información que (b) reclama ser procesado antes de nada. De modo que lo primordial aquí es procesar e integrar nueva información a nuestro fondo de conocimiento, independientemente de que haya algo que hacer más tarde. Una vez más, si aceptamos esto, entonces nuestra dicotomía regulatorio/declarativo sigue en pie. Por si alguien no está satisfecho, debemos

decir que en breve presentaremos un análisis similar que propone Givón (1984) y un criterio idéntico que propone Werner (1988a, 1988b).

3.1.5. La realidad psicológica de las funciones. En un estudio muy reciente, Holtgraves (2004) encontró que, cuando se pide a la gente que agrupe con libre criterio una muestra de enunciados, lo hacen basándose en la valoración emocional que suscita cada cual. De ese modo, las categorías con las que implícitamente se agrupan actos comunicativos son: afecto neutro, afecto positivo, afecto negativo. Los oyentes, entonces, no distinguen entre regulación y declaración sino que emplean otras categorías basadas una dimensión completamente distinta a la funcional. Analicemos el estudio y sus consecuencias con más calma.

Holtgraves elaboró una muestra de escenarios hipotéticos. Los escenarios consistían en pequeñas descripciones de situaciones cotidianas que implicaban a dos personajes. En un momento dado, uno de ellos le hacía un comentario al otro; tal comentario consistía en el enunciado crítico que los participantes del estudio habrían de evaluar.

- (11) Jenny y Emily han sido amigas desde el colegio. Ahora son compañeras de habitación en una residencia universitaria. Emily es muy olvidadiza. Hoy Jenny está convencida de que Emily ha olvidado que tiene una cita con su dentista.
 - Jenny: “No olvides ir a tu cita de hoy con el dentista.” [enunciado crítico]

En una primera parte del estudio, los participantes debían leer escenarios como éste y después formar grupos de enunciados críticos siguiendo un criterio libre. Así alguien podría considerar, por ejemplo, que “No olvides ir a tu cita de hoy con el dentista” es, por alguna razón implícita, muy parecido a “Ten cuidado al pisar, que está recién fregado” pero muy diferente de “No soporto a Pepe Blanco”. Una vez hecha la tarea, Holtgraves sometió los datos resultantes a un análisis jerárquico de cluster. Ello sirvió para identificar los grupos subyacentes a las agrupaciones individuales de cada participante.

Los grupos resultantes fueron los siguientes. El primer cluster incluyó: suponer, asentir, disculparse, preguntar, reafirmar, prevenir, recordar, corregir, presentar. El segundo cluster estuvo compuesto por éstos tipos de acto: fanfarronear, adular, invitar, agradecer, animar, ofrecer, felicitar. El último cluster fue: rechazar, disculpar, suplicar, demandar, quejarse, amenazar, culpar, prometer.

Un simple vistazo a los cluster resultantes permite apreciar que el segundo grupo contiene actos comunicativos que un oyente potencial recibiría de buen grado (o que un hablante potencial no tendría problema en desplegar). A todos nos gusta que nos inviten, nos agradezcan, nos adulen o nos feliciten. El mismo vistazo permite observar que el tercer grupo, en cambio, parece reflejar el panorama opuesto: un oyente potencial no disfrutaría al ser amenazado, culpado o al ser objeto de una queja; más bien sería algo desagradable (y algo parecido le pasaría al hablante). El primer cluster presenta un panorama más neutro. No hay nada demasiado malo ni demasiado bueno en escuchar a alguien recordarnos algo, reafirmarse en una creencia o mostrar su actitud sobre un tema. Así pues, parece que las categorías recogen la experiencia emocional asociada a cada tipo de acto; algo muy alejado de su

función comunicativa. De hecho, las regulaciones y declaraciones parecen repartirse por igual dentro de los cluster (e.g., en el tercer cluster hay tantas regulaciones como declaraciones): las dimensiones son independientes.

Una posible crítica a esta agrupación en cluster es que se basa en un procedimiento estadístico, no en la verdadera percepción de los participantes. Es decir, el procedimiento establece grupos a la fuerza calculando las inter-correlaciones entre todas las categorías: debe encontrar un orden, lo haya o no. Por eso cabe la posibilidad de que la agrupación en cluster sea un mero artificio que, casualmente, casa razonablemente bien con una categoría que existe (valoración emocional).

Para resolver este problema Holtgraves desarrolló una fase más en su estudio. Ordenó al azar los escenarios (enunciado crítico incluido) y pidió a unos nuevos participantes que los valorasen con arreglo a dos escalas: “cómo crees que se sintió el hablante (1 extremadamente mal – 7 extremadamente bien)” y “cómo crees que se sintió el oyente (1 extremadamente mal – 7 extremadamente bien)”. Los resultados fueron más que nítidos: la media del primer cluster fue de 4 en las dos escalas, 6 la del segundo cluster y 2.50 la del tercero. Esto es, el primer cluster queda en el centro de la escala, el segundo muy cerca del polo positivo, el tercero próximo al polo negativo. Por supuesto, las diferencias fueron estadísticamente significativas. Así Holtgraves confirmó que la gente tiende a categorizar los actos comunicativos en función de la carga emocional que poseen; tanto desde la perspectiva del hablante como la del oyente.

La cuestión es cómo eso afecta a nuestra categorización particular, la que distingue actos comunicativos de acuerdo con la función que cumplen -y no con su resonancia emocional. Podemos esgrimir varios argumentos para seguir sosteniendo nuestra categorización. En primer lugar, lo que se pondría en cuestión con este estudio no es que nuestra categorización de funciones sea poco acertada, sino que *toda* taxonomía de funciones carece de realidad psicológica. Parece que lo que prevalece es la experiencia emocional asociada a la comunicación. Así que esto sería un problema para toda taxonomía. El segundo argumento sirve para salvar la legitimidad de diferenciar funciones.

En segundo lugar, el hecho de que la gente agrupe actos en torno a su resonancia emocional no implica necesariamente que la categoría función no tenga realidad psicológica. Lo que puede haber ocurrido es que una dimensión más básica haya ensombrecido la intervención de la otra: es posible que los participantes mostrasen un *sesgo* (o tendencia poderosa) hacia una evaluación emocional. De hecho, la *valencia emocional* (i.e., me gusta/me disgusta) es una reacción básica que regula la respuesta (al menos inmediata) que damos a los estímulos del entorno (e.g., Lang, 1994). Es el mecanismo que nos brinda una primera evaluación de un objeto permitiéndonos diferenciar, en un instante, lo potencialmente beneficioso de lo potencialmente perjudicial. Algo que, sin duda, ha sido moldeado por la evolución (e.g., no en vano los humanos contamos con circuitos neuronales específicos para cada polo: unas áreas cerebrales se activan ante un estímulo negativo, otras ante uno positivo, Lane, Reiman, Bradley, Lang, Ahern, Davidson, & Schwartz, 1997) y que seguramente es compartido por muchos animales. Si asumimos que la valencia emocional es un mecanismo más básico que la distinción entre funciones comunicativas (porque es una reacción visceral clave para la supervivencia), entonces cabe argumentar que lo que sucedió en el estudio de

Holtgraves fue que la valencia emocional sesgó la agrupación/evaluación de sus participantes, sin que ello ponga en tela de juicio la realidad psicológica de una clasificación funcional.

Una manera de probar esta hipótesis sería lograr despejar (de algún modo) el efecto de la valencia emocional y verificar si los cluster resultantes son los mismos. Para ello una posibilidad es eliminar del corpus de enunciados críticos de Holtgraves los (más) negativos y los (más) positivos. De esa manera el corpus sería neutro en cuanto a valencia emocional se refiere. Eso dejaría la cancha libre para explorar el modo como la gente categoriza actos por su valor comunicativo, si es que tal dimensión existe. De momento, sólo podemos decir que el estudio no nos beneficia, aunque tampoco nos perjudica.

En tercer lugar, sabemos gracias a la investigación del propio Holtgraves (2007; Holtgraves & Ashley, 2001) que la gente reconoce la fuerza ilocutiva que se esconde tras los enunciados verbales. Esto lo revisamos en el apartado sobre comprensión del lenguaje. Recordemos que en sus experimentos los participantes leían pequeños relatos en los que un personaje le decía algo al otro. Los participantes verificaban gramaticalmente oraciones que recogían o no el acto ilocutivo detrás de los enunciados y lo hacían mucho más rápidamente cuando sí lo hacían. Lo mismo ocurrió cuando se manipuló el SOA para que fuese de 250 milisegundos. Concluimos entonces que, no sólo es que reconozcamos intenciones comunicativas sino que lo hacemos automáticamente.

Cuarto, más tarde veremos que hay evidencias de una diferenciación funcional en la producción lingüística. En concreto, lo que revisaremos son trabajos en los que se constata que los niños regulan mucho más de lo que declaran en su desarrollo comunicativo; y que en la comunicación de los primates a los que se enseña a hablar (e.g., lengua de signos) predomina también la regulación. A tenor de estas evidencias parece que cada función evoluciona de forma distinta, luego *hay* funciones.

En resumen, hay indicios de que la gente organiza los actos comunicativos de acuerdo con su resonancia emocional, no con arreglo a su función. Sin embargo, eso no significa que las funciones no tengan realidad psicológica: lo que probablemente ocurre es que la valencia emocional es una dimensión mucho más básica. Más importante aún, contamos también con evidencias de que sí captamos las intenciones comunicativas (y de forma automática, además) y de que hay funciones diferenciadas.

3.1.6. No todo son problemas: los fundamentos de la taxonomía. Hemos visto que hay muchos actos ilocutivos, que además difieren en su intensidad; también que hay varios investigadores (Searle, Hancher, Dore, Stiles) que han propuesto taxonomías sobre funciones comunicativas que distan considerablemente de la nuestra; hemos puesto sobre la mesa que regulación y declaración pueden solaparse (en tanto que hay regulaciones que declaran y que hasta las declaraciones buscan, en último término, regular); por último, hemos contemplado la posibilidad de que una clasificación funcional carezca de realidad psicológica. Para defendernos de estos problemas hemos esgrimido los correspondientes argumentos. Así, el balance hasta el momento es positivo. Al menos lo es si el lector ha aceptado nuestros argumentos. Lo que vamos a hacer ahora es revisar las propuestas de otros investigadores que, esta vez sí, coinciden con nosotros al señalar las mismas funciones comunicativas: de hecho,

son la base sobre la que propusimos el supuesto de que existen dos funciones.

3.1.6.1. La clasificación de Givón. Talmy Givón (1984) propone que existen dos tipos de actos ilocutivos. Éstos son los *informativos* y los *manipulativos*. Givón dice literalmente que los actos informativos son los que “transfieren información a otros” mientras que los manipulativos son aquellos que “manipulan la conducta de otros” (p. 496). Sin lugar a dudas, el primer tipo equivale a la función declarativa que hemos descrito aquí mientras que el segundo se corresponde con la regulatoria.

Givón, además, facilita un análisis de las condiciones o parámetros contextuales que propician el uso de una función y otra. En realidad, lo que interesa a Givón es cubrirse las espaldas sobre un supuesto tácito sobre el que basa otras consideraciones. El supuesto es que el lenguaje humano adulto es, casi exclusivamente, informativo (mientras que el discurso infantil o el de los primates superiores sería manipulativo en gran medida). Al percatarse de que no ha dado justificación para tal asunción lo que hace Givón es ofrecer un análisis de las condiciones que pueden propiciar que el uso de funciones informativas o manipulativas sea preferente. A nosotros tal análisis nos servirá para restringir aún más la definición de uno u otro tipo de función.

El análisis contempla, entre otros, los parámetros siguientes: longitud del discurso (número de proposiciones o unidades de significado), rol del contexto deíctico (grado en que el contexto perceptivo es compartido y/o tematizado en el discurso), cambio cultural (estabilidad del cuerpo de conocimiento que se considera compartido por los miembros del grupo cultural) y complejidad social (rigidez con que están estipulados los papeles de cada cual). Como hemos dicho, dependiendo de cómo se configuren estos parámetros, cabe esperar una comunicación más basada en la manipulación o, alternativamente, más basada en la información. Con arreglo a los parámetros citados es posible caracterizar las situaciones comunicativas en las que habitualmente estarían inmersos niños humanos y primates superiores, por un lado, y adultos humanos, por otro.

- Longitud. El discurso es mono-proposicional en el caso de los niños y primates y multi-proposicional en el de los humanos adultos.
- Rol del contexto deíctico. El contexto perceptivo es compartido y es el tema objeto del discurso en el caso de los niños y primates. Con los humanos adultos ni es compartido y ni es el tema objeto de discurso: los temas son ajenos al contexto inmediato.
- Cambio cultural. Para los niños y primates el cuerpo de conocimiento que se presume compartido es estable mientras que para los adultos todo intercambio comunicativo requiere la negociación y actualización del banco de conocimiento que supondrá la base de la comunicación.
- Complejidad social. El papel de cada actor se conoce bien de antemano, en el caso de los niños y primates. Cada cual tiene un estatus, motivación y comportamiento predecibles. En el caso de los humanos adultos, por el contrario, los interlocutores son extraños de manera que su estatus, motivación y comportamiento son desconocidos por las partes.

En una palabra, dice Givón, la comunicación entre niños/primates es propia de una *sociedad*

de íntimos y la de los humanos adultos es la de una *sociedad de extraños*. O dicho de otro modo, en la comunicación infantil/primate se emplea la manipulación, casi en exclusiva¹¹; en la comunicación adulta, la información es dominante.

Tabla 7. Parámetros que determinan el uso preferente de regulación o declaración.

Parámetros	Predominancia regulatoria	Predominancia declarativa
<i>Longitud del discurso</i>	Mono-proposicional	Multi-proposicional
<i>Contexto deíctico</i>	Objeto del discurso	El tópico está fuera
<i>Complejidad cultural</i>	Mínima, base de conocimientos estable	Máxima, base de conocimientos cambiante
<i>Roles sociales</i>	Conocidos	Desconocidos
<i>Tipo de sociedad resultante</i>	De íntimos	De extraños

El lector recordará que nos enfrentamos a dos problemas que ponían en tela de juicio la distinción regulatorio/declarativo. Uno era que las regulaciones a veces declaran (“hace muchísimo frío”) y el otro que todo acto comunicativo, incluso declarativo, persigue en último término modular la conducta de los otros. Para despejar los problemas propusimos una regla que matizaba el uso de los criterios propuestos anteriormente: para un objetivo del hablante dado, hemos de concentrarnos en la tarea principal que ha de llevar a cabo el oyente. Así, aunque deba integrar información nueva, si ésta es mínima y la demanda que se le ha hecho requiere una decisión urgentemente, la regulación que declara es regulación; de igual modo, aunque finalmente deba decidir si satisfacer o no una demanda que se le está planteando, si el oyente ha de procesar e incorporar un volumen considerable de información antes de decidir, la declaración que esconde una regulación es declaración. La regla se ve apoyada ahora por el análisis de Givón. Al realizar una petición indirecta que adopta la forma de declaración (“hace frío”), estamos enunciando un (b) discurso mono-proposicional, (b) que versa estrictamente sobre el aquí y ahora y (c) que implica a interlocutores cuyos conocimientos son compartidos (al menos los necesarios para mantener la comunicación en cuestión) y (d) cuyos roles (los que desempeñan en esas circunstancias concretas) son conocidos. Por tanto, el acto es manipulativo. En cambio, al hacer que cierto volumen de información sea compartido para poder después regular, estamos enunciando un (a) discurso multi-proposicional, (b) que no versa sobre el aquí y ahora y (c) que implica a interlocutores que no comparten todos sus conocimientos (d) ni conocen los roles del otro. Por tanto, el acto es informativo.

El análisis del lingüista constituye una aportación a la distinción declarativo/regulatorio que venimos planteando. Además de los criterios que fueron propuestos (las funciones declarativa y regulatoria difieren en dos dimensiones: objetivo del hablante y la respuesta que le corresponde al oyente), contamos ahora con un análisis de las condiciones contextuales que propician un tipo de comunicación u otro.

¹¹ Más tarde proporcionaremos evidencias que avalan esta suposición.

3.1.6.2. *La clasificación de Werner.* El lingüista (en este caso computacional) Eric Werner propone también una taxonomía de dos funciones comunicativas (1988a, 1988b). Para diferenciarlas, Werner defiende emplear un nuevo criterio. En lugar de concentrarse en el punto ilocutivo (tal como hicieron Searle y Dore), Werner sugiere considerar los efectos que cierto acto comunicativo tiene sobre estado cognitivo del oyente. De este modo, los efectos que distingue son dos: (a) impacto sobre el *estado intencional* del oyente e (b) impacto sobre el *estado informacional* del oyente. Los actos que afectan al estado intencional del oyente son aquellos que reclaman decantarse por cierto curso de acción, entre todos los posibles. Movilizan al oyente en una dirección y por eso Werner dice que afectan a su estado intencional (“¿qué voy a hacer?”). Otros actos afectan al estado informacional porque añaden información al fondo de conocimiento del oyente. Antes no sabía tal cosa, ahora ya lo sabe; por eso Werner dice que el estado del banco de información del oyente ha sido alterado.

Tabla 8. Las funciones comunicativas según su efecto en el oyente, por Werner.

Función	Efecto sobre el estado cognitivo del oyente
<i>Informativa</i>	Afecta al estado informacional: antes no lo sabía, ahora ya lo sé
<i>Directiva</i>	Afecta al estado intencional: debo decidir si hago X o no

Werner sostiene que dar prioridad a esta dimensión representa un avance con respecto a la dimensión punto ilocutivo de Searle. De acuerdo con Werner, (a) el objetivo del hablante es independiente de la interpretación que un oyente finalmente hace de un enunciado y (b) lo que cuenta es esta interpretación (recuérdese, un acto comunicativo no era tal si el oyente no lo captaba y reconocía, al margen de lo que haga el hablante). Un ejemplo que proporciona Werner es el que sigue. Si un presidente declara una guerra, el general habrá de enviar a sus tropas a la guerra le guste o no. De manera que, aunque su intención no sea manipular a su audiencia en cierto sentido, sus palabras (e.g., “soldados, vais a ir a la guerra”) redundarán en tal manipulación -dice Werner. Así, en ausencia de intención del hablante, sigue habiendo efecto en el oyente.

Un análisis de estas características supone cierta divergencia con lo que nosotros venimos asumiendo hasta ahora. Aunque, como Werner, contemplamos la dimensión efecto sobre el oyente también, mantenemos de todos modos la dimensión intención del hablante. Y lo hacemos por varias razones. Una, estamos interesados en la comunicación intencional. El hablante despliega un acto para decir algo a un oyente. Sin la percepción y reconocimiento del oyente de tal acto no habrá acto propiamente dicho. Mas tampoco sin la participación del hablante: el hablante *quiere* decir. No vamos a considerar los actos accidentales, en los que interviene el azar y no la deliberación. Dos, la discusión sobre la intención que plantea Werner es más espinosa de lo que parece. Werner ilustra el problema con el ejemplo del general: el general no quiere ver a sus hombres combatir pero los envía a combatir mediante sus palabras. Aquí hay varias capas de intención. Su motivación no es la guerra (porque queda mucho trecho en la vía diplomática, porque ha visto a demasiados compañeros morir, etc.), es verdad; pero su intención comunicativa *sí* es enviar a la guerra a sus hombres. Aunque le disguste. Si no pretendiese obtener eso con sus palabras entonces diría “no vayáis a la guerra”

o “muchos soldados mueren en la guerra” o “qué rico está este sandwich”.

Ahora bien, considerar como criterio el efecto sobre el estado cognitivo en lugar (o aparte) de focalizar la intención tiene algunas ventajas. Si recordamos, hace un momento tuvimos que enfrentarnos al problema del solapamiento entre funciones: una regulación puede declarar y toda declaración pretende en último término regular. Resolvimos que la clave para distinguir la función predominante en un acto comunicativo determinado era concentrarse en la tarea fundamental del oyente. Así, con “llueve a cántaros” sopesábamos si instalar o no la capota antes de quedar empapados y con “el ascensor está estropeado; acabo de hacer la compra; tengo una lesión en el tobillo...” lo que hacíamos era integrar la información nueva que nos presentaban. De modo que, al igual que Werner, pensamos que considerar el impacto del enunciado en el estado cognitivo del hablante es capital (aunque no descartemos la otra dimensión).

3.1.6.3. *La clasificación de Halliday.* Por último, hay una taxonomía más que sirve de fundamento a la nuestra. El también lingüista M. A. K. Halliday distingue tres macrofunciones en el lenguaje adulto (1973, 1978), dos de ellas se corresponden con las descritas aquí. La *interpersonal* es la función del lenguaje que permite gestionar las relaciones con los demás. Esto es, conseguir que otros hagan algo, conducir su acción, atraer la atención del interlocutor para abrir un diálogo, etc. La *ideacional* alude a la función por la cual el lenguaje permite compartir descripciones y reflexiones sobre la realidad. La última es la *textual*, que se refiere al hecho de que toda emisión lingüística genera un texto (o discurso) que puede o no estar ajustado a la situación concreta. Cabría establecer una correspondencia entre lo que aquí denominamos función regulatoria y la interpersonal de Halliday y entre nuestra función declarativa y la ideacional del citado autor. La función textual ni siquiera es tratada menos que someramente por el propio Halliday.

Nótese que Halliday habla de “macrofunciones” y no “funciones” a secas. Ello se debe a que el autor considera que en el lenguaje adulto hay una reducción funcional: de la diversidad funcional infantil se evoluciona hacia una simplificación. Por ejemplo, en el lenguaje infantil Halliday distingue, entre otras, una función “instrumental” con la que el niño demanda algo (e.g., comida, juguete) sin dirigirse a nadie en particular, una función “reguladora” mediante la cual el niño controla la conducta de personas específicas a su alrededor (e.g., “ven conmigo”) y una función “interaccional” con la que el niño inicia, gestiona y termina la interacción con otros (e.g., “hola”). Todas esas son recogidas en la macrofunción adulta interpersonal.

Más importante es señalar que para el autor inglés las tres macrofunciones adultas están presentes en *todo* enunciado. Así, al decir “hace mucho frío” podemos conseguir que el oyente encienda la calefacción al tiempo que le informamos sobre nuestra percepción o al hablarnos sobre sus problemas económicos una persona consigue que le prestemos dinero (éstos son ejemplos ya conocidos para el lector: los utilizamos para ilustrar casos de enunciados en los que las funciones se solapaban). Nosotros razonamos que, aunque con pertinentes matizaciones, existen dos funciones independientes. Halliday no se pilló los dedos: asume que toda comunicación adulta atiende al mismo tiempo a las dos funciones. Esto no es consistente con lo que veníamos planteando, claro.

Tabla 9. Los componentes funcionales del lenguaje adulto, según Halliday.

Componente	Descripción
<i>Ideacional</i>	Componente mediante el cual se comparte información
<i>Interpersonal</i>	Componente mediante el cual se coordina la acción

¿Cómo integrar lo planteado por Halliday y lo planteado por nosotros? Primero, Halliday al menos distingue los mismos dos componentes. Es decir, el autor asume también que el lenguaje sirve para manipular e informar; aunque para él siempre vayan de la mano. Segundo, Halliday nos previene de que hay casos excepcionales en los que existe una total disociación entre componentes. Un ejemplo que ofrece es “¿cómo estás?” que sirve para demostrar que puede reclamarse algo del oyente (en este caso, una respuesta) sin declarar nada. Lo mismo sucede con las producciones de su hijo Nigel hasta los 19-21 meses (las cuales registró religiosamente en un cuaderno): el niño pedía cosas sin decir nada sobre ellas (e.g., “quiero eso”); Halliday interpreta que el niño tardó en “interiorizar el hecho de que el lenguaje también sirve para decir a los otros cosas que no saben” (1973, p. 33). Por último, el autor sostiene que la función predominante en el lenguaje adulto es la ideativa, hasta el punto de que el concepto que cualquier adulto alberga sobre lenguaje es éste (i.e., medio para transmitir información). Es decir, hay dos componentes simultáneos pero que tienen un papel más o menos prominente en un enunciado particular o en una etapa particular.

Lo que podemos aprovechar de la propuesta de Halliday es que (a) hay dos clases de componentes y (b), aunque todo enunciado posee varios componentes, uno es predominante. Estos componentes son: el componente interpersonal, que tiene que ver con el control de la conducta de los demás, y el ideacional, que está relacionado con la transmisión de información. Aunque Halliday no considera que se produzcan independientemente, al menos su clasificación de componentes es paralela a nuestra diferenciación de funciones y se desprende que en un enunciado o etapa dados una función predomina sobre la otra.

3.1.6.4. Conclusión. Recapitulando lo expuesto en este subapartado, hay varias taxonomías funcionales que han servido de base a la nuestra. Givón distinguía una función manipulativa de otra informativa y aportaba criterios para distinguir las condiciones que suscitan su uso. Werner recomendaba contemplar el efecto que los actos comunicativos tienen sobre el estado cognitivo de la audiencia, diferenciando así la función informativa que afecta al estado informacional de la directiva que hace lo mismo con el estado intencional del oyente. Halliday nos hablaba de los componentes incrustados en todo acto comunicativo (aunque en distinta medida según el caso): el ideacional y el interpersonal. Todas estas propuestas identifican una función regulatoria y otra declarativa y, de un modo u otro, trazan una línea entre ambas. En otras palabras, Givón, Werner y Halliday opinan que las funciones del lenguaje son transmitir información y regular la acción. Lo que nosotros propusimos es una síntesis de sus aportaciones.

3.1.7. Balance del primer supuesto. Hemos asumido que existen dos y no más

funciones del lenguaje y hemos reconocido que existen algunos problemas con este supuesto. Vimos que atendiendo a su fuerza ilocutiva pueden distinguirse no dos sino decenas de funciones (amenazar, suplicar, reafirmar, disculparse, etc.). Pero vimos también que concentrándonos en la intención del hablante y en la reacción correspondiente del oyente, siguen diferenciándose dos grandes categorías. Vimos que otros investigadores han propuesto taxonomías funcionales que distinguen no dos sino cinco, seis y hasta ocho funciones. Pero encontramos cierto solapamiento entre las categorías de unas y un análisis de contenido para diferenciar las categorías en otra. Vimos que una regulación puede declarar y que las declaraciones regulan también: esto es, que las funciones pueden confundirse. Pero resolvimos que si focalizamos el proceso que le corresponde implementar al oyente en cada caso, podía apreciarse que una función domina siempre sobre la otra. Por último, vimos que cuando se pide a la gente que agrupe un listado de enunciados, emplea implícitamente una dimensión distinta a la funcional, sembrando la duda de si la dimensión funcional no es más que una cuestión científica. Pero dijimos que no puede descartarse la realidad psicológica de una distinción funcional hasta que el efecto de la valencia emocional no fuese neutralizado.

La conclusión es, pues, que el supuesto sigue en pie. Asumimos que existen dos y no más funciones comunicativas que difieren en la intención del hablante y en la tarea que le corresponde al oyente en cada caso. Tras haber abordado serios problemas, hemos comprobado que el supuesto mantiene su fundamento.

Tabla 10. Apoyo para el primer supuesto.

Problema	Solución
Existen muchos actos ilocutivos	Todos difieren en las dimensiones clave: intención del hablante, reacción del oyente
Otras taxonomías distinguen más de dos categorías	Tienen problemas. (1) Solapamiento entre categorías; (2) No distinguen funciones sino tipos de contenido
La regulación puede declarar también	La tarea principal del oyente sigue siendo decidir
La declaración sirve para regular	La tarea principal del oyente sigue siendo integrar
La clasificación funcional carece de realidad psicológica	No es posible afirmar eso hasta que el efecto de la valencia emocional sea neutralizado Hay evidencias de que reconocemos intenciones

Además, al final de la sección nos hicimos eco de las propuestas de otros investigadores que sí son convergentes con la nuestra. Lo son porque son precisamente las fuentes de las que bebimos para generar el supuesto, que no es más que una síntesis de estas propuestas.

Tabla 11. Integración entre propuestas convergentes (Givón, Werner, Halliday).

Función	Intención hablante	Efecto sobre oyente	Parámetros contextuales	Componente macrofuncional
<i>Declarativa</i>	Modificar conocimientos del oyente	Integrar la información nueva en su banco de conocimientos (impacto en su estado informacional)	(1) Multi-proposicional; (2) contenido ajeno a contexto deíctico; (3) conocimiento interlocutores no compartido; (4) roles desconocidos	Ideacional
<i>Regulatoria</i>	Modificar conducta del oyente	Sopesar si realizar o no lo que se le pide (impacto en su estado intencional)	(1) Mono-proposicional; (2) versa sobre contexto deíctico; (3) conocimiento interlocutores compartido; (4) roles conocidos	Interpersonal

3.2. ¿Cada función requiere procesos distintos?

Habíamos asumido que regular y declarar implican procesos psicológicos distintos. Y es lo mismo para comprender que para producir tales actos. Esto no estaba planteado en las propuestas de Givón, Halliday o Werner de forma explícita, pero se deduce de ellas. A continuación pasaremos revista a una serie de evidencias que podrían apoyar el mencionado supuesto. Comenzaremos rastreando evidencias de que la regulación aparece antes en el desarrollo comunicativo de los niños, de que en los niños autistas la declaración no llega siquiera a aparecer y de que regulación y declaración reciben un patrón prosódico en la producción lingüística de los niños. El problema será que estas evidencias pueden ser discutidas de tal modo que quedan invalidadas. Después rastrearemos otras evidencias que, esta vez sí, parecerán incontestables. Existe una predominancia regulatoria en el desarrollo comunicativo infantil y en el lenguaje de los primates superiores. Terminaremos haciendo balance del apoyo conseguido para el supuesto y proponiendo cómo podrían recabarse apoyos adicionales para él.

3.2.1. La regulación aparece antes que la declaración. En algunos estudios se ha registrado que los niños que aprenden a comunicarse comienzan empleando la regulación (antes de que la declaración forme parte de su repertorio). En la medida en que pueda demostrarse que regulación y declaración tienen un origen ontogenético diferenciado cabría sostener que son procesos distintos. Ahora bien, el supuesto origen diferenciado no está exento de problemas.

Bates, Camaioni y Volterra (1975/1998; 1976) llevaron a cabo un estudio en el que acompañaron a varios niños (de distintos continentes) durante sus primeros meses de vida. En concreto, exploraron las competencias comunicativas de los niños antes de que cumplieren doce meses. Las investigadoras grabaron periódicamente (en audio y vídeo) las interacciones cotidianas de los niños con sus respectivas madres. El interés de las investigadoras residía en el desarrollo de conductas comunicativas prelingüísticas, esto es, la participación de los niños en intercambios comunicativos antes de la aparición de lenguaje propiamente dicho (i.e., uso

de fonemas, palabras, frases, oraciones) en su repertorio productivo.

De acuerdo con el análisis de Bates et al., los niños del estudio exhibieron conductas comunicativas de dos clases. ¿Cuáles eran estas dos clases de comunicación? En unos casos los niños utilizaban el lenguaje “como medio para controlar la conducta del oyente” (Bates et al., 1975/1998: 277). Las investigadoras acuñaron el término *proto-imperativo* para referirse a este tipo de producción comunicativa. Un ejemplo de proto-imperativo es el que sigue.

- (12) Una de las niñas del estudio se encuentra en la cocina en los brazos de su madre. La niña señala al fregadero asegurándose de que su madre mira también y, al instante, la madre llena un vaso de agua y se lo da a beber a su hija. La niña bebe con impaciencia

En (12), la niña utiliza a su madre para conseguir que ésta le dé de beber. Y lo hace a través de su (pre)lenguaje: no pronuncia las palabras “dame un vaso de agua” sino que señala al grifo con la esperanza de que su madre comprenda su instrucción.

En otros casos los niños utilizaban el lenguaje (más bien, cualesquiera los recursos comunicativos de los que disponían) como medio para “ordenar al oyente atender o a hacerse cargo de cierta información” (1975/1988: 277). Esta clase de producciones fue denominada *proto-declarativos*. En esos casos los niños no deseaban conseguir nada concreto de sus cuidadoras. Más bien lo que querían era atraer su atención hacia algo, a juicio de los pequeños, interesante. Un ejemplo aportado por las investigadoras es éste.

- (13) Una de las niñas está sentada sobre los brazos de su madre bebiendo un vaso de leche. Al terminar de beber, echa un vistazo a su alrededor y comprueba que todos están mirándola. En ese momento, la niña emite un ruido cómico. Acto seguido los adultos que la rodean comienzan a reír. La niña repite el mismo juego varias veces.

En esta ocasión (13) la niña no pretende utilizar a los cuidadores como instrumentos (e.g., dame de beber, acércame la pelota) sino que quiere llamar su atención sobre un objeto. El objeto es el ruido cómico que ella misma emite. El equivalente adulto sería algo así como “qué gracioso es esto” o “no puedo evitar mondarme de risa con este ruido”.

Quizá la distinción se entienda mejor cuando la misma conducta es utilizada como proto-imperativo en un caso y como proto-declarativo en otro. En este caso lo vemos con una situación hipotética (aunque inspirada por las descritas por Bates et al.). Un niño podría señalar un globo que hay en sus proximidades para que un adulto que le acompaña en ese momento se lo acerque (e.g., “pásame el globo”). Pero también el niño podría señalar un globo que vuela por el aire mientras pasea por el parque con su padre. Al señalarlo el niño realizaría una suerte de comentario del tipo “qué alto vuela el globo” o “este globo es igual al que tenemos en casa”; en cambio, parecería poco razonable pensar que el niño quiere que su padre vaya en busca del globo volador, lo atrape y se lo dé.

Se desprende de estas observaciones que ya desde el primer momento los niños emplean dos funciones: una proto-imperativa y otra proto-declarativa. Es interesante que tales funciones se realizan antes incluso de que exista un lenguaje propiamente dicho: hay

ilocución (el acto de solicitar, comentar, etc.) pero no locución (el acto de decir). El lector habrá notado, además, que existe una correspondencia entre regulación y proto-imperativo y, por supuesto, entre declaración y proto-declarativo. Ahora la pregunta es cómo las observaciones de Bates et al. podrían ayudarnos a fundamentar nuestra distinción. Pues bien, existen evidencias de que los proto-imperativos aparecen uno o dos meses antes que los proto-declarativos. En concreto, Bates y sus colegas encontraron que la regulación aparecía en torno a los 9 meses, mientras que había que esperar a los 11 para observar declaraciones por parte de los niños.

Cierto apoyo a esta observación proviene de los registros de Halliday (1973), quien tomó nota de todas las conductas comunicativas de su hijo en su desarrollo comunicativo. El inglés interpretó que su hijo no había dominado el componente ideativo del lenguaje cumplidos los 19-21 meses; en cambio, el componente interpersonal sí estaba dominado: el niño sabía pedir, saludar, rechazar, etc. pero no interpretar, comentar, etc.

¿Cuál es el problema entonces? Un estudio algo más reciente no ha podido confirmar este patrón evolutivo (Gómez, Sarriá, Tamarit, Brioso, & León, 1995). Estos investigadores han observado que no hay tal discrepancia evolutiva sino que ambas funciones se solapan en el tiempo. Tenemos, pues, dos contra uno: hay datos dispares.

¿Cómo resolver el problema? Una solución es subrayar las cualidades metodológicas del trabajo de Bates y su equipo. Las investigadoras contaron con niñas y niños de distintos continentes (Europa-Italia y América-EEUU), lo cual permite generalizar los hallazgos. Esto podría conferir a los resultados de Bates mayor validez, respecto de los encontrados por Gómez y otros. Además, el patrón encontrado por Bates se ve confirmado por las observaciones de Halliday (aunque, por supuesto, la metodología del inglés no sea tan sólida como la de aquélla), de modo que hay más apoyos para la divergencia temporal regulatoria-declarativa que para su solapamiento.

Otra solución posible es simplemente descartar la supuesta divergencia ontogenética de la regulación y la declaración y buscar evidencias en otra dirección. Nosotros preferimos ser conservadores, de modo que optaremos por recabar más evidencias y asumiremos, de momento, que no está del todo probado que exista la citada divergencia.

3.2.2. Sin teoría de la mente no hay declaración: los autistas. Algunos investigadores vienen señalando que las regulaciones no entrañan uso de la teoría de la mente mientras que tal uso sería imprescindible en las declaraciones. Si lo sostienen es porque se ha observado reiteradamente que los niños con autismo, a los que se les atribuye una incapacidad para desarrollar una teoría de la mente, producen proto-imperativos pero no proto-declarativos. En la medida en que (a) las regulaciones no entrañen uso de la teoría de la mente o, lo que es lo mismo, en la medida en que (b) en la población de niños autistas se dé un tipo de comunicación y no el otro, podremos afirmar que se trata de procesos diferenciados. Una vez más, veremos que no podemos afirmar felizmente tal cosa: hay algunos problemas en esta postura.

¿Producen proto-declarativos los autistas? García (1995) examinó las capacidades

comunicativas de una muestra de 30 niños (10 con autismo, 10 con retraso mental, 10 normales). García interactuó con los niños siguiendo un guión preestablecido. El guión incluía hacer sonar un sonajero, desperdigar juguetes por el suelo y hacer pompas de jabón, entre otras conductas. El plan era suscitar algún tipo de conducta comunicativa por parte de los niños. Por ejemplo, ante el sonido del sonajero los niños podrían hacer algún comentario del tipo “qué bien suena”; al ver al experimentador haciendo pompas podrían reclamar el objeto para hacer pompas ellos mismos (“dámelo”); y así sucesivamente. Los resultados mostraron que mientras que no había diferencias significativas entre los grupos en las “conductas de demanda de objetos”, los niños con autismo mostraron un número significativamente menor de “conductas de informar” con respecto a los otros dos grupos. Esto significa que los niños autistas son capaces de desplegar proto-imperativos pero no proto-declarativos. Patrones como éste se han venido encontrando en estudios diferentes (e.g., Baron-Cohen, 1989; Mundy, Sigman, Ungerer, & Sherman, 1986).

Como decíamos, una forma de explicar esto es esgrimir que el proto-declarativo exige poseer teoría de la mente mientras que el proto-imperativo no. Las regulaciones supondrían así emplear a los adultos como *instrumentos* (“yo deseo X, lo pido y *algo* me concede el deseo”) y no como entes vivientes o *sujetos*. Declarar supondría contarle al otro algo que no sabe, lo cual exigiría hacer algún tipo de cómputo sobre qué es lo que sabe y lo que no: o sea, exigiría ver al otro como sujeto. Por eso los autistas, que tienen dificultades para dominar una teoría de la mente, presentan el patrón comunicativo descrito líneas arriba.

Pero, ¿sucede realmente así? Gómez, Laá y Phillips (1993) recomiendan recuperar la definición de proto-imperativo propuesta por Bates et al. (1975/1998; 1976). Según las investigadoras, nos recuerdan Gómez et al., realizar un proto-imperativo consiste en (a) emplear los medios comunicativos de los que se dispone (gestos, vocalizaciones), (b) asegurándose de que el cuidador atiende al mismo objeto que nosotros, (c) para conseguir tal objeto. Para garantizar que los adultos atienden al mismo objeto que ellos, los niños miran alternativamente al adulto y al objeto. Así, Gómez y los suyos nos aclaran que, siendo estrictos, incluso los proto-imperativos exigen percibir al otro como un sujeto (en tanto que para proto-imperar hay que mantener una atención *conjunta* sobre el objeto). Gómez et al. proponen que, en contra de la postura tradicional, los autistas no producen siquiera proto-imperativos. Los niños con autismo no se comunican porque toda clase de comunicación implica uso de la teoría de la mente.

Gómez et al. (1993) desarrollan entonces una situación experimental semejante a la de García (1995) para testar su hipótesis. En la situación, un experimentador interactúa, a su turno, con niños normales, autistas y con retraso mental. En la escena hay un objeto inalcanzable para el niño (e.g., un juguete de lo más atractivo). Para llegar a él el niño tiene dos alternativas. Una es emplear una silla disponible en la sala. La otra es contar con la ayuda del experimentador. Si opta por esta alternativa puede desplegar distintos tipos de comunicación. Primero, puede señalar el juguete sin más y esperar que el experimentador se lo alcance (persona como instrumento). Segundo, puede mirar alternativamente al objeto y al experimentador mientras señala al objetivo (persona como sujeto). Desde una postura convencional no se trazaría esta distinción, por tanto, no habría predicciones. Desde la postura de Gómez et al. se espera que los niños con autismo produzcan menos conductas del segundo

tipo. Los resultados muestran que, aunque autistas, normales y con retraso son iguales a la hora de emplear a los otros como objetos, los autistas los tomaron como sujetos un número significativamente menor de veces. Y ello con respecto a los normales o a los niños con retraso.

En vista de los resultados del experimento de Gómez, Laá y Phillips, cabe decir que los niños autistas no mantienen una verdadera comunicación con los que les rodean. O sea, aunque se venía asumiendo que los niños con autismo usan proto-imperativos y no proto-declarativos, lo cierto es que es que ni los unos ni los otros. Esto supone que tanto uno como otro tipo de comunicación exige habilidades de lectura mental (o teoría de la mente): así que ya no es posible apelar a tales habilidades para diferenciar funciones comunicativas.

Lo que acabamos de decir es consistente con lo que planteamos en el apartado 2.2. En su momento dijimos que para comunicarse hay que ser capaz de manejar el mismo set de evidencias que nuestro interlocutor. Considerar evidencias era necesario para la comunicación, porque la decodificación no agota el proceso interpretativo. Inferir qué evidencias puede estar sopesando el otro exigía ponerse en su lugar, leer su mente. Sea declaración o regulación (recuérdese que el ejemplo de entonces era una regulación: “sé más silenciosa”), la teoría de la mente es un requisito. Sperber y Wilson (2002) llegaban a proponer un sub-módulo comunicativo de la teoría de la mente.

Por tanto, retomando la cuestión principal, seguimos sin recabar las evidencias que buscábamos. Planteamos que si llegaba a demostrarse que los niños con autismo usaban un tipo de comunicación y no otro, tal disociación probaría que regulación y declaración son procesos distintos. Y ello porque cabría deducir que la regulación exige teoría de la mente, cosa que no sería imprescindible para declaración. Pero visto que no sucede tal cosa (Gómez et al., 1993), tendremos que buscar evidencias en otra dirección.

3.2.3. Los patrones prosódicos en la comunicación temprana. Existen otras observaciones de conductas comunicativas infantiles. Lo interesante es que en algunas de ellas se ha constatado una relación sistemática entre función lingüística y patrón prosódico empleado por el niño. Antes de continuar se hace necesario definir qué entendemos por prosodia (aunque será tratado con detalle en el capítulo 2). De momento, será suficiente saber que la prosodia aglutina todos los fenómenos melódicos de una lengua (e.g., Pierrehumbert & Hirschberg, 1990). Estos fenómenos son, por ejemplo, el fraseo (línea melódica), el énfasis (picos en el enunciado) o el uso de las pausas (silencios).

Pues bien, como mencionamos, el lingüista Halliday registró todos los acontecimientos lingüísticos relevantes que protagonizó su hijo Nigel en sus primeros 21 meses de vida (1975, 1978), fijándose también en ciertos rasgos prosódicos. De su laborioso trabajo de registro cabe resaltar lo siguiente. Según Halliday, cuando el niño usaba el lenguaje para materializar sus requerimientos, regular la conducta de otros o gestionar la interacción con otros, su entonación era ascendente. Esto es, allí donde el niño esperaba una respuesta conductual por parte del interlocutor, su tono dibujaba una línea de abajo-a-arriba. En otras ocasiones el niño usaba el lenguaje para referirse y explorar la realidad y la fantasía. En estas ocasiones la entonación era descendente.

Por tanto, la regulación en su versión infantil podría tener siempre una entonación ascendente y la declaración, por su parte, descendente. Así pues, en la medida en que cada tipo de comunicación lleva sistemáticamente asociado un patrón prosódico particular, cabe sostener que son funciones o tipos de comunicación diferentes.

Si bien, hemos de ser cautos al considerar el trabajo de Halliday por varias razones. La primera es que no nos proporciona un análisis detallado de las producciones de Nigel. Es decir, Halliday se basa en su impresión al escuchar a su hijo pronunciar regulaciones y declaraciones no en un análisis de los registros acústicos de éstas (e.g., comparando la frecuencia o la intensidad de las ondas). Tampoco nos dice si otros oyentes validaron su impresión. Esto podría habernos dado una idea de la “realidad psicológica” de las diferencias prosódicas entre las producciones (dado que es posible que aún no existiendo diferencias acústicas significativas entre las producciones, los oyentes sistemáticamente perciban una diferencia¹²). Ninguno de estos controles se llevó a cabo, lo cual hace dudar del análisis del autor. Segundo, se trata de un estudio de caso, no de una descripción de una muestra representativa de niños. Nadie puede garantizar que Nigel no fuese un niño excepcional (la única manera de garantizarlo es explorando un número suficiente de casos, en los que encontrar una regularidad).

Resumiendo, otra prueba que regulación y declaración son funciones diferentes podría ser que los niños aplican sistemáticamente un patrón prosódico distinto a cada cual. El problema es que los datos en este sentido provienen de un estudio con serias limitaciones metodológicas. Es hora de acudir a evidencias más sólidas.

3.2.4. La prevalencia regulatoria en la ontogenia. En cuanto al desarrollo comunicativo infantil, una última posibilidad es conformarnos con demostrar que hay una prevalencia regulatoria. No nos malinterpreten, no es en absoluto un premio de consolación: si un tipo de comunicación se instala en el repertorio del niño con más arraigo (i.e., durante una etapa es dominante) que el otro tipo, entonces podríamos sostener que son tipos o funciones diferenciadas.

Bates et al. (1975/1998) se plantean si, en realidad, los proto-declarativos observados en los niños de su muestra no son más que formas de llamar la atención (*attention-pulling*) de los adultos. Es decir, no se trata de compartir un comentario (“qué alto vuela el globo”) sino de conseguir consideración o atenciones. Así, proto-declarar puede concebirse como una forma de modular la conducta del oyente para un fin muy concreto (“ven aquí y hazme caso”). Al menos, un buen número de proto-declarativos podría reinterpretarse así.

Como apoyo para este análisis cabe presentar dos argumentos. Uno, en la medida en que los supuestos comentarios (proto-declaraciones) de los niños tienen un contenido mínimo, casi nulo, y, en todo caso, éste versa sobre la realidad circundante o contexto deíctico, e implica a interlocutores cuyos conocimientos son ampliamente compartidos¹³ y sus roles

12 A la postre, la acústica y la percepción discurren por caminos distintos, como demuestra el hecho de que un chino-hablante no percibe la diferencia entre /pera/ y /pela/.

13 Es verdad que el adulto posee un cuerpo de conocimientos enorme, que para nada es compartido por el

conocidos, Givón (con arreglo a su análisis de parámetros contextuales) entendería que se trata de una regulación. Nosotros mismos diríamos que si lo que le corresponde al oyente es nada más que decidir si atender o no (sin necesidad de integrar nueva información), se trataría más bien de una regulación. De acuerdo con esto, muchos de los proto-declarativos de los niños de Bates fueron en verdad proto-imperativos.

Dos, incluso cuando siendo algo más mayores (antes de los cinco años) los niños de la muestra de Bates et al. realizan declaraciones inequívocas, éstas casi siempre van acompañadas de elementos que parecen delatar la verdadera intención del hablante:

- (14) “Adivina qué. Hoy voy a comer con mi padre”
- (15) “Héctor ¡Héctor! Mira: el spiderman negro”

Tal como apunta Bates elementos como “adivina qué” o “hey, mira” representan indicios de que lo que realmente quieren los niños no es compartir impresiones sino demandar atención, esto es, lograr que el oyente haga algo concreto. Apelando a los criterios que nosotros proponíamos, aquí diríamos que el objetivo del hablante es controlar la conducta del otro; siendo la tarea del oyente, la correspondiente a la regulación (en gran medida).

Teniendo en cuenta estos dos argumentos cabe decir lo siguiente. Aunque se solape en el tiempo la aparición de regulación y declaración, la primera aparece de forma sensiblemente más rotunda: domina la producción lingüística temprana. Así, sus líneas evolutivas divergen, por lo que entendemos que son procesos diferentes.

3.2.5. La prevalencia regulatoria en la filogenia. Existen indicios de que los primates superiores que aprenden a usar el lenguaje emiten preferentemente mensajes regulatorios. Para empezar, están los datos de Brakke y Savage-Rumbaugh. Estas investigadoras publicaron en 1995 y 1996 los resultados de un estudio muy particular. Las investigadoras habían criado a dos pequeños primates desde su nacimiento en un entorno semejante al de los bebés/niños humanos. Un chimpancé y un bonobo participaron en el estudio durante tres años. El entorno en el que crecieron estos primates se caracterizó por lo siguiente. Los primates vivían en una casa con estancias diferenciadas: un dormitorio, un cuarto de juegos y una cocina. Las cuidadoras convivían con ellos y les hablaban a través del discurso oral o a través de un panel de símbolos. No les proporcionaban una instrucción directa del lenguaje, simplemente les hablaban esperando que el lenguaje (tanto receptivo como productivo) surgiese en ellos como surge en los niños humanos. Los primates, por su parte, tenían la libertad de expresarse o no y de hacerlo del modo que prefiriesen (e.g., a través del panel de símbolos, gestos). El objetivo era examinar su capacidad para aprender lenguaje cuando están inmersos en un medio cultural como el humano. Un objetivo específico era comparar los logros de cada tipo de primate. Toda producción comunicativa de los primates fue registrada en un cuaderno de campo. Lo mismo se hizo con las ocasiones en las que los primates dieron muestras de comprender las emisiones de las cuidadoras.

Por lo que respecta a la producción, con el corpus resultante de emisiones se hizo un

adulto. Si bien, el conocimiento que el adulto pone en juego cuando se encuentra en interacción con el niño representa sólo una parcela de su conocimiento total: justamente el conocimiento poseído conjuntamente.

recuento y una categorización. El primer dato al que vamos a hacer referencia es el total de emisiones: el chimpancé produjo 21.706 y el bonobo 27.344 durante los tres años del estudio. En cuanto a la categorización, se distinguió (entre otras cosas) entre peticiones, no-peticiones y otras. Las peticiones eran emisiones (a través del panel de símbolos) del tipo “jardín, jugar” (i.e., “vayamos a jugar al jardín”) o “cazar” (i.e., “intenta atraparme”). La categoría no-petición englobaba comentarios, afirmaciones y respuestas a preguntas. Por ejemplo, los primates pulsaban el símbolo “perro” en el panel cuando escuchaban los ladridos de un perro en el jardín o veían a uno corriendo entre los matorrales. Por último, la categoría otros recogía, esencialmente, imitaciones, balbuceos y errores. El dato más relevante para lo que estamos tratando aquí fue la proporción de emisiones que pertenecía a cada categoría. En el caso del chimpancé las cifras son éstas: el 67,7% fueron peticiones, el 9,2% no-peticiones, el resto fueron otros. En el caso del bonobo las cifras son, a grandes rasgos, equiparables: el 65,6% fueron peticiones, el 17,6% fueron no-peticiones, el resto fueron otros. Estas cifras nos permiten establecer unas cuantas conclusiones. Una es que los primates superiores son capaces de aprender lenguaje incidentalmente, es decir, sin instrucción explícita. En este sentido (y no en otro), son comparables a los humanos. Otra conclusión es que los bonobos demuestran mejores capacidades para el lenguaje (los resultados de otros estudios en el mismo laboratorio son consistentes con los aquí presentados): hablan más y la proporción de emisiones de tipo no-petición fue significativamente mayor. La última y, para nosotros, más importante conclusión es que dos tercios de la producción fueron peticiones o, en otras palabras, regulaciones.

Merece la pena señalar que hay datos convergentes con éstos y que proceden de otro trabajo. Gómez, Sarriá y Tamarit (1993) criaron un gorila en un entorno humano (a diferencia de Brakke y Savage-Rumbaugh, no le permitieron usar el panel de símbolos). Los investigadores reportan que los proto-imperativos desplegados por el gorila se volvieron más y más sofisticados con el tiempo. En un primer momento, el gorila empujaba al criador con todas sus fuerzas para lograr algo (e.g., que le alcance una pelota en un estante). Más adelante, el gorila guiaba al cuidador hasta el objeto de su interés y esperaba a que el cuidador actuase sobre él. Finalmente, el gorila conjugaba miradas hacia el objeto con miradas hacia el cuidador. Es decir, las primeras fases se corresponden con una concepción del otro como instrumento, tal como decíamos antes (en el subapartado sobre comunicación autista); en la última el cuidador es visto como un sujeto. Ahora bien, aunque el uso de proto-imperativos fue evolucionando, los investigadores afirman que en *ningún* momento llegó a producir gestos proto-declarativos.

¿Qué es lo que nos dicen estas conclusiones sobre nuestra taxonomía de funciones comunicativas? Como hemos visto, los primates superiores que aprenden a producir lenguaje en un entorno humanizado emiten preferente o exclusivamente peticiones. En tanto que hay cierta divergencia entre regulación y declaración parece razonable pensar que dichas funciones son diferentes.

3.2.6. Balance del segundo supuesto. ¿Requieren las dos funciones procesos diferentes? Comenzamos el apartado planteando la posibilidad de que la regulación aparezca en el desarrollo comunicativo sistemáticamente antes que la declaración. Si siguen patrones evolutivos distintos, asumimos, es porque son distintas. Pero había evidencias contradictorias

y preferimos desechar este argumento. Vimos que la producción de proto-declarativos (en el desarrollo de la comunicación) demanda el uso de la teoría de la mente, siendo prescindible en los proto-imperativos. Y lo vimos a partir de las evidencias de que los niños con autismo producen los segundos y no los primeros. Si cada función exige habilidades distintas, entonces son funciones distintas, razonamos. Pero vimos que un análisis más fino de los proto-imperativos ponía de manifiesto que también exigen teoría de la mente. Vimos que los niños aplican un patrón prosódico particular a las emisiones de cada función. Pero reconocimos las limitaciones metodológicas del estudio que lo había descubierto y decidimos ser cautos rechazando el argumento. Este primer conjunto de evidencias no parece muy sólido, a la luz de este balance.

No obstante, vimos también que tanto en la comunicación temprana de niños humanos como en la comunicación de primates superiores criados en entornos hablantes, la regulación es dominante. Dicho de otro modo, la declaración ocupa un lugar marginal en estos tipos de comunicación. Por tanto, cabe decir finalmente que, sí, tienen una evolución diferente: una función se instala en el repertorio de un hablante mucho antes de que la otra se domine inequívocamente. Éstas sí son evidencias que respaldan el supuesto de que las funciones implican procesos distintos.

Ahora la pregunta es, a tenor de las evidencias sólidas y las no tan sólidas, ¿podemos considerar el segundo supuesto suficientemente avalado? Por supuesto, caben dos respuestas. Una opción es pensar que las evidencias no son suficientes y que falta por demostrarse que las funciones implican procesos distintos. Otra opción es considerar que la evidencia reunida es suficiente aval. A nuestro parecer las evidencias son suficientes. En cualquier caso, hay algunas especulaciones que pueden ayudar a reforzar el apoyo conseguido hasta el momento. Las incluimos a continuación.

3.2.7. Evidencias adicionales posibles. Para demostrar que los procesos implicados en la comprensión de regulaciones y declaraciones son exactamente los predichos (i.e., decidir e integrar, respectivamente) necesitaríamos unas pruebas muy específicas.

3.2.7.1. Lesiones cerebrales. Una evidencia que nos serviría provendría de el estudio de las lesiones cerebrales. Si pudiésemos encontrar un trastorno/lesión que (a) afectase la capacidad de decisión y, por extensión, (b) la comprensión de los enunciados regulatorios, (c) pero preservase la capacidad para integrar información, tendríamos ante nosotros la prueba de que regulación y decisión van de la mano. Del mismo modo, si pudiésemos identificar una lesión que dañase selectivamente las capacidades para (a) integrar información y, por ende, (b) la comprensión de declaraciones, preservando (c) la capacidad para seguir órdenes regulatorias, tendríamos una prueba fehaciente de la relación entre integración y declaración.

Lo cierto es que contamos con sugestivos datos en esa dirección. La degeneración fronto-temporal (o FTLT, de la nomenclatura anglosajona *Fronto Temporal Lobar Degeneration*) lleva asociados varios síndromes; es decir, la FTLT puede originar uno entre varios síndromes. Entre ellos están la demencia fronto-temporal (o FTD, de *Fronto Temporal Dementia*) y la afasia progresiva no-fluente (o PA, de *Progressive nonfluent Aphasia*). De acuerdo con Neary et al. (1998), quienes consensuaron los criterios diagnósticos de tales

síndromes, la FTD se caracteriza por una falta de autorregulación y rigidez conductual al tiempo que las herramientas básicas de la percepción y la memoria se conservan. La PA, en cambio, se caracteriza por un trastorno del lenguaje (i.e., alexia, incapacidad para acceder al significado de las palabras) con ausencia de daño en otros dominios cognitivos. Los exámenes neurológicos demuestran que la FTD está asociada a una patología de los dos lóbulos frontales, por lo general simétrica. La PA, por su parte, se manifiesta en forma de atrofia asimétrica, más marcada en el lóbulo izquierdo frontotemporal.

Cabría aventurar que los pacientes de FTD, al demostrar un comportamiento rígido y una falta de autocontrol, tuviesen dificultades para seguir una instrucción sencilla (regulación). Esto, en principio, no sería difícil para los pacientes de PA, quienes conservan sus habilidades sociales y la capacidad para controlar su conducta. El patrón inverso habría de darse ante la comprensión de pequeñas descripciones o explicaciones. Es posible que los pacientes de FTD sean capaces de seguir una historia, mientras que esto mismo podría resultar difícil a los pacientes de PA.

3.2.7.2. Variables críticas. Otra evidencia que nos serviría para demostrar que comprender regulaciones y comprender declaraciones activan procesos distintos (decidir e integrar, respectivamente) sería probar que las mismas variables afectan de modo distinto a cada función. Las variables que podrían considerarse son (a) acceso a la actitud del hablante y (b) autoadministración del input lingüístico. Enseguida revisamos en qué consiste cada variable y cómo, presumiblemente, afectarán a los procesos específicos implicados en la comprensión de regulaciones y declaraciones (demostrando así que los procesos son los propuestos y no otros).

Hemos asumido que el proceso crucial que ejecuta quien recibe una regulación es sopesar si realizar o no lo que se le pide. Es decir, debe preguntarse si está o no dispuesto a satisfacer la demanda que se le hace. Sabemos, porque está documentado, que conocer la actitud de otros hacia cierta tarea afecta a nuestra disposición a realizarla. Por ejemplo, Reeve, Jang, Hardre y Omura (2002) observaron que cuando un instructor de chino mostraba una actitud favorable hacia el aprendizaje del idioma, los aprendices demostraban una mayor disposición a aprenderlo. En la medida en que (a) tener acceso a la actitud de otros afecta a nuestra disposición a hacer algo y (b) quienes reciben una regulación deben sopesar si hacer o no algo, el acceso a la actitud de otros debería afectar a la regulación.

El proceso que pone en marcha quien recibe una declaración es procesar e integrar cierto volumen de información. Tener acceso a la actitud de otros afecta a nuestra disposición a hacer algo pero no (directamente) al modo como finalmente lo hacemos, a la ejecución. Por esa razón tener acceso a la actitud del hablante debería ser irrelevante en la declaración.

Hemos asumido que el proceso que pone en juego quien recibe una declaración es el de integrar información. Es decir, esta persona debe identificar ideas y relaciones para conformar una representación mental e integrarla en su fondo de conocimiento. Sabemos, porque hay evidencias de ello, que la posibilidad de autoadministrar el input lingüístico afecta a nuestra capacidad para incorporar nuevas ideas. Por ejemplo, Rawson y Kintsch (2005) observaron que cuando se permite a unos aprendices leer dos veces un texto sobre dióxido de

carbono aprenden más que quienes lo leen sólo una. En la medida en que (a) poder autoadministrar el input lingüístico influye sobre el modo como incorporamos ideas nuevas y (b) quienes reciben una declaración deben integrar nuevas ideas en su fondo de conocimiento, la autoadministración del input debería tener un impacto sobre la declaración.

Decíamos que el proceso que debe realizar el que recibe una regulación es decidir si hacer o no lo que se le demanda. Autoadministrar el input lingüístico afecta al modo como incorporamos información a nuestro fondo de conocimiento pero no a nuestra disposición a hacer algo. En consecuencia, poder autoadministrar el input no debería ser relevante para la regulación.

En definitiva, hay dos variables que afectarían de forma distinta a la regulación y a la declaración. El tener acceso a la actitud del hablante debería beneficiar a la regulación y no a la declaración. El poder autoadministrar el input lingüístico debería beneficiar a la declaración y no a la regulación. Lo cierto es que, aunque con otro objetivo (el nuestro era diseñar ayudas), los experimentos que serán descritos en el capítulo 4 pusieron estas dos predicciones a prueba. En la medida en que sean confirmadas habremos obtenido el apoyo adicional que andábamos buscando para el segundo supuesto.

En cuanto a cómo recrear esas condiciones (acceso a actitud, autoadministración) remitimos al lector al próximo capítulo. En él veremos que las características específicas de las modalidades oral y escrita del lenguaje son precisamente éstas. Lo escrito proporciona control (i.e., posibilidad de autoadministrar el input) mientras que lo oral presenta expresividad (que encarna las actitudes del hablante hacia lo que dice).

3.2.7.3. Conclusión. Hemos propuesto métodos para recabar apoyo adicional al supuesto de que cada función implica procesos diferentes. Uno de ellos está basado en las descripciones de dos síndromes asociados al FTLD y que, de acuerdo con nuestras especulaciones, afectarían selectivamente a la regulación y a la declaración. Por supuesto, habría que poner a prueba las especulaciones. Otro método consiste en explorar como las mismas variables críticas afectarían de modo diferente a cada función. Los experimentos que presentaremos en el capítulo 4 servirán para probar las especulaciones que hemos hecho al respecto. Aunque, recordemos, el objetivo que nosotros perseguíamos con la tesis era diferente.

4. Resumen, predicciones, consecuencias.

El marco de investigación que presentamos en el capítulo 0 asumía que existen dos funciones del lenguaje. En este capítulo nos preguntamos qué argumentos respaldan tal supuesto. De modo que en este capítulo discutimos sobre las funciones del lenguaje.

Comenzamos especificando que el lenguaje tiene dos grandes propiedades. Es un sistema de representación y es un sistema de comunicación. Como sistema de representación el lenguaje nos permite tener control sobre nuestro conocimiento. Poder manipular el conocimiento es una ventaja. Como sistema de comunicación el lenguaje permite que

obtenemos información valiosa de segunda mano y que cooperemos con otros para alcanzar metas complejas. Eso es también ventajoso. Nos propusimos identificar cuál de las dos propiedades es más importante y para ello buscamos cuál podría haber sido seleccionada por la evolución. Un sistema de representación no habría necesitado el diseño que, de hecho, tiene el lenguaje. Por eso resolvimos que es la comunicación la propiedad que seleccionó la evolución. La comunicación es la propiedad fundamental del lenguaje.

La comunicación, a su vez, sirve a distintas funciones. Para averiguar a cuáles tuvimos que definir qué entendemos por comunicación. Para eso propusimos una síntesis de varias propuestas. La síntesis se articula en los supuestos siguientes. Primero, la unidad básica de la comunicación son los actos comunicativos, que son todos aquellos actos que un hablante despliega para decir algo y que un oyente reconoce como tal. Segundo, en un acto comunicativo cabe distinguir, al menos, dos componentes. Uno es semántico y se refiere a lo que un hablante dice. Otro es pragmático y se refiere a lo que un hablante pretende con lo que dice, esto es, la función que el enunciado cumple. Tercero, hay distintos tipos de acto comunicativo dependiendo de la función que desempeñan. Cuarto, para que la comunicación tenga lugar los interlocutores deben ser capaces de llevar a cabo procesos que se apoyan en ciertos conocimientos, pistas y habilidades. Dos proposiciones adicionales, sobre las que no había tanto consenso, fueron: que las funciones son dos, regular y declarar, y que cada cual implica procesos distintos. Propusimos dos criterios para definir función: objetivo que persigue el hablante y reacción que le corresponde al oyente. Los actos regulatorios persiguen manipular la conducta del oyente, quien debe sopesar si realizar o no lo que se le pide. Los declarativos persiguen transmitir información al oyente, quien debe integrar dicha información.

Planteamos que era necesario buscar apoyo para los dos últimos supuestos. En cuanto a si son éstas y no otras las funciones, tuvimos que hacer frente a varios problemas. Primero, existen muchos actos del habla ilocutivos, por qué reducirlos dos categorías. Aprovechamos entonces para recordar los criterios que habíamos propuesto: todos los actos ilocutivos difieren en el objetivo del hablante y la respuesta del oyente y sólo hay dos objetivos y respuestas posibles. Segundo, existen otras clasificaciones que también distinguen un número reducido de funciones, pero más de dos. No obstante, una vez examinadas de cerca descubrimos que había solapamiento entre las funciones que distinguían. Hechos los ajustes pertinentes podían diferenciarse dos categorías básicas, justo las propuestas por nosotros. Tercero, existen regulaciones que declaran. Pero dijimos que si nos concentramos en la tarea del oyente ésta seguía siendo sopesar si satisfacer o no lo que se le demanda. Cuarto, todas las declaraciones buscan regular en último término. Sin embargo, apelando una vez más al criterio de lo que le corresponde al oyente, pudimos seguir manteniendo la categoría declaración. Quinto, hay indicios de que la gente no opera con una distinción de funciones sino que la valencia emocional prima sobre esa distinción. No obstante, los indicios provienen de un estudio con ciertas limitaciones y, más importante aún, hay también evidencias de que sí distinguimos funciones. Superados los problemas el balance fue positivo. Terminamos la sección retomando las fuentes que sirvieron de base a los dos supuestos finales. Vimos así que hay tres propuestas que coinciden al señalar dos funciones, una regulatoria y otra declarativa.

Por lo que respecta a si las funciones implican procesos diferentes, repasamos distintas

evidencias. Las primeras no demostraron ser todo lo robustas que prometían. En algunos estudios se ha observado que la regulación aparece antes en el desarrollo comunicativo, lo que sugería que las funciones tenían un patrón evolutivo diferente. Pero no se vio confirmado. Algunos estudios apuntaban a que los niños autistas no declaraban pero sí regulaban, lo que apuntaba a que cada función requería habilidades específicas. Pero un análisis más fino reveló que tampoco regulaban. Un lingüista había observado un patrón prosódico particular para cada función en su hijo. Pero el estudio presentaba limitaciones metodológicas. Después rastreamos evidencias más sólidas. Vimos que, aunque regulación y declaración aparecen a la vez en el desarrollo comunicativo, la regulación prevalece. O sea, que finalmente cabe hablar de patrón evolutivo distinto. Vimos también que los primates que aprenden a hablar se sirven prioritariamente de la regulación. O sea, ahí hay también una prevalencia regulatoria. Por último, planteamos dos métodos para recabar apoyos adicionales. Uno estaba basado en las descripciones sobre síndromes asociados al FTLD. Especulábamos con la posibilidad de que, dado que uno afecta al autocontrol y no al lenguaje y el otro presenta el cuadro inverso, cabría esperar que uno afectase particularmente a la regulación y el otro a la declaración. Otro método planteaba que ciertas variables deberían afectar a cada función de manera específica. En concreto, predijimos que (a) dado que en la regulación el proceso clave es decidir si queremos hacer o no algo y hay evidencias de que las actitudes de otros afectan a nuestra disposición a hacer algo, tener acceso a la actitud de otro facilitará a la regulación; y que (b) puesto que en la declaración el proceso clave es la integración de información y hay evidencias de que la autoadministración del input lingüístico afecta al modo como incorporamos nuevas ideas, la posibilidad de autoadministrar el input facilitará la declaración. Aunque con otro objetivo, esto es justo lo que se pondrá a prueba en los experimentos del capítulo 4. En conclusión, hay evidencias sólidas (y otras menos sólidas) de que existen funciones diferenciadas; si alguien no las considera suficientes, las que recabaremos con los experimentos pueden contribuir a despejar sus reservas.

Retomemos ahora la cuestión central del capítulo, ¿son regular y declarar las funciones del lenguaje? Integrando varias propuestas propusimos que el lenguaje sirve para regular, que es gestionar la acción mediante el lenguaje, y declarar, que es compartir información mediante el lenguaje. Aunque es posible plantear algunas críticas, todas pudieron contraargumentarse. Hay evidencias de que existen funciones diferenciadas siendo justamente las que hemos señalado.

4.1. Predicciones y consecuencias.

De acuerdo con lo que hemos planteado en el capítulo 1 es posible hacer dos predicciones. Su confirmación puede servir como apoyo adicional para el supuesto de que cada función implica procesos distintos,

La primera predicción se basa en éste razonamiento. Por un lado, asumimos que comprender una regulación implica sopesar si satisfacer lo que se nos demanda. Por otro, hay evidencias de que la actitud de otros puede influir sobre nuestra disposición a hacer algo. Por tanto, cabe esperar que si se tiene acceso a la actitud de los otros la regulación se vea facilitada.

La segunda predicción reposa sobre éste razonamiento. Por un lado, asumimos que comprender una declaración supone integrar la información que nos es transmitida. Por otro, sabemos que la autoadministración del input lingüístico puede favorecer la adquisición de ideas nuevas. Por consiguiente, es razonable esperar que si se puede autoadministrar el input la declaración se vea facilitada.

Por supuesto, si las funciones no estuviesen diferenciadas psicológicamente, entonces ninguna predicción se confirmaría. Si regulación y declaración implican el mismo proceso, las variables críticas que seleccionemos (en este caso, acceso a la actitud y autoadministración) afectarán por igual a las dos funciones. Dicho de otro modo, si no existen dos niveles en una de las variables (regulación/declaración) es imposible que se de una interacción con los niveles de la otra (acceso actitud/ autoadministración). Un panorama como éste no sería consistente con los indicios que revisamos y que apuntan a una diferenciación funcional a nivel psicológico.

Confirmar estas predicciones supondría ganar nuevos apoyos para la distinción de funciones que venimos proponiendo. ¿Cómo explicar que las mismas variables afecten de modo distinto al lenguaje si no es apelando a las dos funciones?

Capítulo 2.

Las modalidades del lenguaje: control escrito y expresividad oral.

0. Introducción.

El marco de investigación que propusimos en el Capítulo 0, la aproximación funcional, asume que cada modalidad del lenguaje tiene características específicas, las cuales pueden resultar ventajosas. En concreto, afirmamos que estas características específicas son el control, en el caso de la modalidad escrita, y la expresividad, en el caso de la modalidad oral; y que ambas son potencialmente ventajosas. Ahora la pregunta es ¿cabe pensar así? El presente capítulo tiene como objetivo el argumentar por qué pensamos de este modo.

Vamos a comprobar que no todos opinan de igual manera. Repasaremos las propuestas de diferentes investigadores, las cuales identifican otras características como específicas de lo oral o lo escrito. Más concretamente, veremos que hay quienes creen que la única diferencia entre lo oral y lo escrito radica en una cuestión de decodificación¹⁴: lo escrito exige decodificar las palabras impresas, cosa que no exige lo oral. Otros tantos se dedican a señalar diferencias aisladas, muchas veces sin ánimo de presentarlas como las únicas. En todo caso, son características dignas de consideración en tanto que aparentemente específicas. Por eso hemos elaborado un catálogo a partir de estas contribuciones puntuales. Como resultado de este repaso descubriremos que cada modalidad parece tener no una sino un buen número de características específicas. Si las modalidades difieren en algo más que en el control y expresividad, ¿por qué seguir afirmando que éstas son sus características específicas? Tendremos que hacer una criba para quedarnos sólo con el control y la expresividad.

En la segunda sección trataremos de hacer una primera criba. Nos preguntaremos cuáles de todas esas características son relevantes, es decir, cuáles tienen un impacto en la comprensión (semántica/pragmática) del lenguaje. Así iremos reuniendo evidencias que indican que tal o cual característica tiene un efecto sobre la comprensión. Al final descubriremos que todas las características del catálogo son potencialmente relevantes, de modo que ninguna será descartada. Será necesario buscar otra forma de cribado. Por cierto,

¹⁴ Entendida aquí como la traducción de las palabras impresas y no como la decodificación en el sentido de Wilson y Sperber (2002) que vimos en el capítulo anterior.

esto servirá también para demostrar que, como asumíamos, el control y la expresividad pueden resultar ventajosos.

En la tercera sección haremos una segunda criba. Esta vez cuestionaremos si las características señaladas como específicas son realmente tal cosa. Recuérdese que específico significaba inherente (presente al margen de la coyuntura particular) y exclusivo (presente en una modalidad y no en la otra). Así, es posible las características que no dependan de la modalidad sino de factores coyunturales o que puedan presentarse en ambas modalidades y no sólo en una. Sólo dos características lograrán superar la segunda criba: el control escrito y la expresividad oral. De manera que el segundo supuesto de la aproximación funcional quedará respaldado.

A tenor de lo expuesto, la tabla de contenidos de lo que sigue líneas abajo es ésta. En una primera sección nos preguntaremos cuáles son las características específicas de lo oral y de lo escrito. Propondremos las que nosotros consideramos como tal y las que otros han señalado como tal. En una segunda sección averiguaremos si de todas esas características reunidas alguna es irrelevante y, por tanto, no merece consideración. En vista de que (casi) todas son potencialmente relevantes, en una tercera sección nos veremos obligados a cuestionar si realmente son específicas de cada modalidad. Esto es, nos preguntaremos si las características reunidas son inherentes y exclusivas de cada modalidad. Para terminar ofreceremos un resumen de todo lo dicho. Lo que habrá de habernos quedado claro es que el control y la expresividad son las características específicas de cada modalidad que, además, pueden resultar ventajosas bajo ciertas condiciones.

1. ¿Cuáles son las características específicas de lo oral y lo escrito?

¿Qué diferencia lo oral de lo escrito? A lo largo de las últimas dos décadas (y, probablemente, desde hace muchas más) muchos autores han tratado de responder a esta pregunta. Para ello han identificado características que consideran específicas de una modalidad y otra. Nosotros hemos hecho acopio de esas características que se han venido proponiendo con objeto de aclarar en alguna medida la cuestión.

1.1. Propuestas.

1.1.1. ¿Qué decimos nosotros? Como ya planteamos en el capítulo introductorio, hay dos características que nosotros consideramos propias de lo escrito y lo oral. Éstas son el control (escrito) y la expresividad (oral).

La modalidad escrita otorga *control* al receptor del lenguaje. Control se refiere a la posibilidad (a) de regular el ritmo de presentación del input, (b) de consultar de nuevo el input (releer) y (c) de concentrarse en cada momento en el segmento de interés (párrafo, palabra, etc.). Esto es posible, claro está, por la naturaleza permanente de la escritura. La modalidad oral es volátil, lo que significa que las palabras se desvanecen una vez se han pronunciado (salvo que uno las mantenga activas en su memoria, algo virtualmente imposible cuando el discurso adquiere cierta extensión).

La modalidad oral presenta *expresividad*. La expresividad se refiere a la explotación de recursos diversos. En general, estos recursos se ubican dentro de la prosodia, término que designa al conjunto de fenómenos melódicos de una lengua (e.g., Pierrehumbert & Hirschberg, 1990). Fenómenos como la variación entonacional (e.g., énfasis) o temporal (e.g., pausas). La expresividad puede también referirse a otras formas de paralenguaje, como el uso de gestos manuales (e.g., Bavelas, 1994) o gestos faciales (e.g., Ekman, 1997), entre otras cosas. La expresividad tiene la capacidad para encarnar (a) la intención comunicativa de un hablante y (b) su actitud hacia lo que está diciendo (Wilson & Wharton, 2006). Nada de esto (i.e., fenómenos melódicos, gestos), como es obvio, aparece en lo escrito (aunque algunos dispositivos gráficos puedan transmitir cierta expresividad).

Es de resaltar que no somos los únicos en hacer este análisis. Ferreira y Anes (1994) coinciden al señalar estas características como las específicas de lo oral y lo escrito. Para ellos las diferencias entre la modalidad oral y la escrita del lenguaje son también el control y la expresividad, ni más ni menos¹⁵. Luego nuestro análisis no es exclusivo. En su capítulo de 1994, Ferreira y Anes discuten sobre distintas técnicas para el estudio del procesamiento (on y off-line) del lenguaje oral, partiendo de la base que el lenguaje oral debería recibir más atención por la investigación psicolingüística de la que recibe en comparación con lo escrito. Y si defienden eso es porque no es igual procesar lo oral que lo escrito: lo primero tiene expresividad y lo segundo control.

Así, para nosotros (y Ferreira & Anes (1994) nos respaldan) las características específicas de lo escrito y lo oral son el control y la expresividad, respectivamente. Pero es muy posible que el lector esté pensando en muchas otras características que, al menos en apariencia, pueden atribuirse específicamente a lo oral o a lo escrito. ¿Por qué no considerarlas? Veamos si otros autores recogen esas características en las que más de uno estará pensando. Más tarde veremos si merece o no seguir teniéndolas en cuenta.

1.1.2. ¿Qué se dice en el campo de la adquisición de la lectura? Aprender a leer es un auténtico desafío. Es un desafío para quienes lo aprenden, los cuales deben incorporar una habilidad muy compleja que engloba muchas otras -algo que no encuentran en absoluto fácil. Aprender a leer es “antinatural” llegan a decir algunos (Gough & Hillinger, 1980). Y, como no resulta nada fácil aprenderlo, los educadores deben aplicarse brindando toda su asistencia a los aprendices. Así que es un desafío también para ellos.

Y, ¿por qué exactamente es tan difícil para los aprendices dominar la lectura? Los mismos investigadores responden que lo es porque exige ser competente al decodificar palabras impresas, algo muy sofisticado que no se desarrolla sino en el contexto de lo escrito. Un niño de seis años, apuntan Gough y Hillinger, ha aprendido un promedio de más de cuatro palabras al día desde su nacimiento y, sin embargo, no puede leer una palabra. De modo que lo escrito implica algo completamente específico, para lo que pronunciar y entender palabras

15 Aunque los investigadores señalan tres diferencias, la segunda y la tercera pueden considerarse una sola. La tercera característica que ellos apuntan es que, mientras los límites entre constituyentes se marcan con pausas en lo oral, en lo escrito deben marcarse con espacios y comas. Creemos que el uso de las pausas se ubica en la expresividad.

(habladas) no ayuda, para lo que no hemos nacido.

De hecho, lo que plantea Gough (Gough, Hoover & Peterson, 1996; Hoover & Gough, 1990) es que comprender un texto requiere dos habilidades independientes. Una es la de comprender el lenguaje natural (e.g., entender un relato que alguien nos cuenta). La otra es la de decodificar palabras (e.g., casa, /k/ /a/ /s/ /a/, “casa”, “edificio para habitar”). Sin la primera no sabríamos qué es una casa o qué significa “la casa de Marta era más grande que la de Juan”. Sin la segunda no podríamos saber que casa es “casa” y así sucesivamente. Eso es comprender, ni más ni menos. De ahí que el modelo de Gough y asociados se llame *simple view of reading* (o visión simple de la lectura). Más concretamente su modelo reposa sobre dos proposiciones. Primero, comprender un texto incluye dos componentes, que son los que ya hemos mencionado. Segundo, ninguno de ellos es suficiente por sí mismo. Es decir, siempre es necesario haber dominado en algún grado cada habilidad.

Lo interesante para lo que nos ocupa es que según Gough “más allá de la frontera de la decodificación escuchar y leer requieren esencialmente los mismos procesos” (Gough et al., 1990: 2). Estábamos buscando las características específicas de lo oral y lo escrito; nosotros propusimos dos y Gough plantea otra diferente: la decodificación. Así, la *única* característica específica de lo oral desde la simple view of reading es que no exige decodificar mientras que la única característica específica de lo escrito es que exige decodificar. Eso supone no sólo considerar una nueva dimensión (decodificación/no-decodificación) sino negar las que habíamos propuesto (control/no-control, expresividad/no-expresividad).

Conviene recordar que la percepción del habla no es en absoluto simple. En realidad, es también un sistema muy sofisticado. De hecho, los sonidos del habla se perciben como discretos cuando en realidad forman un continuo; y tampoco hay correspondencia entre los fonemas, que percibimos de forma estable, y las señales acústicas, que son variables y dependientes del contexto acústico (e.g., Massaro, 1994). Ahora bien, las estructuras involucradas en percepción del habla se desarrollan de forma rápida y natural mientras que la decodificación del lenguaje escrito es un proceso lento y difícil. O sea, percibir el habla y decodificar palabras no es para nada comparable.

De modo que hay que añadir una característica específica más a cada modalidad. Al control y la expresividad hay que sumar la necesidad decodificación y la no necesidad de decodificación. Nótese que para la simple view of reading, la decodificación no es una característica que se agrega a las anteriores sino que es la única que distingue lo escrito de lo oral. Cabe encontrar unas cuantas características específicas más, según otros investigadores.

1.1.3. ¿Qué dicen otros psicólogos y lingüistas? Psicólogos y lingüistas han hecho sus aportaciones a la cuestión que nos ocupa. Aunque -¡y esto es muy importante!- cada cual desde su disciplina y cada cual con fines particulares, muchos son los que han identificado otras características como específicas de lo oral y lo escrito. Los lingüistas, a partir del análisis de diferentes corpus (atendiendo a su léxico y sintaxis, mayormente), y los psicólogos, a partir de reflexiones/estudios sobre las diferencias que existen entre el procesamiento oral/escrito o la comunicación oral/escrita. Pasamos revista a todo ello ahora mismo. Advertimos de que para facilitar la lectura hemos tratado de ordenar las nuevas

características específicas por niveles de procesamiento: nivel de input, nivel léxico-sintáctico, nivel semántico, nivel pragmático, nivel interactivo, nivel de pensamiento.

1.1.3.1. Nivel de input. En este nivel hemos englobado las propiedades físicas del estímulo lingüístico. Por eso aquí hemos ubicado el canal sensorial, el control y la decodificación.

La característica específica (quizá) más evidente de las distintas modalidades es el *canal sensorial* que cada cual emplea. Al menos es algo capital para los que diseñan materiales instructivos que incluyen palabras e imágenes (e.g., Mayer, 2001; Sweller, van Merriënboer & Paas, 1998). El lenguaje escrito utiliza el canal visual mientras que el oral utiliza el auditivo. Si se prefiere, podemos decir que la señal lingüística entra en el sistema cognitivo a través del ojo, en el caso de lo escrito, y a través del oído, en el caso de lo oral. Se trata de algo trivial a simple vista pero que puede, bajo unas condiciones precisas, afectar al procesamiento de forma decisiva (algo que, como decimos, interesa mucho a quienes diseñan materiales con palabras e imágenes). Lo que ocurre es que, al menos inicialmente, el input escrito es mantenido y manipulado por un sub-procesador de la memoria de trabajo (entendida como el sistema de mantenimiento y manipulación de la información necesaria para realizar tareas complejas, e.g., Baddeley, 1997) llamado agenda visoespacial. El input oral es mantenido y manipulado por otro sub-procesador llamado bucle fonológico. Resulta que la agenda visoespacial es también el sistema encargado de procesar información pictórica (animaciones, ilustraciones). El corolario es fácil de suponer: si debemos procesar al mismo tiempo imágenes y texto (escrito) un sólo sub-procesador habrá de soportar toda la carga, cosa que no ocurre cuando lo verbal se presenta oralmente. Esto es lo interesante para quienes diseñan materiales instructivos. Después volveremos sobre ello.

En el nivel de input habría que ubicar un par de características más, de acuerdo con los psicólogos. Nos referimos al *control* del input, que está disponible en lo escrito y no en lo oral, y a la *decodificación*, que es necesaria con lo escrito pero no con lo oral. Sobre éstas ya hemos hablado antes. Si aparecen de nuevo es porque pueden considerarse propiedades físicas del estímulo lingüístico: el control depende de si el input es permanente o evanescente y la decodificación de si aparece impreso o no.

Resumiendo, a nivel de input cabe encontrar varias características específicas de lo oral y lo escrito. Lo oral es auditivo, no es susceptible de control y no implica decodificación. Lo escrito es visual, es susceptible de control e implica decodificación.

1.1.3.2. Nivel léxico-sintáctico. En este nivel hemos ubicado propiedades puramente lingüísticas, relacionadas con el uso particular del léxico y la sintaxis de cada modalidad. En concreto, vamos a ver que las modalidades pueden variar en tres dimensiones más: densidad léxica, uso de elementos auxiliares y uso de mecanismos de integración.

Halliday (1987), a quien ya nos hemos referido anteriormente, estuvo interesado en explorar las diferencias entre lo oral y lo escrito desde un punto de vista puramente lingüístico. Lo que hizo fue tomar unas grabaciones de lenguaje (oral) cotidiano y transformarlas. ¿Cómo? Halliday transcribió las grabaciones y aplicó sucesivas reescrituras a

las transcripciones. Por ejemplo, aquí Halliday trabajó sobre la transcripción de un relato sobre una mascota:

- (16) tuve que esperar a que naciera y después tuve que esperar a que cumplierse ocho o nueve semanas de edad
- (17) tuve que esperar a que naciera y tuviera unas ocho semanas de edad
- (18) aproximadamente ocho semanas después de su nacimiento

Así, dice Halliday, podían crearse pasajes que serían reconocidos por la mayoría como orales (16) y pasajes que serían reconocidos por la mayoría como escritos (18). Halliday aplicó el mismo procedimiento pero en sentido inverso. Tomó fragmentos de artículos como los de *Scientific American* y los oralizó.

Una vez hechas las transformaciones Halliday descubrió algo que llamó su atención. La *densidad léxica* de lo escrito es mucho mayor que la de lo oral. La densidad léxica es el número de items léxicos por el número total de palabras o de frases -según el índice. Halliday encontró una densidad de en torno al 33% para lo oral y el 51% para lo escrito empleando el índice items léxicos/total palabras. Halliday cita análisis de otros colegas con resultados comparables. O sea, que la carga léxica es mayor en lo escrito, con respecto a lo oral.

El lingüista Chafe (1982, 1985) estuvo también interesado en analizar, desde un punto de vista lingüístico, las características específicas de lo oral y lo escrito. Una de las diferencias que encontró entre ambas modalidades es la presencia relativa de lo que podemos denominar *elementos auxiliares*, que son más frecuentes en lo oral. Se trata de partículas que no aportan contenido al discurso. Veámoslas en acción con un ejemplo

- (19) “*Bueno, coge Luis, que estaba mosqueado -vamos, no veas cómo-, que tenía un super cabreo por el cacharro que le habían vendido... pilla y va a la tienda y le salta al dependiente, dependiente que había allí: <<devuélveme el dinero, nosequé, que lo que me habéis vendido es un pedazo de cacharro>>; y así dos o tres veces...*”

Aquí cabe diferenciar varios tipos de elementos auxiliares. Para empezar están las coletillas. Éstas son partículas que aparecen al principio y al final de una oración o frase (e.g., “bueno”, “coge”, “va”) y que, además de no ofrecer información alguna, parecen no encajar en la estructura sintáctica convencional (e.g., no es lo mismo “era un niño muy *bueno*...” que “*bueno*, era un niño que...”). Los cualificadores son elementos que intensifican o exageran lo que se está diciendo (e.g., “pedazo”, “nosequé”, “nosecuántas”). Otras partículas reflejan el (aparente) deseo del hablante de asegurarse de que el oyente sigue el hilo de lo que el primero dice. Cosas como “no veas”, “sabes cómo te digo” o “quiero decir que”. Otras partículas más tienen que ver con la vaguedad del habla (e.g., “dos o tres”, “como 100 ó 200 personas”) que se contrapondría a la precisión escrita (“dos”, “154”). Otras más son las disfluencias, que no son más que vacilaciones al expresarse. Pausas largas (“...”), palabras incompletas (“el bi, bich, bicho”) o palabras repetidas (e.g., “dependiente, dependiente”) son, todas ellas, disfluencias.

Dice Chafe que las partículas auxiliares parecen estar más presentes en lo oral. Pero

¿exactamente en qué se basa para afirmarlo? Una de las cosas que hizo Chafe (1982) fue analizar una muestra de textos académicos y varias conversaciones de sobremesa en busca de diferencias oral/escrito. Si en las conversaciones había una media de 8 expresiones tipo “sabes”, “quiero decir”, etc., en los textos académicos éstas brillaban por su ausencia. Además, por cada cualificador escrito Chafe encontró 12 orales. Más todavía, la vaguedad fue tres veces más utilizada en lo oral que en lo escrito. Ravid y Berman (2004), por su parte, solicitaron a un grupo de niños y adultos que elaborasen textos orales y escritos. Los participantes veían primero un vídeo (sin palabras) que contaba una historia cotidiana. Después se les proponía un tema relacionado con la historia del vídeo y se les pedía bien que escribiesen sobre él o bien que hablasen sobre él. Los textos (escritos y orales) producidos fueron analizados lingüísticamente. Las investigadoras encontraron una presencia de coletillas y calificadores significativamente mayor en los textos orales. En suma, en ambos trabajos se constata que en lo oral abundan elementos auxiliares, algo poco frecuente en el lenguaje escrito.

Dentro del nivel léxico-sintáctico cabe encontrar más diferencias, según los lingüistas. De acuerdo con ciertos análisis lingüísticos, el lenguaje escrito utiliza estructuras sintácticas más complejas que las del habla: se caracteriza por su *integración* (en contraposición a la fragmentación oral). De nuevo el análisis de Chafe (1982, 1985) nos ayudará a entender en qué consiste exactamente esta diferencia. Donde el lenguaje oral repite: “los guitarristas no dejan de tocar durante los descansos, los guitarristas nunca están contentos con el volumen de su amplificador”; el escrito integra las frases: “los guitarristas no dejan de tocar durante los descansos y nunca están contentos con el volumen de su amplificador”. Donde el lenguaje oral enumera: “ninguno de los alumnos me escucha, ninguno de mis amigos me escucha, ninguno de mis familiares me escucha”; el escrito empaqueta: “ninguno de mis alumnos, amigos o familiares me escucha”. Cuando lo oral despliega: “teníamos un perro que no paraba de molestar”; lo escrito comprime: “teníamos un perro molesto”. Son ejemplos de frases refundidas, series y participios, respectivamente. Todos estos mecanismos (y algunos más) son los que, según Chafe, emplea el lenguaje escrito para empaquetar más información en menos espacio. Y son los que él se encontró en los corpus analizados.

En síntesis, a nivel léxico-sintáctico los lingüistas han encontrado otras tres dimensiones de variación oral/escrita: densidad léxica, uso de elementos auxiliares y mecanismos de integración. Lo oral tiene poca densidad léxica, muchas partículas auxiliares y poca integración. Lo escrito tiene mucha densidad, pocas partículas y mucha integración. Y seguimos sumando.

1.1.3.3. Nivel semántico. Considerando el texto en su conjunto los lingüistas han podido identificar, a partir de sus análisis de varios corpus, nuevas características propias del lenguaje oral y del escrito. Aquí vamos a revisar la densidad informativa y el modo como se marca la idea central.

Hace un instante hemos visto que el lenguaje escrito, dada su mayor complejidad sintáctica, comprime más información en menos espacio. Es fácil imaginar que si cada línea del texto escrito condensa varias ideas, lo escrito consiga expresar un mayor número de ideas en total: exhibe una mayor *densidad informativa*. Esto es lo que Ravid y Berman (2004)

encontraron examinando los textos orales y escritos que los participantes de su estudio compusieron. Las investigadoras clasificaron los elementos del texto en dos tipos: contenidos y partículas auxiliares (las que antes vimos). La proporción de contenidos en los textos escritos fue de casi el 90%, mientras que sólo el 49% de las palabras pronunciadas en los textos orales fueron destinadas a ese fin.

Otro aspecto que, de acuerdo con algunos lingüistas, distingue los textos orales de los escritos es el *modo como se expresa el meollo*, la idea central. Siguiendo a Tannen (1982), quien analizó conversaciones en cenas de acción de gracias, cabe esperar estrategias distintas en cada caso. En los textos escritos el meollo se lexicaliza, o sea, se enuncia explícitamente; y en los orales se dramatiza o se insinúa, o sea, se revela indirectamente. La primera estrategia es la que la autora llama evaluación externa, siendo de evaluación interna la segunda que hemos descrito. Veámoslo con un ejemplo (inspirado en uno que Tannen extrajo de sus grabaciones):

- (20) “Pues resulta que tengo una alumna de siete años. Y la niña lleva siempre vestiditos, zapatitos, también tiene un bolso muy mono, lleva sus pendientes a juego... Me sigues, ¿no? Además tiene unos andares así y se da unos aires... La verdad es que yo no sé ni qué cara poner cuando se acerca a hablar conmigo...”

Si estamos de acuerdo, el meollo aquí parece ser que la citada niña se comporta de una manera poco adecuada a su edad y esto es algo que desconcierta y divierte a su profesor: la niña se cree toda una señorita. Si aceptamos esta idea como la central del texto, entonces hemos de aceptar también que no se ha dicho con palabras sino que sólo se ha insinuado. Se ha aplicado entonces una estrategia de evaluación interna, en palabras de Tannen. La estrategia de evaluación externa incluiría, por ejemplo, una expresión final del tipo “la niña actúa como si tuviese 30 años cuando en realidad tiene siete, lo cual me descoloca bastante” o “vamos, que la niña se debe de creer que tiene 30 años o algo parecido”. Por supuesto, para poner en marcha la evaluación interna en toda su plenitud, el hablante puede servirse de claves prosódicas y paralingüísticas. Es decir, se sirve de los rasgos expresivos a los que ya nos hemos referido antes y en los que vamos a insistir en el sub-subapartado que viene a continuación.

En síntesis, en este nivel los lingüistas han identificado dos nuevas dimensiones de variación: la densidad informativa y el modo como la idea central está marcada. De acuerdo con ellas, unas características específicas de lo oral son la poca densidad y la indicación implícita del meollo; las específicas de lo escrito son la alta densidad y la indicación explícita.

1.1.3.4. Nivel pragmático. En el capítulo anterior vimos que comprender el lenguaje es más que recrear el estado de cosas descrito de palabra. Entender lo que nos dicen es también entender qué se nos quieren decir, qué pretenden con lo que dicen y qué actitud tienen para con lo que dicen. Todo eso, dijimos, forma parte de la comprensión pragmática del lenguaje. Vimos que, de acuerdo con Wilson y Sperber (2002), esas comprensiones requerían el concurso de evidencias: conocimientos de la memoria y pistas contextuales. En este nivel de análisis vamos a ubicar las dimensiones que determinarían en algún grado la comprensión que depende de las pistas contextuales. Concretamente y, siguiendo a investigadores como Clark

(1996), Clark y Brennan (1991) o van Dijk (1977), hay toda una serie de claves presentes en el lenguaje natural (la conversación oral) y no disponibles en lo escrito que podrían determinar la comprensión pragmática. Se trata de la expresividad, el entorno perceptivo común y las propiedades atribuibles al hablante/situación. Las repasamos a renglón seguido.

La primera característica distintiva de lo oral que cabe situar en este nivel es la *expresividad*. Ésta ya fue mencionada antes, puesto que era una de las que nosotros consideramos como propia de lo oral (y ausente en lo escrito). Es momento ahora de describirla con más detalle.

La expresividad se refiere, generalmente, al uso de la prosodia y puede extenderse también al uso de otras formas de paralenguaje (como los gestos manuales o faciales). La prosodia, como ya anticipamos, se refiere a todo el abanico de fenómenos melódicos de una lengua (e.g., Beckman, 1996; Ferreira, 1993; Pierrehumbert & Hirschberg, 1990; Wilson & Wharton, 2006); fenómenos que se forman a partir de combinaciones particulares de las tres dimensiones del sonido (i.e., frecuencia, amplitud, longitud, las cuales se corresponden grosso modo con las percepciones agudo/grave, alto/bajo, largo/corto).

Los fenómenos melódicos pueden agruparse en dos categorías: la variación entonacional y la temporal. Dentro de las variaciones entonacionales podemos encontrar fenómenos como (a) el énfasis, que es la prominencia relativa de algunas sílabas (e.g., la diferencia entre “savia” y “sabía”); (b) el fraseo, que es la secuenciación de tonos altos y bajos (e.g., “¿SaBÍas que Óscar VA a ser paDRE?”); o el (c) rango tonal, que es la distancia entre la frecuencia más alta y la línea base de frecuencia. Dentro de las variaciones temporales puede encontrarse el fenómeno del (d) ritmo (e.g., “¿sabías... queee Óscar-va-a-ser-padre?”) que se consigue haciendo cierto uso de la duración y las pausas.

Hay fenómenos prosódicos más específicos. Un ejemplo es el alargamiento de final-de-frase (Ferreira, 1993). Resulta que la misma palabra (con su pausa subsiguiente) es más larga si aparece al final de una frase que si aparece en medio de ella. Así, “gas” (y su pausa subsiguiente) tiende a ser más largo en (21) que en (22):

- (21) Por culpa del impago del gas Rusia corta el grifo de nuevo
- (22) El gas de Rusia es un arma de persuasión poderoso

Lo interesante es que la prosodia puede desempeñar varias funciones. Básicamente sirve para (a) informar sobre la emoción y la actitud del hablante para con lo que dice y para (b) favorecer cierta interpretación de un enunciado de entre varias posibles (Wilson & Wharton, 2006). Así, no es lo mismo “perdí el autobús” que “¡PERDÍ el AUTOBÚS!”: resignación en un caso y enfado en el otro; tampoco es igual “señor muerto, esta tarde llegamos” que “señor, muerto está: tarde llegamos”: interpretaciones totalmente distintas en cada caso. Hay quien le atribuye a la prosodia funciones adicionales como el proveer información sobre el final inminente de un enunciado (Grosjean & Hirt, 1996), lo que libera recursos atencionales, o el facilitar la adquisición del lenguaje (lo que se conoce como hipótesis del *bootstrapping*). Más tarde exploraremos qué hay de cierto en todo esto.

Hay otros recursos expresivos de los que un hablante (que no escritor) puede servirse. Entre ellos están los gestos manuales y los faciales. Nosotros vamos a considerar únicamente los primeros. Por gesto (manual) entendemos todo movimiento manual que, igual que las palabras, se despliega para decir algo a otra persona. Aunque esta definición suene a perogrullo, lo cierto es que no todos suscriben que los gestos cumplen una función comunicativa. Investigadoras como Goldin-Meadow (1999) o Bavelas (1994) opinan que los gestos sí sirven a esta función pero investigadores como Krauss (1998) o McNeill (1985) recelan o se desvían en algún grado de esta asunción. Krauss (1998) sostiene que los gestos no son comunicativos sino que ayudan al proceso de recuperación léxica, esto es, facilitan el acceso a ciertas palabras que queremos pronunciar. McNeill (1985), aunque acepta el valor comunicativo de los gestos, concentra su interés en los gestos como sistema de representación (recuérdese lo que planteábamos en el capítulo anterior: el lenguaje es, a la vez, un sistema de representación y de comunicación). En la segunda sección del capítulo veremos si las evidencias apoyan a unos u otros.

Los gestos que uno puede producir para comunicar son muy variados. Los hay independientes del discurso hablado o asociados a él. El ejemplo más claro de gesticulación independiente del discurso es la lengua de signos. Muchas personas sordas (y todo oyente que lo haya aprendido) se comunican empleando la lengua de signos. Toda la información contenida en las palabras y sus relaciones es, en el caso de la lengua de signos, soportado por los gestos manuales. Otro ejemplo son los denominados emblemas (Ekman, 1997), que pueden emitirse de forma aislada también. Emblemas son el gesto que consiste en juntar las yemas de los dedos índice y pulgar y que significa convencionalmente “okay” o el que consiste en hacer girar el dedo índice junto a la cabeza y que convencionalmente significa “locura”. Nosotros estamos interesados en los gestos asociados al discurso.

Los gestos que acompañan a lo que uno dice pueden, en función de su relación con ello, ser redundantes o informativos. Los redundantes son los que no aportan información que los medios puramente lingüísticos no provean ya. Por ejemplo, uno puede decir “este hombre está loco” mientras produce el emblema al que antes nos referimos (i.e., hace girar su dedo índice junto a su cabeza): su mano no haría sino repetir lo que su boca está pronunciando¹⁶. Los informativos son los que amplían (y hasta contradicen) lo que uno dice de palabra. Éstos son los que vamos a considerar aquí.

Dentro de los informativos cabe hacer más distinciones. Para empezar, estos gestos pueden desempeñar una función semántica o pragmática. Esto es, pueden proporcionar contenidos proposicionales o pueden servir para aclarar la intención del hablante o su actitud hacia el discurso. Una distinción habitual (aunque cada quien utiliza una nomenclatura) dentro de los semánticos es la que diferencia gestos deícticos, icónicos y metafóricos (e.g., Goldin-Meadow, 1999). Los deícticos son aquellos que especifican, a través del señalamiento, cuál es el referente del enunciado que se está produciendo. Decir “acércame ese bolígrafo” mientras se señala uno entre los varios bolígrafos que se encuentran sobre la mesa sería un ejemplo de indicador¹⁷. Los icónicos especifican la forma y tamaño de un objeto. Dibujar en el aire un cuadrado utilizando ambas manos mientras decimos “la plaza tenía esta forma” sería un

16 Aunque eso podría enfatizar lo que uno dice de palabra, luego no sería estrictamente accesorio.

17 Los proto-imperativos del capítulo 1 serían gestos deícticos, con arreglo a esta categorización.

ejemplo de gesto delineante. En los metafóricos el referente es abstracto y no hay correspondencia directa entre las propiedades físicas (ya que no tiene) del referente y los movimientos manuales; la relación, por tanto, es de analogía. Por ejemplo, podría suceder que alguien que nos habla de una discusión que mantuvo con un amigo apriete el puño mientras comenta que “Emilio seguía empeinado, no admitía otra opinión” o que ponga sus dos manos abiertas, palmas hacia afuera, mientras comenta que “llegó un momento en que tuve que calmar los ánimos” (lo primero sugiere perseverancia, obstinación; lo segundo, prudencia, pacificación). Dentro de los pragmáticos los hay que aclaran la fuerza ilocutiva del enunciado al que acompañan (Kendon, 1995), que gestionan la interacción entre los participantes en el intercambio comunicativo (Bavelas, Chovil, Lawrie & Wade, 1992) o que indican el estatus del enunciado al que acompañan con respecto al resto del discurso (McNeill, 1985). Más adelante comprobaremos si, en efecto, la gesticulación sirve a estas funciones semánticas y pragmáticas.

En segundo lugar está el *entorno perceptivo común*, disponible en lo oral y ausente en lo escrito. La comunicación se inscribe en un escenario concreto. Mientras los interlocutores hablan puede comenzar a nevar, el camarero puede hacer una indicación de que el bar va a cerrar o una lámpara puede caer sobre los contertulios. Todos esos eventos son percibidos por los dos participantes en la comunicación si y sólo si ambos comparten el escenario donde tiene lugar la comunicación (y disponen de las capacidades perceptivas pertinentes). En lo oral el escenario es compartido de modo que los eventos son percibidos por todos los presentes. En lo escrito, sin embargo, este acceso queda cancelado. Esta circunstancia hace que enunciados como “olvidé mi abrigo”, “acábate eso” o “si lo sé no vengo” tengan pleno sentido en unos casos (i.e., “quería mi abrigo para protegerme de la nieve”, “termina tu copa porque el bar cierra”, “si llego a saber que una lámpara se nos viene encima no piso este sitio”) pero no en otros.

La tercera característica que muchos consideran, en este nivel propia de lo oral son las *propiedades atribuibles a la situación y al hablante*. Como decíamos, la comunicación se enmarca en un contexto. Cuando sabemos exactamente con quién estamos hablando y bajo qué circunstancias precisas, la comunicación puede ser mucho más fluida. Por ejemplo, si gracias a que tenemos enfrente a nuestro interlocutor, podemos recabar pistas suficientes para determinar que es un policía de paisano, podremos actuar como corresponde ante sus interpelaciones (e.g., “¿qué estás fumando?”). Si no disponemos de esas pistas lo que haremos, en un exceso de insensatez, será mandar a paseo al personaje entrometido (“y a ti qué te importa”). Lo mismo si las evidencias (disponibles o no) nos dicen en qué clase de situación nos encontramos. No es igual decir “viajo a Kabul” si lo haces en octubre de 2001 que si lo haces en febrero de 1989, no es igual que tu jefe te diga “pídeme ayuda cuando lo necesites” cuando acabas de llegar a la empresa que lo haga cuando llevas varios años y acabas de pifiarla otra vez...

Siguiendo a varios psicólogos hemos aceptado que hay nuevas dimensiones, relacionadas con la comprensión pragmática, en las que las modalidades divergen. Lo oral proporciona acceso a la expresividad, a un entorno perceptivo común y a las propiedades del hablante y la situación. Lo escrito, por su parte, cancela el acceso a esas tres fuentes de evidencia.

1.1.3.5. *Nivel interactivo*. En este nivel queremos ubicar la posibilidad de interpelar al otro, de interactuar con nuestro interlocutor. Ésta es para muchos una característica central en el lenguaje (oral) natural (e.g., Clark, 2001, 2002; Pickering & Garrod, 2004) pero, por supuesto, no disponible en lo escrito. Hay dos formas de entender la interacción: interacción abierta y encubierta.

La *interacción abierta* es la posibilidad de interpelar a nuestro interlocutor sin restricciones. Es decir, que (a) los dos participantes puedan hablar y que (b) puedan hacerlo cuando les venga en gana. Esa posibilidad permite dos cosas. (Grosso modo podrían corresponder con la reparación estratégica y la implícita de Pickering y Garrod). Una, solventar cualquier malentendido con sólo solicitar una simple aclaración (e.g., “¿a qué te refieres?”; “¿quieres decir que X?”): la negociación explícita. Algo como eso podría facilitar el que los conocimientos y supuestos de ambos participantes en la comunicación converjan, lo que a buen seguro redundará en una mejor comunicación. Dos, las evidencias de incompreensión que uno va ofreciendo (e.g., comentario: “ayer, charlando con mi padre, todo fue de perlas”; respuesta: “pues qué aburrido hablar sólo de perlas...”), le pueden servir al otro para entrever que algo no va bien y tomar las medidas oportunas. Eso, es de suponer, debe beneficiar la comunicación. O sea, mantener una interacción abierta sirve para que una vez identificados los problemas la reparación sea más fácil y para que la propia identificación de los problemas lo sea también.

Además, es posible hablar de una interacción restringida, la *interacción encubierta*. En ésta sólo uno de los interlocutores habla pero el otro “dice cosas” vía paralenguaje (y aquí sí valen las señales involuntarias¹⁸). Primero, la audiencia puede mostrar expresiones faciales o gestos útiles para el hablante. Si el hablante es mínimamente sensible a las reacciones no verbales del oyente, podrá adaptar su mensaje en consecuencia. Esto es, el hablante puede adecuar el mensaje a las necesidades que el oyente pueda ir experimentando, de acuerdo con la interpretación que aquél hace de las claves no verbales proporcionadas por éste. Si un profesor percibe caras de circunstancia en las primeras filas, escucha murmurar a sus alumnos o ve que uno se ha colocado los auriculares de un iPod seguramente deducirá que su explicación no está teniendo el efecto deseado: los alumnos no están entendiendo y/o se están aburriendo. En virtud de sus observaciones el profesor puede decidir en ese momento elevar el volumen de su voz, contar un chiste o rebajar la densidad de su explicación para devolver a los alumnos al cauce normal de la lección (incluido el gracioso del iPod). No sólo eso. Segundo, tener acceso a información sobre la audiencia (e.g., edad, nivel socio-económico, género) permite conocer qué cosas son compartidas y qué cosas no lo son. Si tienes cuatro años no conviene hablarte sobre lo difícil que es llegar a fin de mes; lo mismo si eres un multimillonario del boom inmobiliario.

En definitiva, en un nivel interactivo (tan importante para algunos psicólogos) pueden encontrarse nuevas características específicas de las modalidades oral y escrita del lenguaje. La primera concede a sus interlocutores la posibilidad de mantener una interacción abierta y

18 En el capítulo 1 decíamos que un acto comunicativo sólo es tal si quien lo produce lo hace queriendo decir algo; o sea, de forma deliberada. En la interacción encubierta es útil servirse de las claves no verbales del otro, aunque sean involuntarias.

encubierta. Sería específico de lo escrito el negar ambas posibilidades.

1.1.3.6. Nivel de pensamiento. Cabe rastrear otra serie de diferencias entre modalidades atendiendo al impacto en el pensamiento y el impacto cultural que una y otra han tenido. Esto es lo que mantienen algunos investigadores. En concreto, nos referimos a la idea de que lo escrito ha podido modificar las representaciones de la realidad que sustentamos, como algunos han señalado.

El efecto más inmediato de la escritura es la propia *noción de fonema*. Las personas que no han aprendido a leer en un sistema alfabético son incapaces de resolver tareas en las que se manipulan fonos (Morais, Bertelson, Cary & Alegría, 1986). Por ejemplo, no pueden eliminar el sonido “/k/” de la palabra “/kasa/” (“casa” sin “c” es “asa”). O sea, que una característica específica de lo escrito sería cambiar el modo como percibimos los sonidos lingüísticos: las señales acústicas se convierten en fonemas una vez hemos aprendido a leer.

Pero es posible ir más allá. David R. Olson publicó en 1977 un conocido artículo en el cual propuso que el pensamiento occidental es el resultado de una transición del habla (*utterance*) al texto (*text*). Gracias a la escritura, el pensamiento informal, intuitivo se convierte en un *pensamiento más refinado*. En este último la explicitud es máxima: las ideas se definen escrupulosamente y todas sus implicaciones lógicas se vuelven claras y verdaderas.

El autor señala tres hitos fundamentales en la evolución utterance-to-text. El primero: el alfabeto griego. Permitió reflexionar sobre el significado de las palabras en sí mismo, abstraído de su contexto. Segundo, la invención de la imprenta hizo necesaria la construcción de textos autónomos, esto es, textos suficientemente explícitos como para permitir que lectores diferentes, en momentos diferentes, pudiesen comprender su significado. El tercer hito fue la consagración de la técnica ensayista inglesa. Se plantearon reglas para conformar textos autónomos convirtiendo al texto con máxima explicitud en un estándar. La escritura se transformó así en una forma de explorar sistemáticamente la realidad y ampliar las representaciones que tenemos de ella.

De acuerdo con estos psicólogos hay otra dimensión en la que las modalidades oral y escrita varían: su impacto en el pensamiento. Lo escrito ha podido suscitar un cambio en nuestra percepción de los sonidos lingüísticos y en nuestro modo de concebir el mundo, mucho más explícito y sistemático ahora. Todo esto sería imposible sin la modalidad escrita.

1.2. Catálogo de características específicas: resumen de las propuestas.

Aunque inicialmente planteamos sólo dos características específicas de lo oral y lo escrito, hemos ido recopilando muchas más. Contamos ahora con un auténtico catálogo basado en las contribuciones de lingüistas y psicólogos.

Lo escrito y lo oral difieren en 14 dimensiones. Éstas son el canal sensorial que cada cual emplea, la posibilidad de controlar el input, la necesidad de decodificar, la densidad léxica, la presencia de elementos auxiliares, el grado de integración, la densidad informativa, la estrategia para marcar la idea central, la expresividad, si el entorno perceptivo es o no

común, si los eventos actuales se perciben por las dos partes o no, si es posible atribuir ciertas propiedades al hablante/situación, la posibilidad de mantener una interacción abierta, la posibilidad de mantener una interacción encubierta, la noción de fonema, el pensamiento refinado. En la tabla siguiente presentamos de qué manera las modalidades difieren en estas dimensiones.

Tabla 12. El catálogo de diferencias oral/escrito.

Dimensiones	Oral	Escrito
<i>Canal sensorial</i>	Auditivo	Visual
<i>Control</i>	No disponible	Disponible
<i>Decodificación</i>	No necesaria	Necesaria
<i>Densidad léxica</i>	Baja	Alta
<i>Elementos auxiliares</i>	Frecuentes	Casi ausentes
<i>Integración</i>	Baja	Alta
<i>Densidad informativa</i>	Baja	Alta
<i>Indicación idea central</i>	Interna (implícita)	Externa (explícita)
<i>Expresividad</i>	Disponible	No disponible
<i>Entorno perceptivo</i>	Disponible	No disponible
<i>Propiedades situación/hablante</i>	Disponible	No disponible
<i>Interacción abierta</i>	Disponible	No disponible
<i>Interacción encubierta</i>	Disponible	No disponible
<i>Noción de fonema y pensamiento refinado</i>	Imposible	Resultado

Como puede apreciarse en la tabla, las acotadas diferencias señaladas al empezar se han convertido en un listado abrumador. Esto pone en tela de juicio nuestra caracterización de las modalidades. Para nosotros difieren en dos dimensiones, para Gough y la simple view of reading en una sola, para un grupo de lingüistas y psicólogos en todas las de la tabla. Tenemos que resolver esta falta de consenso en la caracterización de las modalidades. Para ello vamos a realizar una criba con la esperanza de que podamos reducir el catálogo en algún grado. Un primer requisito para formar parte del catálogo es que la característica debe ser relevante (ventajosa/perjudicial) en la comprensión del lenguaje. Las que no lo sean, serán descartadas.

2. ¿Son relevantes las características identificadas?

Hemos recopilado un buen número de características (supuestamente) específicas de lo oral y lo escrito. La extensión del catálogo contrasta con nuestro planteamiento inicial, en el que sólo había dos características (o con la simple view of reading, que sólo contempla una). Es necesario, como hemos anticipado, hacer una criba para reducir el listado. Lo que vamos a

hacer ahora es rastrear evidencias empíricas que pongan de manifiesto que las características identificadas son relevantes en la comprensión (semántica y/o pragmática) del lenguaje. Si alguna de las características no supone una ventaja o un obstáculo a la comprensión del lenguaje será desestimada.

Los trabajos empíricos que revisaremos provienen de diferentes campos. Concretamente, de los campos de la comprensión de textos, del aprendizaje a partir de materiales instructivos, de la comunicación humana y del procesamiento de oraciones.

2.1. Nivel de input.

2.1.1. Canal sensorial. Decíamos que una característica específica de las modalidades oral y escrita era el canal sensorial que cada cual utilizaba. Decíamos también que esto era particularmente importante para quienes diseñan, con objeto de que promuevan aprendizaje, materiales instructivos que incorporan tanto información verbal como pictórica. En concreto, de acuerdo con el *principio de modalidad*, Mayer (2001) y Sweller et al. (1998) sostienen que dependiendo del canal sensorial en que la información verbal es presentada, un sub-procesador u otro de la memoria de trabajo se pone a funcionar. Las palabras en modalidad auditiva son procesadas por el bucle fonológico mientras que las palabras en modalidad visual son procesadas por la agenda visoespacial, la cual se encarga también de procesar información pictórica. Mayer y Sweller et al. argumentan que si un aprendiz debe aprender de un material con palabras e imágenes y las palabras se presentan visualmente el aprendizaje será menor que si se presentan auditivamente; y ello porque palabras e imágenes habrán de procesarse por un único sub-sistema de la memoria de trabajo y el otro quedará inutilizado. La cuestión es si tales argumentos se ven respaldados por la evidencia empírica.

La respuesta es que sí. En el capítulo próximo el respaldo empírico del principio de modalidad será tratado con mayor profundidad; así que aquí haremos referencia nada más que al balance global. En más de una veintena de experimentos el principio de modalidad se ha visto confirmado (cf. Ginns, 2005). En los experimentos los participantes deben aprender de materiales instructivos que incorporan imágenes (animaciones, ilustraciones) y palabras. Por ejemplo, un material puede consistir en una animación sobre cómo funcionan los frenos de tambor de un coche. La animación presenta distintos eventos que son descritos por las palabras que la acompañan. Las palabras se presentan bien auditivamente, bien visualmente. Tras consultar el material los participantes deben recordar la información que se les ha presentado o deben resolver problemas (transferencia) utilizando los conocimientos que han adquirido (e.g., “imagina que pisas el pedal de freno y no funciona, ¿qué podría haber ocurrido?”, “¿qué podrías hacer para mejorar el funcionamiento de un freno?”; Mayer & Moreno, 1998). Los resultados de experimentos como éste muestran (casi en la totalidad de los casos) que quienes recibieron la información verbal auditivamente rinden mejor en las tareas de recuerdo y transferencia que quienes la reciben visualmente.

Merece la pena destacar que el mismo resultado se ha encontrado en circunstancias diferentes. Con imágenes estáticas (i.e., ilustraciones) y dinámicas (i.e., animaciones), con presentaciones secuenciales (i.e., imágenes-palabras-imágenes-palabras) y simultáneas (i.e., imágenes y palabras a la vez) y con medidas on-line (e.g., carga mental que experimentan los

aprendices) y off-line (e.g., recuerdo y transferencia). Las medidas on-line son particularmente reveladoras porque confirman el principio de forma más contundente: demuestran que, en efecto, quienes reciben la información verbal visualmente experimentan una carga mental (saturación) mayor que la de sus compañeros.

Nos preguntábamos si la primera diferencia entre lo oral y lo escrito, el canal sensorial de cada una, era o no un factor relevante. Hemos constatado que sí, al menos cuando se trata de aprender de materiales instructivos que incorporan tanto información verbal como gráfica.

2.1.2. Control. Una de las dos características que nosotros señalamos como específica de lo escrito es el control. Éste, decíamos, está disponible en lo escrito pero no en lo oral. Ahora se trata de averiguar si resulta relevante disponer de este control para comprender el lenguaje. ¿Importa poder controlar el input a nuestro antojo? Varios hallazgos, que provienen de la investigación en la comprensión de textos, indican que sí.

Antes de revisarlos es preciso asegurarse de que, en efecto, la gente hace uso de este control que se le concede. Lo cierto es que así es. Sabemos que los lectores presentan un patrón de inspección del texto impreso muy particular (Rayner, Juhasz & Pollatsek, 2005). Lo interesante es que no es un patrón lineal, como el que seguimos inevitablemente cuando escuchamos un discurso hablado. Al registrar los movimientos oculares de los lectores encontramos que existen fijaciones, movimiento sacádicos y regresiones. Las fijaciones dan cuenta del tiempo que permanece el ojo enfocando una palabra por primera vez y suelen durar entre 200 y 250 milisegundos (mseg). Los movimientos sacádicos son desplazamientos residuales del ojo: son momentos en los que el ojo hace pequeñísimas fijaciones (20-40 mseg) pero en las que no hay visión. Las regresiones son fijaciones sobre palabras que fueron leídas anteriormente, es decir, son movimientos de re-fijación. Más interesante aún: las fijaciones son más frecuentes sobre palabras de contenido, como nombres y verbos, que sobre palabras funcionales, como artículos o preposiciones (Just, Carpenter & Woolley, 1982). Así que inspeccionamos el texto enfocando en cada momento el segmento de interés y dedicando el tiempo que estimamos oportuno. Todo esto es imposible cuando el estímulo lingüístico nos viene dado.

Y ahora retomando la cuestión, ¿importa que inspeccionemos el input libremente? En primer lugar, releer el mensaje tiene efectos positivos en su comprensión. Rawson y Kintsch (2005) realizaron un pequeño meta-análisis e implementaron dos experimentos explorando si releer un texto tiene efectos positivos en su comprensión. De su trabajo se desprende que los que leían dos veces seguidas recordaban más información y respondían mejor a preguntas sobre los textos que aquéllos que los habían consultado una única vez (suena a perogrullo pero había que probarlo). Y consultar de nuevo el estímulo lingüístico favorece su comprensión aunque se presentase inicialmente en modalidad oral. Mayer (1983) permitió a los participantes de su experimento consultar varias veces el mismo texto. En este caso el texto no se presentó por escrito sino en una cinta de cassette. Los participantes, dependiendo de la condición, podían escuchar el texto una, dos ó tres veces. Después debían recordar toda la información posible sobre él y resolver problemas que requerían poner en juego los conocimientos adquiridos. En las dos tareas los que escucharon tres veces rindieron sensiblemente mejor que los que lo hicieron sólo una vez. Considerando ambos trabajos

podemos afirmar que tener la posibilidad de releer puede ser determinante. Puede serlo porque quienes consultan dos/tres veces el material lo comprenden mejor que quienes sólo lo hacen una vez.

En segundo lugar, regular el ritmo de presentación favorece la comprensión. Concretamente, lo que está documentado es que cuando el ritmo viene impuesto, cuesta más comprender. Kintsch y Keenan (1973) realizaron un experimento en el que varios participantes autoadministraban la presentación de unas oraciones y, otros, recibían cada oración a un ritmo de 330 mseg por palabra (los lectores libres de restricciones temporales ocuparon en promedio 970 mseg/ palabra). El recuerdo inmediatamente posterior de las oraciones presentadas fue peor en la condición de tiempo predeterminado y este patrón fue más acusado cuanto más densas fueron las oraciones.

En tercer lugar, focalizar en cada momento el segmento de interés es también positivo para comprender. De hecho, los lectores que más comprenden son los que desvían su mirada adelante y atrás en función de sus necesidades. Hyönä, Lorch y Haakinen (2002) pidieron a un grupo de participantes que leyese unos textos para más tarde realizar un resumen sobre lo leído. Durante la consulta de los textos los movimientos oculares de los participantes fueron registrados. Lo que observaron Hyönä et al. es que los mejores resúmenes pertenecían a los participantes que habían vuelto la vista a atrás para focalizar oraciones importantes y habían dedicado un tiempo más o menos largo en función de la importancia de las oraciones. Los peores resúmenes pertenecían a los participantes que habían leído linealmente: sin volver a atrás y destinando el mismo a todas las oraciones (al margen de su importancia). A juzgar por los resúmenes, volver la vista a atrás o concentrarse en determinados segmentos en los momentos oportunos es clave para integrar adecuadamente el texto. Otra prueba de que volver atrás puede ser determinante es la que sigue. Just et al. (1982) compararon distintas formas de presentar un texto. Hubo dos formas de presentación que nos interesan: acumulativa y ventana estacionaria. En la forma acumulativa las palabras del texto fueron presentadas en su posición normal (de izquierda a derecha, de arriba a abajo) pero no aparecían todas a la vez: el lector pulsaba un botón para ver la siguiente palabra; así sucesivamente hasta completar el texto. Las palabras viejas permanecían visibles a medida que iban apareciendo las nuevas (“el”, “el gato”, “el gato era”, ...). En la condición ventana estacionaria las palabras se presentaban siempre en el centro de la pantalla. Por supuesto, cada vez que aparecía una nueva palabra del texto, la anterior desaparecía (“el”, “gato”, “era”, ...). La forma acumulativa permitía releer fragmentos previos, cosa imposible en la condición ventana estacionaria. Just et al. encontraron que los lectores en la condición ventana estacionaria emplearon más tiempo en leer que los asignados a la forma acumulativa para igualar la comprensión alcanzada por sus compañeros.

Una puntualización a lo que estamos planteando en este subapartado es la siguiente. La disponibilidad de control es beneficiosa, como vemos, pero hay datos que apuntan a que es así sólo si el material a aprender es complejo. Es decir, el control puede servirnos sólo si lo que debemos procesar comprende muchas ideas y relaciones. Esto es lo que se aprecia al comparar lo que sucede cuando se usan unos y otros tipos de materiales. Cuando el material es simple, no disponer de control es irrelevante (e.g., Gernsbacher, Varner & Faust, 1990; Kintsch & Kozminsky, 1977); cuando es complejo el control marca diferencias significativas:

tenerlo es mejor (Green, 1981).

En suma, releer, regular el ritmo de lectura y enfocar en cada momento el segmento de interés está asociado al aprendizaje. En concreto, el control facilita la adquisición de nuevas ideas y relaciones. Una restricción es que esto parece ser sólo así cuando el material a aprender es suficientemente complejo.

2.1.3. Código. Lo escrito está impreso y lo oral se escucha. El que esté impreso hace que sea necesario decodificar lo escrito, algo muy distinto de lo que pasa al escuchar. Para los que suscriben la simple view of reading (Gough et al., 1996; Hoover & Gough, 1990), como sabemos, esta característica específica de lo escrito es crucial. De hecho, es la única cosa que, a su modo de entender, diferencia lo escrito de lo oral. Ahora la pregunta es si la circunstancia en cuestión es relevante.

Existen evidencias de que la decodificación es un proceso capital en la comprensión del lenguaje escrito. Esto es, se trata de un proceso que tiene un peso en la comprensión de un texto de forma que a mejor decodificación mejor es la comprensión. Las evidencias provienen de distintas clases de estudios. Por ejemplo, hay estudios en los que se entrena a los participantes en la habilidad para decodificar y se evalúa su comprensión lectora después en busca de mejoras significativas. Tan y Nicholson (1997) utilizaron este método. Reunieron una muestra de niños entre 7 y 10 años y los asignaron a varias condiciones experimentales. En dos condiciones de entrenamiento los niños debían leer en voz alta varias palabras extraídas de un texto (una condición leía la lista de palabras, la otra leía frases que contenían esas mismas palabras) hasta hacerlo a una velocidad suficiente (hasta automatizarlo, según los investigadores). En otra condición se preguntaba a los niños por el significado de las mismas palabras en cuya lectura los demás fueron entrenados. De esta manera activaban representaciones semánticas asociadas a las palabras (algo que inevitablemente harían también los niños de las condiciones de entrenamiento) pero no mejoraban su velocidad y precisión al leerlas. Más tarde todos los participantes debían leer el texto en voz alta, contestar a varias preguntas sobre él y se les pedía también que trataran de resumirlo. Los resultados revelan que los niños entrenados en la lectura de palabras (sea en lista o en contexto de frase) rindieron mejor que los otros en todas las tareas: respondieron mejor a las preguntas e incorporaron más ideas centrales en sus resúmenes.

Otro tipo de estudios recluta a buenos y malos comprendedores y los compara en tareas relacionadas con la decodificación. Esto es lo que pusieron en práctica Bell y Perfetti (1994) con alumnos de college. Distinguieron dos grupos: uno en el que los sujetos tenían buenas puntuaciones en un test de comprensión lectora (Nelson-Denny) y otro con sujetos que habían manifestado tener problemas en lectura y cuyas puntuaciones estaban por debajo de la media en el citado test. Después pidieron a todos ellos que resolviesen varias tareas. Entre ellas, lectura de palabras y pseudopalabras (considerando rapidez y precisión). Los buenos comprendedores mostraron un rendimiento significativamente mejor que los pobres comprendedores tanto en rapidez como en precisión al leer palabras.

Ambos trabajos ponen de manifiesto que el proceso de decodificación juega un papel determinante en la comprensión de un texto escrito. En un caso, reforzar la habilidad de

decodificación supone un beneficio en la comprensión lectora; en el otro, se constata que los que comprenden mejor lo hacen, entre otras cosas, porque son mejores decodificadores. A la luz de estas evidencias podemos decir que la decodificación es una diferencia clave.

2.2. Nivel léxico-sintáctico.

2.2.1. Densidad léxica. Si recordamos, Halliday (1987) había analizado lingüísticamente tanto textos que todo el mundo reconocería como orales como textos que todo el mundo reconocería como escritos. Al comparar las dos muestras encontró que los escritos tienen una densidad léxica (o número de ítems léxicos por número total de palabras) mucho mayor que la de los orales. Y Halliday citaba otros trabajos que habían encontrado lo mismo. O sea, lo escrito tiene más palabras contenido que lo oral: en consecuencia, el lector ha de traducir más palabras diferentes que el oyente. ¿Es eso importante?

Se puede razonar del siguiente modo. Si lo escrito contiene más ítems léxicos es más probable que presente un vocabulario más variado, procesar palabras variadas exige tener un vocabulario rico de tal forma que si uno no lo tiene comprende peor. Un razonamiento como éste encuentra apoyo en estudios como el de Nation y Snowling (1998). Las investigadoras reunieron una muestra de niños malos y buenos comprendedores. Les solicitaron, en un primer estudio, que decidiesen si pares de palabras eran o no sinónimas y, en un segundo estudio, que generasen ejemplares para unas categorías dadas (e.g., animal: perro, gato, manatí, etc.). Los malos comprendedores puntuaron menos en ambas pruebas. Por tanto, el volumen del vocabulario que uno posee determina en algún grado la comprensión que alcanza, en tanto que los que comprenden bien son los que tienen más vocabulario.

Lo escrito, con seguridad, presenta un vocabulario variado. Eso puede dificultar la comprensión en tanto que exige poseer un rico vocabulario y no todos lo tienen.

2.2.2. Elementos auxiliares. De acuerdo con los análisis lingüísticos de Chafe (1982) o Ravid y Berman (2004), los elementos auxiliares están muy presentes en lo oral y son escasos en lo escrito. ¿Es esta diferencia un factor crítico en la comprensión del lenguaje?

Antes de responder a la pregunta es preciso hacer memoria. Los elementos auxiliares eran partículas que no aportaban contenido proposicional al discurso. Unos eran las coletillas (e.g., “bueno”, “pues”), otros los cualificadores (e.g., “nosecuántos”, “tatatá tatatá”), otros las vaguedades (e.g., “como 100 o 200 personas”, “un montón”) y otros más las disfluencias (e.g., “er”, “...”).

En el caso de las vaguedades cabe aventurar lo siguiente. Es muy posible que respondan a la ignorancia del hablante y su deseo de no comprometerse. Si no sé si los asistentes al concierto eran exactamente 100, 127 ó 205 siempre puedo decir que había “unos 100-200” y así no faltar a la verdad permitiendo al oyente, de paso, hacerse una idea aproximada de la cantidad de gente congregada. Y lo mismo si digo que había “un montón” de gente. O sea, emplear estos elementos me sirve para esbozar algo sin comprometerme: podría no decir pero el oyente no tendría siquiera un esbozo y podría aventurar una cifra pero el oyente tendría una (probable) mentira. Así que las vaguedades podrían ser resultado de la

ignorancia del hablante y su deseo de no descubrirse (aunque, siendo estrictos, esto necesitaría alguna prueba empírica; por ejemplo, dar más o menos información a unas personas sobre un tema y pedirles que hablen sobre él registrando la presencia de vaguedades: debería ser mayor cuando la información proporcionada es mínima). Y cabe esperar un efecto particular de estos elementos sobre el oyente. Obviamente, la representación mental que puede construir a partir de nuestras palabras es determinada si ofrecemos datos ciertos pero indeterminada si sólo esbozamos la cuestión. Está claro que si no te digo cuántas personas había nunca podrás saberlo.

Respecto de los cualificadores no contamos con pruebas empíricas (que nosotros sepamos) acerca de su impacto en la comprensión del lenguaje. Sólo cabe remitir informalmente a la experiencia que, creemos, todo el mundo ha tenido. Un ejemplo: una amiga nos habla sobre lo que ha tenido que hacer para conseguir la camiseta que quería en plenas rebajas de enero y lo agotada que está por eso mismo.

- (23) “Ve corriendo hasta la parada porque el autobús 13 pasa sólo cada hora. Pelea por ser de las primeras en entrar en la tienda, cuya puerta está abarrotada de gente deseosa de gastar sus cuartos adentro. Tatatá tatatá. Encuentra la camiseta de tu talla entre el montón descuidado de ropa que hay. Tatatá tatatá... Total, ¡que estoy muerta!”

Si estamos de acuerdo, aquí nuestra amiga parece sugerir que, aunque no nos diga todo, hay mucho más, lo que contribuye a su agotamiento final. Es decir, no sólo debes buscar una camiseta de tu talla sino que debes procurar identificar los rasgos distintivos de la camiseta que quieres, agarrar una a una las prendas del montón buscando los rasgos distintivos, ir amontonando en otro lado las prendas que no sirven, asegurarte que la prenda elegida no tiene desperfecto alguno, etc. Así que los cualificadores pueden intensificar lo que uno dice. Cabría esperar que quien recibe cualificadores encaje el discurso de un modo distinto de aquel que no los recibe. No obstante, no sabemos de ningún trabajo empírico que explore esta cuestión.

En cuanto a las coletillas y disfluencias cabe decir lo siguiente (y, esta vez sí, fundamentado en estudios empíricos). Es razonable pensar que coletillas y disfluencias aparecen allí donde el hablante desea encontrar la palabra acertada o plasmar con claridad una idea compleja. Es decir, le sirven como mecanismos para conseguir elocuencia. Hay datos que apoyan esta hipótesis. Por ejemplo, Smith y Clark (1993) pidieron a un grupo de participantes que respondiese unas cuantas preguntas de cultura general (e.g., “¿Cuál es la capital de Chile?”, “¿Cuál es el río más grande de América?”) mientras sus respuestas eran grabadas en audio. Al analizar las grabaciones resultantes Smith y Clark encontraron que había un número significativamente mayor de “umm” y “eer”¹⁹ en las respuestas que incorrectas que en las correctas. Así cuando los hablantes no pueden encontrar “Santiago” (o lo que sea) en su memoria, emplean sistemáticamente partículas que aquí hemos denominado disfluencias. Hay datos coherentes con esto que proceden de la investigación sobre la gesticulación. Krauss (1998) hizo un repaso de sus experimentos sobre la relación entre la producción de gestos manuales y la recuperación de ciertas palabras de la memoria. En un experimento solicitó a un grupo de personas que hablasen sobre experiencias personales mientras se les grababa en

¹⁹ En inglés, idioma utilizado en las investigaciones revisadas, las disfluencias se escriben *um* y *uh*. Nos ha parecido apropiado “traducirlas” al castellano.

vídeo. Después unos actores debían reproducir el “guión” que los primeros habían diseñado espontáneamente: debían imitar las grabaciones. Al comparar los gestos producidos por unos y otros, los investigadores encontraron que ciertos gestos (los que uno hace cuando trata de expresarse con claridad) aparecían mucho más en las grabaciones originales, en comparación con las de los actores. Los actores no debían encontrar la palabra exacta, esa les venía dada por el guión; por eso gesticularon menos. En otro experimento Krauss y sus asociados solicitaron a un grupo de participantes que explicase lo que veían en unas fotografías bajo dos condiciones: libre y gesto restringido. Quienes no pudieron gesticular pronunciaron sus explicaciones a un ritmo más lento que los otros. Todo ello sugiere que algunos gestos pueden servir para facilitar el acceso a ciertas palabras²⁰. Esta clase de gestos podría considerarse como una forma corporal de disfluencia verbal. Hallazgos como los de Smith y Clark o Krauss sugieren que si un hablante se apoya en coletillas o disfluencias es porque lo necesita para decir lo que quiere decir.

Y ¿qué hay del impacto de coletillas y disfluencias sobre el oyente? Sobre esto existen resultados, cuando menos, sorprendentes. En un experimento se puso de manifiesto que las disfluencias facilitan el procesamiento de las palabras a las que acompañan. Corley, McGregor y Donaldson (2007) utilizaron los ERP (*event related potentials* o potenciales evocados) como medida on-line de procesamiento lingüístico. Los ERP reflejan la actividad cerebral asociada a la realización de ciertas tareas. En concreto, los ERP permiten, mediante el pico N400, conocer el momento en el que se produce la interpretación del significado en el cerebro del sujeto. El pico N400 aparece pronto al escuchar palabras frecuentes o predecibles dado un contexto y algo más tarde al escuchar palabras infrecuentes o impredecibles. Lo que hicieron Corley et al. fue confeccionar un set de oraciones parecidas a la siguiente y solicitaron a sus participantes que trataran de comprenderlas:

- (24) Todo el mundo tiene manías y la mía es morderme [“eer”]
 - (a) las uñas
 - (b) la mano

En principio, el pico N400 debe aparecer temprano en las oraciones tipo (a) y retrasarse un poco en las de tipo (b). Esto debería suceder así debido a que (a) es una palabra predecible, a diferencia de (b). La cuestión es que si las disfluencias facilitan el procesamiento de las palabras subsiguientes, la diferencia debería desaparecer cuando éstas son incluidas. Eso es lo que pusieron a prueba Corley et al. insertando o no disfluencias (“eer”) después de la palabra “morderme”. Lo que indicaron los resultados es que cuando no había disfluencias las palabras impredecibles (como b) provocaban un N400 más tardío que las predecibles (como a); cuando había disfluencias las diferencias se disiparon. Además, Corley et al. pidieron a sus participantes que resolviesen una prueba sorpresa de memoria. Los participantes hubieron de reconocer, entre un listado de palabras, aquellas que sí habían escuchado. Las palabras precedidas por disfluencia fueron mejor recordadas. Esto es lo que ocurrió con las disfluencias. Quedaría por saber qué hubiese ocurrido si, en lugar de un “eer” (disfluencia) un “pueees” o un “o sea” (coletillas) hubiese sido utilizado. Aunque no conocemos la respuesta sí podemos conjeturar, sobre la base de otro hallazgo, que el resultado tal vez no se extrapolaría.

²⁰ Esta conclusión tiene sus detractores, con evidencias que avalan la postura enfrentada. Luego lo discutiremos.

Fox-Tree (2001) solicitó a sus participantes que pulsasen un botón justo cuando escuchasen ciertas palabras clave registrando el tiempo que tardaban en hacerlo. Las palabras aparecían en extractos de grabaciones de conversaciones informales. La manipulación fue que en ocasiones la palabra iba precedida por un “umm” o “eer” y en otras ocasiones no (por supuesto, se habían insertado arbitrariamente las mismas disfluencias junto a palabras no clave). Tras analizar los tiempos empleados en detectar las palabras, Fox-Tree pudo apreciar que la detección era más rápida cuando la palabra estaba precedida por un “eer” que cuando la palabra iba sin él. Este resultado es coherente con el de Corley et al. Pero “umm” no tuvo ningún efecto. Es decir, aunque la disfluencia “eer” parece afectar al procesamiento, no cabe decir lo mismo de “umm”. Esto invita a ser prudentes a la hora de extrapolar lo sucedido con “eer” a las coletillas.

En resumen, aunque no podemos afirmar nada con seguridad sobre vaguedades y cualificadores, sabemos que la disfluencia “eer” tiene cierto impacto en el procesamiento y recuerdo de palabras. El que la disfluencia “umm” no haya tenido efecto invita a ser cautos a la hora de generalizar el efecto. También hay evidencias que sugieren que la producción de coletillas y disfluencias responde a un mecanismo del hablante para alcanzar elocuencia.

2.2.3. Integración. En la primera sección se atribuyó al lenguaje escrito, a la luz del análisis de algunos corpus, el uso de estructuras sintácticas más complejas que las que suelen utilizarse en el oral. La complejidad sintáctica servía, como vimos, para empaquetar más información en menor espacio (Chafe, 1982, 1985). Dicho de otro modo: la complejidad hace que el texto sea denso. La pregunta que surge es si la densidad es un factor clave en el procesamiento y comprensión del lenguaje.

Una vez más, tenemos a nuestra disposición evidencias experimentales que apuntan a que sí. Proceden de la investigación en comprensión de textos. Kintsch y Keenan (1973), cuyo experimento fue citado antes, nos ofrecen datos en este sentido. Los investigadores pidieron a un grupo de participantes que leyese una serie de oraciones y párrafos para tratar de recordarlas justo después. Como antes relatamos, en una condición los participantes recibían las oraciones/párrafos a un ritmo fijo; en otra, en cambio, podían consultarlas libremente. Las oraciones/párrafos fueron manipuladas para un número variable de proposiciones manteniendo constante el número de palabras. Así, podía haber oraciones de 15 palabras pero conteniendo 4 ú 8 proposiciones: “Rómulo, el legendario fundador de Roma, tomó a la mujer de Sabine por la fuerza” versus “la pérdida de Cleopatra fue su insensata confianza en los volubles políticos del mundo romano”. Los resultados revelaron que para recordar la información los participantes debían dedicar un tiempo extra a medida que la complejidad (número de proposiciones) aumentaba; además, los mejores recuerdos estuvieron asociados a un mayor tiempo de procesamiento. Por lo tanto, la densidad impone tasas al procesamiento.

Con un procedimiento similar Kintsch, Kozminsky, Streby, McKoon y Keenan (1975) llevaron a cabo varios experimentos más. En esta ocasión utilizaron párrafos en lugar de oraciones. Además, variaron el número de argumentos contenidos en las oraciones manteniendo fijo el número de proposiciones. Es decir, modificaron nuevamente el grado de densidad de las oraciones pero empleando otra estrategia. Los resultados de los dos primeros

experimentos ponen de manifiesto que se requería más tiempo para procesar los párrafos que aglutinaban un mayor número de argumentos. Además, cuanto mayor era el número de argumentos, peor era el recuerdo del contenido de los párrafos. Estos resultados son otra muestra de que a mayor densidad, mayor carga de procesamiento. La carga se refleja en el tiempo invertido y en el nivel de recuerdo alcanzado.

En definitiva, la densidad, que se consigue a base de aplicar mecanismos de integración, es un factor crítico. Lo es porque cuanto más crece, mayor es el tiempo invertido en procesar y peor el recuerdo.

2.3. Nivel semántico.

2.3.1. Carga informativa. Antes describimos escuetamente un estudio en el que las investigadoras encontraron que el lenguaje escrito incluye una mayor proporción de contenidos, una mayor carga informativa (Ravid & Berman, 2004). Los participantes debían producir textos (orales y escritos) sobre un vídeo que habían visto previamente. El número de ideas incorporado en las producciones escritas fue significativamente mayor; algo que no es de extrañar habida cuenta de la integración del lenguaje escrito. Ahora es el momento de valorar si un número mayor de ideas supone un factor crítico.

El estudio de Kintsch et al. (1975) nos será útil una vez más. En su trabajo manipularon el número de argumentos que pequeños textos, igualados en número de palabras, contenían. Como vimos, a medida que el número de argumentos crecía el recuerdo era peor y había que dedicar tiempo extra para asimilar un mayor número de argumentos.

A partir del experimento de Linderholm, Everson, van den Broek, Mischinski, Crittenden y Samuels (2000) también pueden obtenerse datos relevantes para lo que nos ocupa. Linderholm et al. pusieron a sus participantes a leer textos para resolver tareas de aprendizaje posteriormente. Se distinguió entre textos complejos y simples, en función (entre otras cosas) del número de conexiones causales entre los eventos descritos. Ante los textos complejos el rendimiento en las tareas de aprendizaje fue significativamente peor, en comparación con el exhibido ante los simples.

En consecuencia, podemos afirmar que la carga informativa es un factor relevante. Hemos visto que hay datos que indican que a medida que crece el volumen de información a procesar, la comprensión se resiente o necesitamos un esfuerzo adicional para compensar las cosas.

2.3.2. Señalando la idea central. Una de las características que algunos creían propia de lo escrito es que éste delimita explícitamente el meollo del mensaje (a diferencia de lo oral que lo hacía vía implícita). Esto es lo que proponía Tannen (1982) después de analizar conversaciones informales. ¿Marcar la idea central es un factor crítico?

Sabemos que hacer explícitas las ideas centrales de un texto es, de hecho, una variable crítica en la comprensión: quienes reciben estas marcas aprenden más que quienes no las reciben. Mayer, Dyck y Cook (1984) reclutaron varios participantes para hacerlos leer varios

textos y resolver tareas después. Mayer et al. insertaron títulos que contenían (parte de) la idea central en algunos textos. Uno de los títulos fue “El ciclo del nitrógeno y los humanos” y sirvió como anticipo a cinco párrafos que hablaban sobre productos derivados del ciclo del nitrógeno y las ventajas y riesgos que suponen para los humanos. Los participantes que leyeron textos con títulos rindieron mejor en diversas tareas que quienes leyeron los mismos textos sin títulos.

Algo parecido sucedió en uno de los experimentos de McNamara, Kintsch, Songer y Kintsch (1996). En su caso no emplearon títulos sino oraciones completas que recogían la idea central de uno o varios párrafos. Incluyeron, por ejemplo, oraciones del tipo “algunos trastornos del corazón son adquiridos después del parto” como preámbulo a un párrafo en el cual se desarrollaban contenidos en torno a esa idea. Los participantes que disfrutaron de las versiones revisadas siguiendo este procedimiento recordaron mejor las ideas principales del texto.

O sea, si marcamos explícitamente la idea central de un discurso, éste es comprendido mejor. Eso significa que otra de las diferencias que algunos habían encontrado entre las modalidades es relevante.

2.4. Nivel pragmático.

2.4.1. Prosodia. Algunos psicolingüistas habían advertido que hay toda una serie de claves presentes en lo oral y ausentes en lo escrito (e.g., Clark, 1996; Clark & Brennan, 1991; van Dijk, 1977). La primera de ellas es la expresividad que, convenimos, incluía la prosodia y los gestos manuales. Ahora veremos si la prosodia determina o no la comprensión semántica y/o pragmática.

2.4.1.1. La prosodia se produce. Antes de atender a esa cuestión quizá sea necesario demostrar que los fenómenos prosódicos (esos que describimos anteriormente, e.g., fraseo, ritmo) son una realidad. Queremos decir que tal vez alguien esté pensando que, en condiciones naturales, la prosodia es más bien plana y, sólo bajo circunstancias excepcionales, los fenómenos descritos aparecen. Bien, pues sabemos que los fenómenos aparecen sistemáticamente. Cassidy y Kelly (1991) examinaron algunas cualidades fonológicas que distinguen los nombres de los verbos (en inglés). Concretamente, exploraron la posibilidad de que los nombres sean por sistema más largos (i.e., mayor número de sílabas) que los verbos. Lo primero que hicieron fue analizar un corpus lingüístico en el que se incluían varios tipos de materiales, tales como intercambios comunicativos adulto-adulto y adulto-niño. En efecto, los nombres eran mayores que los verbos. En una fase posterior del estudio Cassidy y Kelly proporcionaron un listado de pseudopalabras a una muestra de adultos (e.g., “parturro”, “lifá”). La tarea era construir oraciones empleando las pseudopalabras. Los participantes utilizaron preferentemente las pseudopalabras largas como nombres y las cortas como verbos (e.g., “el parturro se puso a lifear”). Finalmente se solicitó a un grupo de niños (de cuatro años y medio) que hiciesen algo parecido. En esta ocasión se mostraban viñetas en las que aparecía un objeto y una acción (e.g., “un objeto que comienza a vibrar”); junto a ellas se presentaban parejas de pseudopalabras que habían de asociarse a la acción y al objeto representados. También los niños preferían utilizar las palabras cortas para los verbos y las largas para los

nombres. Esto serviría para corroborar la idea de que la prosodia (en este caso, la longitud de algunas emisiones) marca el estatus de ciertas palabras. En su tercer experimento Ferreira (1993) pidió a sus participantes que leyesen en voz alta y con naturalidad varias oraciones. Las oraciones contenían algunas palabras clave. Esas palabras aparecían en distintos contextos según la oración. Se trataba de ver si, al margen de la pronunciación intrínseca de la palabra, habría cambios en su prosodia debidos a la posición que ocupaban en las diferentes oraciones.

- (25) “El chófer creía que podía *parar* el coche.”
- (26) “Aunque el chófer creía que podía *parar*, los pasajeros estaban preocupados. ”

Y así ocurrió. En la segunda versión (26) la duración de las palabras fue mayor y también lo fueron sus pausas subsiguientes. Esto indica que las palabras se alargan al situarse al final (26) de una frase y que en esa posición producen pausas mayores. Tales variaciones sistemáticas señalan la separación entre frases. Fernald y Mazzie (1991) quisieron averiguar si existe una prosodia convencional que marca las frases o si la prosodia a la que están expuestos los niños tiene algo particular (lo que se conoce como *motherese* o habla materna). En ese último caso, podríamos atribuir la característica de marcar las unidades lingüísticas sólo al *motherese* pero no a la prosodia convencional. En su primer experimento los investigadores solicitaron a un grupo de madres que contasen una historia a partir de unas viñetas sin texto. El público fueron sus niños (en una de las condiciones) y un experimentador (en la otra). Las viñetas se elaboraron de tal modo que en cada cual había un nuevo elemento diferenciado del resto (se trataba de varias escenas mostrando cómo una persona se viste y cada vez se coloca una nueva prenda). El modo como las madres pronunciaron ese nuevo elemento (“el sombrero”, “la chaqueta azul”, etc.) fue el objeto del análisis. Los resultados mostraron que cuando la audiencia eran los niños, las madres imprimían un énfasis especial a tales palabras. No fue así cuando el experimentador ejercía de público. El énfasis se midió objetiva (considerando el tono, la intensidad, la duración) y subjetivamente (juzgado por un grupo de oyentes). El experimento apoyaría la idea de que el *motherese* marca las unidades significativas pero no el habla adulta. Un segundo experimento sirvió para matizar esta conclusión. Fernald y Mazzie pidieron a un grupo de adultos que explicase a otro grupo cómo ensamblar las piezas de un electrodoméstico. Esta vez sí apareció una prosodia particular cuando los instructores se referían a elementos nuevos (a.g., “ahora debes coger *la varilla roja* e insertarla en... ”). O sea, bajo determinadas circunstancias la prosodia dirigida tanto a niños como a adultos marcaría específicamente algunos elementos. Smith y Clark (1993), cuyo estudio explicamos antes, encontraron una relación sistemática entre disfluencias y pausas. Las pausas que seguían a un “umm” eran más largas que las que seguían a un “eer”. Y lo confirmaron consultando corpus diferentes. Todos estos estudios muestran que, también en condiciones normales, los fenómenos prosódicos son una realidad.

2.4.1.2. *La prosodia se capta*. Pero hay algo más que objetar. Alguien puede dudar de que los oyentes seamos capaces de captar esos fenómenos. Aunque existan de hecho, si no los percibimos, ¿cómo esperar que tengan un impacto en el modo como procesamos el lenguaje? Para tranquilidad de los recelosos debemos decir que incluso los niños más pequeños son capaces de captar sutilezas prosódicas. Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward y Piwoz (1992) aportan pruebas contundentes. Los investigadores reunieron una

muestra de niños de 6 y 9 meses. Elaboraron una serie de grabaciones en audio que contenían la voz de sus madres hablándoles. En unos casos las pausas entre frases se preservaron tal cual y en otros se editaron digitalmente para encajarlas en lugares poco apropiados:

- (27) la terraza del salón [stop] era bastante estrecha [stop] pero tenía unas vistas impresionantes [stop]
- (28) la terraza del [stop] salón era bastante [stop] estrecha pero tenía unas vistas [stop] impresionantes

Es decir, presentaron a los niños grabaciones que reproducían la voz de sus madres con pausas coincidentes o no coincidentes con los límites entre frases. El mismo fragmento en versión coincidente y no coincidente se reproducía en altavoces situados a ambos lados de la cabeza del pequeño. Los niños de 9 meses prefirieron orientarse hacia el altavoz con pausas coincidentes; no hubo una preferencia clara en los niños de 6. El mismo resultado se encontró en varios experimentos. Por consiguiente, a partir de cierta edad (bastante corta, por cierto) somos sensibles a la prosodia.

2.4.1.3. Prosodia en la adquisición del lenguaje. Hechas las aclaraciones pertinentes, es momento ahora de ver qué impacto tienen los fenómenos prosódicos. Un primer efecto de la prosodia es facilitar la adquisición de una lengua, lo que se conoce como hipótesis del bootstrapping. Según ésta, la capacidad para servirse de la prosodia sería útil para romper el continuo sonoro en unidades significativas (palabras, frases, oraciones), lo que ayudaría en la adquisición de una lengua nueva. La prosodia marcaría así el límite entre frases, las palabras nuevas o la frontera entre una oración y la siguiente. Prueba del papel de estas marcas en la adquisición del lenguaje son las evidencias de que aprendemos mejor la sintaxis de un lenguaje nuevo cuando tenemos acceso a ellas. Morgan, Meier y Newport (1987) idearon un lenguaje artificial. En su primer experimento, pidieron a un grupo de universitarios que lo aprendiese bajo tres condiciones: sin prosodia, con prosodia coincidente y con prosodia no coincidente. Los de la segunda condición aprendieron mejor el idioma artificial que sus compañeros (fueron capaces de abstraer reglas gramaticales, agrupar elementos dados al azar, etc.). Así pues, una de las funciones constatadas de la prosodia es ayudar en la adquisición de un nuevo idioma.

La prosodia tiene un papel en la adquisición de una lengua, a juzgar por estos resultados. Ahora bien, este papel está circunscrito a la fase de adquisición de una lengua: una vez aprendida, el andamio que sirvió para construirla se retira.

2.4.1.4. Prosodia en la desambiguación de oraciones. La prosodia tiene más funciones contrastadas. Otra de ellas es que ayuda a desambiguar oraciones ambivalentes. Un paradigma experimental en este campo consiste en lo siguiente. Se presenta en audio una oración ambigua a los participantes. Se manipulan ciertos rasgos prosódicos de la grabación de dicha oración. Los participantes deben decidir cuál de las posibles interpretaciones es la que el hablante pretende (para lo cual disponen de tiempo ilimitado). Se mide si ciertas variaciones prosódicas sesgan la interpretación, esto es, aumentan las probabilidades de que determinada interpretación sea elegida.

- (29) Martin sostenía que el jefe mintió cuando la investigación comenzó

Aquí caben dos interpretaciones posibles: Martin dijo eso al comienzo de la investigación o el jefe mintió al comienzo de la investigación y eso es lo que sostenía Martin. ¿Cómo resolver la ambivalencia? Carlson, Clifton y Frazier (2001) manipularon los límites entre las frases que componen la oración y presentaron las distintas versiones resultantes a un grupo de participantes. Utilizaron la oración del ejemplo y otras tantas similares. Para modificar los límites entre frases alargaron/acortaron la última sílaba antes del límite y prolongaron/recortaron las pausas entre frases. Se esperaba que cierta configuración hiciese a los participantes decantarse por la primera interpretación y que otra lo hiciese por la segunda. Tras escuchar las oraciones los participantes leían dos paráfrasis de la oración, cada cual recogiendo una de las interpretaciones posibles. En los resultados se aprecia que las distintas configuraciones determinaron la interpretación.

Schafer, Carter, Clifton y Frazier (1996) hicieron algo parecido. En su caso manipularon el énfasis que se hacía sobre algunos elementos de la oración -en lugar de los límites entre frases.

- (30) El sol brilló sobre la hélice del avión que el mecánico estaba examinado cuidadosamente

De nuevo dos interpretaciones son igualmente válidas: el mecánico podría haber estado examinando la hélice pero también el avión. ¿Qué interpretar? Schafer et al. variaron el énfasis realizado sobre las palabras hélice y avión. Presentaron varias oraciones (con sus respectivos énfasis) de esta clase a los participantes, los cuales tuvieron seguidamente que responder preguntas tipo “¿Dónde incidía el sol?”. Los resultados revelaron que la frase final (“que el mecánico...”) se acoplaba a la palabra que recibía mayor énfasis.

Tomados en conjunto, estos experimentos evidencian que el modo como se pronuncian los límites y como se remarcan algunas palabras determina el modo como se interpretan oraciones ambivalentes. La prosodia sirve para desambiguar. Si bien, no es menos cierto que este tipo de oraciones no es muy común. Tiene interés para quienes las utilizan experimentalmente (les sirve para explorar cómo intervienen distintos elementos en el procesamiento del lenguaje y cuándo ejerce cada cual su influencia); pero quizá son situaciones alejadas de la experiencia cotidiana.

2.4.1.5. Prosodia en la predicción del fin de enunciado. Otra función de la prosodia es permitir realizar predicciones sobre el fin de un enunciado. Gracias a la prosodia, los enunciados dibujan una línea entonacional particular (el fraseo, lo llamamos). Dado que somos sensibles al fraseo podemos predecir si una línea toca a su fin o si, por el contrario, se intuye que aún debe proseguir para ser completada. Y ello al margen del análisis semántico de la oración. Según algunos autores la posibilidad de saber si un enunciado toca o no a su fin permite regular la inversión de recursos atencionales. Un ejemplo nos ayudará a entender este fenómeno.

- (31) Hace un rato mi hermana se dio un chapuzón
- (32) Hace un rato mi hermana se dio un chapuzón en la piscina
- (33) Hace un rato mi hermana se dio un chapuzón en la piscina del club

Como puede observarse, la oración (31) puede ser una oración completa pero también puede continuar hasta configurar las oraciones (32) y (33). Es decir, al escuchar “chapuzón” el oyente se halla ante una duda: ¿seguirá hablando el emisor o ha terminado ya su enunciado? Al menos se halla ante tal dilema si la situación experimental les fuerza a tomar una decisión. Grosjean y Hirt (1996) presentaron grabaciones conteniendo oraciones de este tipo a sus participantes. Los investigadores presentaban las palabras de la oración (31) pero extraídas de la pronunciación de la oración (31), (32) ó (33). O sea, un narrador grababa previamente las tres oraciones. Después se editaban los archivos de audio de modo que se suprimía todo lo que va después de la palabra “chapuzón”. Los participantes escuchaban las grabaciones editadas y debían estimar el número de palabras que faltaban para el fin del enunciado: 0, 3 ó 5 (eligiendo entre los formatos 31, 32 ó 33 de las oraciones). En el análisis de resultados se comprobó que sus predicciones se basaban en la prosodia del fragmento escuchado. Cuando escuchaban los fragmentos correspondientes a las oraciones tipo (32) y (33) sabían que el enunciado aún debía continuar; cuando escuchaban la grabación de (31) sabían que estaban ante el fin del enunciado. En un segundo experimento los investigadores encontraron que los participantes eran igualmente capaces de realizar estas estimaciones sólo escuchando la palabra “chapuzón” o parte de ella.

Aunque los autores afirman que saber cuando un enunciado va a terminar libera recursos atencionales, aún está por ver. Queremos decir que no hay evidencias de que ese supuesto ahorro repercuta en el procesamiento de fragmentos más extensos de discurso.

Así, otra función de la prosodia es anunciar el fin de un enunciado. Ello podría redundar, además, en una mayor disponibilidad de recursos atencionales, aunque no lo sabemos a ciencia cierta.

2.4.1.6. Prosodia encarnando actitudes. Una última función de la prosodia es su capacidad para transmitir las actitudes y emociones del hablante para con lo que dice. Revisamos pruebas de que así es.

Brennan y Williams (1995) idearon un curioso diseño (que describimos sucintamente antes) con el fin de explorar el modo como los oyentes se apoyan en la prosodia para inferir estados mentales de los hablantes. Su experimento se realizó en dos fases. En la primera varios participantes respondieron oralmente una serie de preguntas sobre cultura general (e.g., “¿Quién fue el primer presidente de los EEUU?”). Sus respuestas fueron audio-grabadas. Después de responder a las preguntas los participantes debían calificar en una escala el grado de éxito que, a su juicio, tendrían más tarde en otra prueba posterior. Esta prueba posterior consistía en identificar la respuesta correcta a la pregunta original de entre varias opciones (“G. Washington”, “Th. Jefferson” o “B. Clinton”). Al graduar el éxito posterior los participantes expresaban su *feeling of knowledge* (i.e., sensación de que sabes) o FOK. En la segunda fase del experimento otros participantes escuchaban las respuestas en audio de los primeros. Se les presentaba una grabación del tipo:

- (34) “¿Cuál es la respuesta a la pregunta nº 3?” (estos participantes no escucharon la pregunta real planteada a los primeros para evitar tener acceso a cierta información)
- (35) “umm... creo que... sí, eer Thomas Jefferson”

Los participantes en la segunda fase tenían que calcular el *feeling of another's knowledge* (i.e., sensación de que otro sabe) o FOAK a partir de la grabación. Esto es, debían apoyarse en las pistas que la grabación les facilitaba para inferir el grado de seguridad que el participante original tenía al dar la respuesta. Los resultados revelan que hubo una covariación sistemática entre las pistas y las respuestas FOAK. En otras palabras, los participantes en la segunda fase intuían el FOK de los primeros a partir de los rasgos prosódicos exhibidos en la grabación (e.g., (35)). En una tercera fase se replicó la segunda introduciendo algunas modificaciones. En esta ocasión se editaron los archivos de audio. Independientemente del modo original como los contestadores habían pronunciado sus respuestas, Brennan y Williams insertaron pausas de diferente duración (largas/cortas) y modularon el fraseo (ascendente/descendente) de las respuestas. Esto proporcionó mayor control al anterior experimento (toda variable extraña asociada a la duración de las pausas y/o la entonación había sido despejada). Los resultados replicaron los anteriormente citados. El trabajo de estos investigadores pone de manifiesto que la prosodia informa sobre la actitud del hablante hacia el mensaje; más en concreto, la prosodia revela el grado de seguridad que tiene el hablante respecto de lo que dice, ello es percibido por el oyente y tales percepciones le sirven de base para inferir la seguridad del hablante.

El oyente también puede captar la cordialidad con la que el hablante dice lo que dice a través de su prosodia. Haskard, S. Williams, DiMatteo, Heritage y Rosenthal (2008) realizaron un estudio sobre el impacto de la prosodia en la percepción que los pacientes tienen de los médicos que les tratan. Varias interacciones médico-paciente de un hospital de California fueron audio-grabadas. Tras la interacción los pacientes cumplieron una serie de cuestionarios. Específicamente se les interrogó sobre la medida en que creían que su médico contaba con ellos para tomar decisiones, su grado de satisfacción con el médico y grado en que consideraban que su médico les ofrecía información relevante sobre su enfermedad, tratamiento, etc. Las grabaciones de las interacciones eran más tarde presentadas a otro grupo de participantes. Antes de hacerlo, Haskard et al. aplicaron un tratamiento distorsionador a las grabaciones: el filtrado de contenido. Este tratamiento permite anular todo rastro de inteligibilidad sin afectar la prosodia de la grabación. El segundo grupo de participantes debía evaluar al médico (a partir de las grabaciones filtradas) de acuerdo con varias cualidades: cálido-alentador, competente-interesado, hostil-irrespetuoso, entusiasta. Finalmente, la correlación entre las puntuaciones de los cuestionarios de los pacientes y los juicios de los oyentes de la segunda fase fue calculada. En los resultados se pudo apreciar que hubo correlación positiva y significativa entre el grado de satisfacción de los pacientes y los rasgos calidez-alentador y competente-interesado. También entre el grado en que creían que su médico contaba con ellos para tomar decisiones y los rasgos calidez-alentador, competente-interesado y entusiasta. Por último, hubo también correlaciones positivas y significativas entre el grado de información que estimaban se les estaba dando y los rasgos calidez-alentador y competente-interesado y negativa con respecto al rasgo hostil-irrespetuoso. En definitiva, las características prosódicas de la intervención del médico determinaban las inferencias que los

oyentes generaban sobre su actitud, lo que coincidía con la opinión de los pacientes tratados.

Baum y Nowicki (1998) desarrollaron un subtest sobre la capacidad para captar la emoción de un hablante a partir de su prosodia. El subtest forma parte de una prueba más amplia de sensibilidad hacia la conducta no verbal (*Diagnostic Assessment of NonVerbal Accuracy*, DANVA). El subtest incluye 24 grabaciones en audio que son presentadas a los participantes para que éstos las evalúen. Las grabaciones son versiones de un mismo enunciado interpretadas por actores expertos (más de 20 años de experiencia), un hombre y una mujer. El enunciado es “me voy fuera de la habitación ahora, volveré más tarde”, juzgado como neutral por más del 85% de los entrevistados sobre su neutralidad. El mismo enunciado se presenta a los participantes en varias versiones. Las versiones difieren en la emoción que intentan reflejar (feliz, triste, enfadado, asustado) y en la intensidad con la que lo hacen (alta, baja). El subtest incluye tres enunciados de alta intensidad y tres de baja por cada emoción: 24 items (de una muestra de 133 grabaciones las 24 finales se seleccionaron por ser las más fieles al reflejar la emoción e intensidad pretendidos, tal como juzgó el 80% de una muestra de oyentes). El test se aplicó en nueve estudios con niños y adultos. Lo que nos interesa es que los niños de 4-5 años acertaron al identificar la emoción e intensidad correctas en 11 de los 24 items (46% de acierto) y los adultos de 14-19 años acertaron 19 de 24 (79% de acierto). Esto significa que tanto niños como adultos son capaces de reconocer la emoción con la que un hablante dice lo que dice a partir de su prosodia; y lo hacen con un porcentaje de acierto moderado-alto.

Los trabajos de Brennan y Williams, de Haskard et al. y de Baum y Nowicki ponen de relieve que los oyentes son capaces de inferir las actitudes de los hablantes. En concreto, los oyentes son capaces de inferir la seguridad que uno tiene en lo que dice, la cordialidad con la que dice las cosas y la emoción con la que pronuncia una oración neutra.

2.4.1.7. Resumen. Una de las características específicas de la modalidad oral según algunos (entre ellos, nosotros) era la expresividad, que brilla por su ausencia en lo escrito. La expresividad se refería al uso de la prosodia y de los gestos. La prosodia es lo que abordamos en este subapartado. Lo que nos interesaba era saber si la prosodia tiene algún impacto en la comprensión del lenguaje, si era relevante. Una vez garantizado que la prosodia no es algo extraordinario (sino que se produce en condiciones normales), repasamos evidencias de que, en efecto, la prosodia es una variable crítica. La prosodia facilita la adquisición de una lengua nueva, sirve para desambiguar oraciones, anuncia el fin de un enunciado liberando recursos atencionales y encarna las actitudes del hablante para con lo que está diciendo.

2.4.2. Gestos (manuales). Sabemos que la expresividad se materializa gracias al uso de la prosodia y otras claves paralingüísticas. Quizás las más importantes entre éstas otras son los gestos (manuales). Ahora nuestro cometido es averiguar si los gestos tienen algún peso en la comprensión del lenguaje.

2.4.2.1. Los gestos se producen. Al igual que hicimos con la prosodia, conviene hacer un par de aclaraciones previas. La primera es dilucidar si los gestos se producen en condiciones normales o si son un artefacto, sólo observable en circunstancias excepcionales. Además, debe aclararse si se producen para comunicar y no para otros fines.

Hay evidencias de que los gestos se producen espontáneamente, sin necesidad de que concurran circunstancias excepcionales. Un simple ejemplo. Bavelas et al. (1992) reunieron a un grupo de estudiantes universitarios (normales y corrientes). Les pidieron que contasen cómo podía sacarse un libro de la biblioteca de la universidad, un capítulo de una serie de dibujos animados que habían visto o una anécdota de sus vidas. El análisis de las grabaciones en vídeo de los relatos de los estudiantes reveló una media de unos 10 gestos por minuto (si refundimos distintos tipos de gestos en las condiciones de dos experimentos). Otro ejemplo es uno de los experimentos de Krauss (1998) del que hablamos hace un momento. En él había también participantes extraídos de una población normal y corriente que gesticulaban profusamente para relatar lo que veían en una fotografía.

Una cuestión controvertida, que se anticipó en alguna medida antes, es el objetivo con el que se producen los gestos. Una posibilidad es que quien produce gestos lo haga con un fin comunicativo, lo cual redundaría en un beneficio para el receptor (que contaría con más evidencias para interpretar lo que el hablante quiere decir). A ésta alternativa se suman varios investigadores. Bavelas es un ejemplo. En su estudio, antes referido (Bavelas et al., 1992), nos proporciona algunos indicios a favor de esta postura. En el primer experimento los participantes debían relatar cómo se toma prestado un libro de la biblioteca de la universidad y un capítulo de la serie de animación *Road Runner*. Hubo dos condiciones: en parejas y en solitario. Los primeros debían compartir el relato y los segundos lo hacían a solas. La tasa de producción de (cierta clase de) gestos fue significativamente mayor en la condición de pareja que en la solitaria. En el segundo experimento los participantes debían contar una anécdota triste y otra divertida sobre sus vidas y debían contárselo a otro participante situado frente a ellos. Hubo dos condiciones: una cara-a-cara y la otra con un biombo entre los participantes. Una vez más, la condición cara-a-cara mostró una tasa de producción gestual mayor que la de la condición con biombo. Si los gestos persiguiesen un fin no comunicativo, ¿cómo explicar que se produzcan más cuando tenemos un interlocutor que cuando no lo tenemos?

Otra posibilidad es que los gestos sirvan a funciones que favorecerían no al receptor sino al propio productor de los gestos. Una hipótesis, dentro de esta postura, es que la gesticulación le sirve al hablante porque mediante ella le es más fácil recuperar de la memoria algunos items léxicos. Recordemos que Krauss (1998) ofrecía todo un cuerpo de evidencias a favor de esa hipótesis (e.g., que la frecuencia de habla al describir un objeto es menor si nos prohíben hacer gestos al tiempo). Si bien, hay también evidencias que ponen en cuestión la hipótesis de los gestos como ayuda a la recuperación léxica. Alibali, Kita y Young (2000) controlaron los contenidos sobre los que sus participantes debían hablar pero variaron la complejidad de lo que debían decir. Todos debían decir cosas sobre unos mismos objetos (con sus correspondientes etiquetas léxicas) pero unos debían describirlos y otros debían explicar un fenómeno que se daba entre los objetos. Los explicadores producían más gestos que los descriptores. O sea, la gesticulación serviría no para recuperar léxico sino para expresar, dar forma a ideas complejas.

Otra hipótesis, pues, dentro de la posición no comunicativa es que la gesticulación favorece al hablante porque ayuda en la formulación de ideas complicadas. De hecho, hay evidencias que apoyan esta posibilidad. McNeill (1985) ofrece una completa revisión a este

respecto. McNeill apunta evidencias como éstas: la proporción de gestos producidos por un hablante es infinitamente mayor que la del oyente, la gesticulación se produce justo antes de la pronunciación de una frase y no después, los afásicos de Broca pierden la capacidad para combinar items verbales para crear un todo coherente y la de realizar ciertos tipos de gestos, el lenguaje verbal y los gestos se desarrollan en paralelo. Todas estas evidencias (las aportadas por McNeill y las de Alibali et al.) sugieren que los gestos asisten al hablante en la formulación de lo que quiere expresar o, dicho de otro modo, que los gestos son pensamiento como lo son las palabras. Esto es consistente con una concepción multirepresentacional de la mente (e.g., Karmiloff-Smith, 1992).

¿Con qué quedarnos entonces? Hemos visto, para empezar, que los gestos se producen en condiciones normales. Eso demuestra que son una realidad (como lo era la prosodia) y no un artefacto que aparece bajo circunstancias extraordinarias. En cuanto a si cumple una función comunicativa o de apoyo al hablante (recuperación léxica o formulación de ideas complejas) hemos comprobado que hay evidencias que apoyan cualquiera de las hipótesis. Para resolver este problema podemos razonar de la siguiente manera. Sea cual sea el objetivo de los gestos si (a) quienes están junto a quien los produce pueden captarlos e interpretarlos y (b) eso tiene algún efecto en el modo como interpretan los mensajes lingüísticos, entonces cabe decir que *terminan* desempeñando una función comunicativa. Aceptar esto supondría asumir que los algunos actos comunicativos son involuntarios (y en el primer capítulo dijimos que nos concentraríamos en los deliberados únicamente). Vamos a aceptarlo en este caso (y en el de la interacción encubierta, si recordamos). Por tanto, vamos a asumir que, aunque no se produzcan para tal fin, los gestos cumplen una función comunicativa si los receptores los captan y les ayuda a comprender mejor un mensaje. Ahora toca demostrar esas dos premisas.

2.4.2.2. *Los gestos se captan.* La segunda cuestión a aclarar antes de averiguar si los gestos tienen o no un impacto en la interpretación del lenguaje es si las personas somos capaces de captar e interpretar los gestos. Podría ocurrir que se produzcan e incluso que persigan un fin comunicativo y, sin embargo, que los receptores no sean capaces de servirse de ellos para hacer inferencias sobre el mensaje del hablante.

Alibali, Flevares y Goldin-Meadow (1997) diseñaron un experimento para poner a prueba esta cuestión. Seleccionaron deliberadamente a participantes sin ninguna clase de entrenamiento en codificación de gestos. Los participantes fueron mitad estudiantes universitarios, mitad profesores de Primaria y Secundaria. La tarea era identificar qué estrategia empleaban unos niños para resolver problemas matemáticos a partir de las explicaciones de los niños (basadas en discurso y en gestos). Así que lo primero fue reunir un conjunto de problemas aritméticos y plantearlos a un grupo de niños para que los resolvieran. Tras hacerlo se pidió a los niños que explicasen qué habían hecho para ello. Sus explicaciones, que se basaban tanto en palabras como en gestos, fueron grabadas en vídeo. Los vídeos resultantes fueron analizados por Alibali et al. con objeto de identificar las explicaciones redundantes e informativas de los niños. Las primeras eran las que decían lo mismo vía discurso y vía gesticulación. Las segundas eran las que presentaban alguna divergencia discurso-gesto, es decir, aquellas en las que los gestos aportaban información sobre cómo habían resuelto el problema aritmético. La manipulación experimental, por tanto, fue vídeo redundante versus vídeo informativo. Los estudiantes y profesores vieron las

grabaciones y manifestaron qué estrategia de resolución habían empleado los niños. Estudiantes y profesores fueron igualmente capaces de rendir bien en ambas condiciones. Esto es, son tan capaces de entender las explicaciones verbales de los niños como de entender sus gestos.

Resultados como éstos ponen en evidencia que las personas no entrenadas, las personas normales y corrientes, somos capaces de captar e interpretar los gestos que otra persona produce. Lo que falta por determinar (que es lo que haremos ahora) es si eso tiene un impacto en la interpretación del lenguaje.

2.4.2.3. *Gestos deícticos*. Al hablar por primera vez de los gestos manuales distinguimos entre gestos que aportan contenido proposicional al discurso y los que aclaran la intención o actitud del hablante. Distinguimos así entre gestos semánticos y pragmáticos. Los primeros, a su vez, podían subdividirse en categorías más específicas. Diferenciamos entonces entre gestos deícticos, icónicos y metafóricos. Los primeros son los que trataremos aquí. Recordemos que los gestos deícticos son los que señalan un objeto sobre el cual quiere llamarse la atención. Veamos si son o no relevantes consultando dos trabajos en los que los gestos se manipulan para propiciar un mayor aprendizaje.

Buisine y Martin (2007) pidieron a un grupo de estudiantes universitarios que aprendiesen cómo usar varios aparatos a partir de un material instructivo. El material consistía en imágenes que mostraban los componentes del aparato (e.g., mando a distancia, fotocopidora) acompañadas por descripciones verbales por parte de un agente pedagógico (o *animated pedagogical agent*). Hubo tres versiones del mismo material. Una en la que el agente señalaba (con gesto deíctico) un elemento del aparato al tiempo que lo nombraba (“el gran botón redondo del centro”), una en la que sólo había gesto y otra en la que sólo había nombrado. Tras consultar el material los participantes debían resolver varias tareas. En concreto, recordar toda la información posible sobre lo que habían consultado y hacer un dibujo del aparato incluyendo sus componentes. Los resultados mostraron que quienes habían recibido la versión con gesto y nombrado recordaban mejor que quienes habían recibido sólo gesto o sólo nombrado. En la tarea de dibujo no hubo diferencias significativas. Estos resultados indican que el gesto deíctico puede tener algún impacto en el modo como la gente comprende (en este caso, un material instructivo sobre aparatos).

Pero hay también evidencias de lo contrario. Craig, Gholson y Driscoll (2002) solicitaron a sus participantes que aprendiesen cómo se producen las tormentas a partir de un material. El material consistía en una animación (que mostraba nubes, masas de aire en movimiento, etc.) acompañada de una narración que describía los eventos mostrados. Había además varias versiones de un agente pedagógico que aparecía junto a animación. En unos casos el agente señalaba con su dedo (gesto deíctico) el aspecto concreto de la animación que la narración estaba describiendo; en otros, el agente aparecía junto a la animación sin más; en una condición control ni siquiera aparecía. Una vez consultado el material los participantes tenían que recordar todo lo posible sobre el material y resolver tareas de transferencia (e.g., “Imagina que ves nubes en el cielo pero no relámpagos. ¿Por qué no se ven?”). Pues bien, no hubo ninguna diferencia significativa entre los grupos en ninguna de las medidas. En contra de los resultados de Buisine y Martin, aquí los gestos deícticos tuvieron un valor nulo.

A la luz de los resultados de los estudios considerados aquí parece apresurado concluir que los gestos deícticos puedan tener un impacto. Siendo cautos los consideramos irrelevantes.

2.4.2.4. Gestos icónicos. Los gestos icónicos eran aquellos que representan la forma y tamaño de un objeto. Pongamos que uno está hablando sobre sus habilidades como pescador y comenta, ante el asombro de sus contertulios, que pescó un pez “así de grande” mientras coloca sus manos palma contra palma y las separa un trecho de unos 60 centímetros. Además de exagerar lo que está haciendo uno es emplear un gesto icónico. Son o no relevantes estos gestos.

Existe un experimento que aporta algunos indicios a favor del sí. Valenzeno, Alibali y Klatzky (2003) reclutaron a niños de cuatro años y medio y les presentaron una lección sobre la noción de simetría. La lección se presentaba en vídeo bajo dos condiciones. La mitad de los niños vio un vídeo en el que una profesora explicaba verbalmente, a partir de un set de figuras, por qué algunas eran simétricas y por qué otras no, detallando en cada caso el tipo de violación; la otra mitad vio el mismo vídeo pero con la profesora apoyándose no sólo en palabras sino también en gestos. Gestos, debemos aclarar, de tipo icónico, en tanto que especificaban la forma del objeto, el punto donde se encontraba su centro y la comparación entre lo que queda a cada lado del mismo. Después de ver el vídeo con la lección, los niños habían de evaluar la simetría de seis nuevas figuras (e.g., un reloj, una taza). Evaluar significaba decir (a) si es o no simétrica y (b) explicar por qué. Los resultados mostraron que si bien los niños de ambas condiciones rindieron igual al juzgar si cierta figura era o no simétrica, los de la condición con gestos ofrecían un porcentaje ligeramente más alto de explicaciones adecuadas que los niños de la otra condición. Esto significa que quienes vieron los gestos aprendieron levemente mejor que sus compañeros.

Aunque sugerente, los resultados del experimento no son, para nada, concluyentes. Un primer problema es que la ventaja de una condición sobre la otra es mínima (un 65% de explicaciones adecuadas contra un 55%, lo que ni siquiera se sometió a análisis estadístico). Un segundo problema es que la lección versaba sobre un contenido particularmente pictórico o gestualizable. No es que los gestos de la profesora hayan matizado o reforzado sus palabras sino que la propia lección eran sus gestos. Con las evidencias con las que contamos es nuevamente arriesgado concluir que esta clase de gestos sea relevante.

2.4.2.5. Gestos metafóricos. Esta categoría hacía referencia a los gestos que representan entidades abstractas. Un puño cerrado que indica perseverancia, una mano que gira para indicar redundancia, una mano que ondula para indicar relajación, etc.

En este caso contamos con las evidencias recogidas en estudio, con la limitación de que el tipo de gesto no estuvo bajo control. Es decir, en el estudio que vamos a repasar enseguida se puso a prueba el impacto de gestos metafóricos pero también de gestos de otro tipo, estando todos mezclados. Krauss, Morrel-Samuels y Colasante (1991) llevaron a cabo cinco experimentos sobre el valor comunicativo de los gestos. En todos ellos se empleó como material fragmentos de grabaciones en vídeo en las que personas normales y corrientes

hablaban sobre personas, fotografías o paisajes. Durante sus relatos las personas emitían espontáneamente toda clase de gestos, entre los que había metafóricos. El tercer experimento es de interés para lo que nos ocupa. Los participantes vieron fragmentos de esos vídeos bajo tres condiciones: voz y gestos, sólo gestos o sólo voz. Después de ver varios fragmentos se pedía a los participantes que reconociesen, dado un pool de fragmentos (algunos viejos y otros nuevos), cuál de ellos era el que anteriormente había visto y/o oído. Al comparar el reconocimiento demostrado por los participantes en cada una de las condiciones, Krauss et al. observaron que la condición sólo gestos era peor que cualquiera de las otras dos y que no había diferencias entre voz y gesto y sólo voz. En una palabra, ver los gestos no aportaba nada, la clave era escuchar la voz. Los investigadores interpretan este resultado (y otros paralelos en su set de experimentos) como prueba de que la gente interpreta lo que un hablante quiere decir sobre la base de sus palabras, sin que los gestos tengan mucho que decir al respecto.

Nuevamente hemos de concluir diciendo que los gestos metafóricos tienen escaso valor. Esto significa que la respuesta que andábamos buscando es no: ésta no es una variable crítica.

2.4.2.6. Gestos pragmáticos. A los gestos pragmáticos se les atribuyen diversas funciones comunicativas. Antes nos hicimos eco de algunas propuestas a este respecto: hay quien cree que sirven para encarnar la fuerza ilocutiva de un enunciado, otros piensan que indican el estatus que ocupa el enunciado en curso con respecto al resto del discurso, etc. Pero para nosotros no bastan las conjeturas: necesitamos saber si realmente afectan la interpretación de los oyentes de esa manera.

Que sepamos, no hay demostración empírica de tal cosa. Sólo disponemos de descripciones, muy detalladas eso sí, de los gestos y sus supuestas funciones. Kendon (1995), por ejemplo, realizó un análisis de cuatro gestos que son típicos en ciertas regiones de Italia. Uno de ellos es el *mano a borsa*, que consiste en juntar las yemas de los cinco dedos de la mano y orientar ésta hacia arriba meneando la muñeca levemente. De acuerdo con el análisis del autor, el *mano a borsa* hace explícito el acto ilocutivo concreto que el hablante realiza al pronunciar determinado enunciado; concretamente, revela que el enunciado pretende poner en cuestión, criticar o desafiar la validez de algo enunciado por otra persona. Pero ¿qué hay de cierto en todo eso? Como decíamos, no estamos en condiciones de responder.

2.4.2.7. Resumen. La expresividad se tomó por una de las características específicas de la modalidad oral. La expresividad, entendimos, puede incluir los gestos manuales además de la prosodia. En este subapartado lo hemos dedicado a los gestos. Hemos aclarado que los gestos se producen espontáneamente (en ausencia de circunstancias excepcionales) y que, además, personas no entrenadas son sensibles a ellos. La pregunta, como siempre, era si los gestos juegan un papel importante en la comprensión semántica y/o pragmática del lenguaje. La respuesta, a la luz de las evidencias, es que no. No disponemos evidencias suficientes de que los gestos, ya sean deícticos, icónicos, metafóricos o pragmáticos, jueguen un papel relevante.

2.4.3. Eventos perceptibles. En la modalidad oral tenemos acceso al mismo entorno

perceptivo que nuestro interlocutor. En esos casos, todo lo que rodea a la comunicación es percibido por ambos interlocutores y eso, asumían algunos (e.g., Clark, 1996), era importante para la interpretación del lenguaje. Ya sabemos lo que toca ahora.

Trueswell, Skerina, Hill y Logrip (1999) prepararon un sencillo setting para llevar a cabo su estudio. Colocaron varios objetos sobre una mesa y solicitaron oralmente a sus participantes que tomaran alguno de ellos y lo cambiaran de posición. Tan simple como eso. La sofisticación viene ahora: introdujeron ambigüedades en las instrucciones verbales y examinaron cómo éstas eran procesadas por los participantes registrando sus movimientos oculares. Un ejemplo de instrucción ambigua era: “Pon la rana sobre la servilleta en la caja²¹”. ¿Se refiere a la-rana-sobre-la-servilleta o quiere decir que coloque la rana primero sobre la servilleta y luego todo junto en la caja? Ahí reside la ambigüedad: son posibles tanto una interpretación tipo sintagma nominal (SN: “la rana sobre la servilleta”) como otra tipo complemento circunstancial de lugar (CCL: “sobre la servilleta en la caja”). Como control, se usaron también versiones inequívocas: “Pon la rana, que está sobre la servilleta, en la caja.”; que hace decantarse por la primera interpretación posible (la de SN). Otra fuente de aclaración fue el contexto: sobre la mesa los participantes podían encontrar una ó dos ranas. Cuando había dos, sólo una estaba sobre una servilleta. ¿Qué movimientos oculares cabe esperar en cada caso?

- Ante la instrucción ambigua y una única rana el sujeto no debería saber a qué atenerse: al no saber si la instrucción se interpreta como SN o como CCL, el sujeto habría de dirigir su mirada tanto a la rana como a la servilleta (además de a la caja, que deberá mirarse en toda condición).
- Ante la instrucción ambigua y dos ranas la instrucción se utiliza para desambiguar el contexto: el sujeto debería mirar hacia la rana sobre la servilleta y punto.
- Ante la instrucción inequívoca y una rana los ojos del sujeto deberían focalizar la rana sin titubeos.
- Finalmente, ante la instrucción inequívoca y dos ranas la instrucción servirá para decidirse sin problema por la rana pertinente: una vez más, el sujeto mirará la rana buena sin vacilar.

En el primer experimento los participantes fueron niños de cinco años; en el segundo, adultos. Veamos qué sucedió con los adultos. Trueswell et al. encontraron exactamente el patrón previsto que hemos desglosado hace un instante. Esto significa que las propiedades observables en la situación se utilizan también como claves para interpretar el lenguaje.

En el experimento de Trueswell et al. sin compartir el entorno perceptivo del hablante el oyente no habría podido reaccionar como lo hizo. De modo que sí hay evidencias a favor de la relevancia de esta variable.

2.4.4. Propiedades atribuibles al hablante y a la situación. El hecho de que nuestro interlocutor sea uno en particular y la situación comunicativa sea conocida por ambos interlocutores coloca a la modalidad oral ante otra posible ventaja. ¿Lo será? Aquí van unos

21 Quizá resulte extraña esta construcción en español. Por supuesto, en inglés es una oración completamente normal (aunque ambigua): *Put the frog on the napkin in the table.*

estudios sobre comunicación.

Hay evidencias que apuntan a que sí. Primero, es posible reconocer la intención comunicativa del hablante a partir del estatus que éste ostenta: no es igual lo que dice el jefe que lo que dice un compañero, por ejemplo. Esto es lo que Holtgraves (1994) confirmó experimentalmente. El investigador llevó a cabo cuatro experimentos en los que manipuló el estatus del hablante y la formulación de sus palabras y registró la facilidad con la que la intención de éste fue reconocida. Holtgraves presentaba a los participantes textos que narraban escenas cotidianas. En todos ellos un personaje pedía a otro realizar determinada acción: es decir, la intención comunicativa del primero era “solicitar” (en contraposición a “afirmar” o “comentar”; en otras palabras, regulatorio y no declarativo). Los sujetos leían textos como (36).

- (36) “La compañía constructora Carter estaba reparando carreteras rurales durante el verano. Hacía calor y todo el mundo bebía un montón de agua. Un día particularmente caluroso, el bidón de agua de la cuadrilla estaba casi vacío a media mañana. En el descanso el capataz/ un obrero le dijo a un/ otro obrero...
 - i. ¿Podrías rellenar el bidón de agua?
 - ii. El bidón de agua está casi vacío.”

¿Cómo se manipuló el estatus del hablante? Es fácil apreciar que el estatus del hablante se modificó intercambiando el emisor del fragmento final (i/ii): el hablante podía ser el capataz o un obrero. ¿Cómo se manipuló la formulación de sus palabras? Holtgraves utilizó además dos tipos de peticiones: unas expresadas en formato convencional (i) y otras en formato no-convencional (ii). El investigador consideraba que las peticiones en formato convencional (i) se interpretan fácilmente como tal. Si bien, las no-convencionales (ii) son ambiguas. El estatus del hablante sirve en estos casos de pista para facilitar (desambiguar) el reconocimiento de la intención. A la postre, los capataces tienen el derecho y la costumbre de dar órdenes a sus obreros, de forma que cabe esperar que un capataz pronuncie peticiones. Tras leer los escenarios los lectores habían de realizar varias tareas (dependiendo del experimento) que exigían haber aprehendido la intención comunicativa del hablante. En el experimento 2, por ejemplo, el investigador hizo que sus participantes leyeran una oración justo después de haber revisado los escenarios. Se esperaba que juzgasen si la oración era o no una construcción posible en inglés. Tales oraciones podían contener la expresión explícita de la solicitud (“rellena el bidón de agua”) o una idea cualquiera (“hoy hace mucho calor”). El tiempo (medido en mseg) empleado en decidir si la oración era o no correcta fue grabado. Los resultados revelaron que ante el formato convencional (i) la oración se verifica igual de rápido tanto si es el capataz el hablante como si tal posición la ocupa un obrero. Sin embargo, ante el formato no-convencional de la petición (ii) los participantes verificaron la oración más rápidamente cuando había sido enunciada por el capataz. Los resultados de los otros tres experimentos (que utilizan diferentes variables dependientes) fueron convergentes. Este hallazgo revela que cuando existe cierta ambigüedad (formato no-convencional), el conocimiento acerca del estatus del hablante sirve de pista que facilita la interpretación del enunciado. Nos sirve, además, para mostrar cómo las propiedades atribuidas al hablante es una fuente de información relevante para la comprensión pragmática.

Segundo, para determinar qué quiere decir un enunciado incluso los niños se sirven de pistas sobre el hablante. Cimpian y Markman (2007), cuyo estudio mencionamos en el capítulo 1, presentaron oraciones ambiguas a niños de cuatro años con la intención de que las interpretasen. Las oraciones eran como ésta: “no pueden ver los colores muy bien”; y se presentaban en un contexto, por ejemplo, un veterinario lo dice mientras se ve a dos perros en la escena (los niños veían las viñetas y se les leía el texto). Las oraciones podían interpretarse de dos maneras. Una genérica (“todos los perros ven los colores no muy bien, es propio de su especie”) y otra particular (“estos dos perros ven los colores no muy bien, es un problema que les he detectado a ellos dos”). La manipulación de Cimpian y Markman fue variar las características del hablante y su situación. La misma oración podía ser pronunciada, por ejemplo, por un veterinario en la clínica o por un profesor en una librería señalando una fotografía en un libro sobre perros. La cuestión era averiguar si los niños eran capaces de servirse de las pistas contextuales para interpretar en un sentido u otro la oración ambigua. Así ocurrió. Los niños realizaron preferentemente una interpretación genérica con el profesor y una particular con el veterinario (u otros personajes oportunos). De modo que incluso los niños de cuatro años son sensibles a las propiedades atribuibles a un hablante y las utilizan para interpretar el lenguaje.

En definitiva, la modalidad oral, según algunos, permite conocer las propiedades atribuibles del hablante y la situación. Hay pruebas de que eso beneficia a la comprensión del lenguaje.

2.5. Nivel interactivo.

2.5.1. Interacción abierta. Habíamos quedado en que, según algunos (e.g., Pickering & Garrod, 2004), en el lenguaje oral los dos interlocutores pueden intervenir y pueden hacerlo en el momento que se les antoje, lo que puede reportar beneficios. Decíamos que eso agiliza la comunicación por dos razones. Una, posibilita la negociación explícita de cualquier asunto que haya quedado oscuro. Dos, al hablar ambos interlocutores aportan evidencias de cuál es su estado mental, lo que le dice al otro qué rumbo seguir. Ahora es momento de comprobar si la posibilidad de mantener una interacción facilita la comprensión del mensaje.

Por un lado, hay evidencias de que la interacción abierta facilita la comunicación. Krauss y Wertheimer (1966) formaron parejas de participantes y les pidieron resolver una simple tarea. Uno de ellos tenía una lámina con una figura (muy enrevesada) y el otro una lámina en la que aparecían varias figuras (también muy enrevesadas) entre las que estaba la de su compañero. Lo que tenía que hacer el primero (instructor) era darle las indicaciones pertinentes al segundo (identificador) para que éste identificase la figura común dentro del set que tenía. Instructor e identificador se comunicaban a través de un teléfono. Una de las manipulaciones fue que sólo la mitad de las parejas podía mantener una interacción abierta; en la otra mitad de las parejas sólo el instructor podía hablar mientras el identificador escuchaba. La variable dependiente fue el número de palabras de las expresiones referenciales empleadas por el instructor. Por expresión referencial se entendía el grupo de palabras que uno usaba para referirse a cierto aspecto de una figura o a una figura (e.g., “¿has visto *el pedacito que sobresale bajo el pico derecho?*”). El resultado fue que los instructores de la condición de interacción abierta emplearon un número significativamente menor de palabras

en sus expresiones referenciales, en comparación con los de la comunicación restringida.

Por otro lado, disponemos de evidencias de que la interacción abierta facilita el aprendizaje. Un extenso y reciente estudio de VanLehn, Graesser, Jackson, Jordan, Olney y Rosé (2007) nos será muy útil para investigar este asunto. El estudio comprendía siete experimentos. En todos ellos se comparó la eficacia de distintas formas de enseñanza: tutoría, tutoría virtual y lectura de un texto. En todos los casos, grupos formados por alumnos universitarios tenían que aprender física aplicada (las leyes de Newton). Los alumnos asignados a la condición de tutoría (humana o virtual) mantenían sesiones de conversación abierta con su tutor: podían preguntar lo que necesitasen y obtener la correspondiente respuesta en tiempo real. El tutor también hacía sus preguntas, que eran respondidas en el acto por los alumnos. Los alumnos asignados a la condición texto consultaban materiales sobre (casi) el mismo contenido. Los investigadores crearon textos ad hoc basados en sesiones de tutoría prototípicas. Es decir, aunque los contenidos de las sesiones de tutoría no pueden mantenerse bajo control (ya que los participantes pueden preguntar lo que se les antoje y recibir respuesta contingente), sí cabe analizar las transcripciones para identificar los asuntos que típicamente son tratados. Con esa información es posible componer textos que reflejen lo que se discute (en promedio) en las tutorías dejando a los lectores en condiciones semejantes a las de los tutorizados. Ahora bien, los contenidos se les ofrecen en una forma fija; los tutorizados, sin embargo, pueden interpelar libremente a su tutor. Todos los alumnos resolvían diversas tareas al terminar la sesión de tutoría/lectura; en ellas ponían en juego sus conocimientos de física adquiridos. El análisis de resultados pone en evidencia un panorama desigual: cuando a los alumnos se les pedía que rindiesen a un nivel medio, los participantes tutorizados no obtenían mejores puntuaciones en las tareas (i.e., no comprendían más profundamente) que quienes habían leído; por el contrario, cuando se les pedía que rindiesen a un nivel experto, los lectores rendían peor que los tutorizados. Esto significa que, cuando se espera que aprendamos mucho más de lo que sabemos sobre una materia, la posibilidad de interactuar con nuestro interlocutor es crucial: posibilita un mayor aprendizaje.

Acabamos de evaluar si la interacción abierta es un factor crítico. Hemos visto que sí, puesto que agiliza la comunicación y favorece el aprendizaje.

2.5.2. Interacción encubierta. Dijimos que la interacción puede ser también encubierta. Esto es, aunque el oyente no pueda hablar, puede proporcionar (de forma voluntaria o no) pistas no verbales muy útiles para el hablante. El hablante puede así ajustar su discurso en consonancia con las necesidades que percibe. ¿Realmente sucede de esta manera?

Sin embargo, sabemos de un estudio sobre comunicación que invita a pensar que la interacción encubierta no juega un rol de peso. Clark y Krych (2004) solicitaron a un grupo de participantes que construyese diez modelos con fichas Lego. A una parte de los participantes se les asignó el papel de instructores; al resto se les asignó el de montadores. Los primeros debían guiar a los segundos en la construcción de los modelos Lego establecidos. Clark y Krych les entregaron los modelos a los instructores. Los montadores no tenían acceso a los modelos por lo que debían dejarse guiar por los instructores para poder replicarlos a partir de un puñado de fichas que les fueron entregadas. La interacción instructor-montador tuvo lugar

bajo distintas condiciones. Éstas se configuraron a partir de dos variables: (a) espacio de trabajo del montador visible/invisible para el instructor y (b) cara del montador visible/invisible para el instructor. Con visibilidad del espacio de trabajo nos referimos a la posibilidad de ver lo que el montador está haciendo: qué ficha selecciona, dónde la coloca, en qué posición la coloca, su dificultad para encontrar la ficha adecuada, etc. Ver o no la cara del montador significa tener acceso a sus expresiones faciales (e.g., cara de perplejidad), que podían funcionar como pistas cruciales. Las variables dependientes fueron, entre otras, el tiempo empleado para ensamblar los modelos, la corrección con la que éstos fueron contruidos y el tiempo dedicado a chequear el modelo Lego una vez construido. Los resultados mostraron que la visibilidad del espacio de trabajo del montador fue una variable significativa a diferencia del factor visibilidad de la cara del montador, que no tuvo un impacto significativo. Los montadores cuyo instructor tenía acceso visual a su espacio de trabajo (i.e., lo que estaban haciendo) requirieron menos tiempo para montar los modelos y chequear su corrección. En contra de lo que cabría esperar, la visibilidad de la cara de los montadores no fue de utilidad para los instructores. Podrían haber modificado sus instrucciones de acuerdo con la expresión facial de los montadores, lo cual habría redundado presumiblemente en un proceso constructivo más fluido. Pero no ocurrió así.

Para explicar los resultados de Clark y Krych pueden esgrimirse varios argumentos. Una posibilidad es que los montadores no fuesen muy expresivos. Sin embargo, sabemos que la producción de gestos (al menos manuales) es algo natural, que surge espontáneamente (e.g., Bavelas et al., 1992). Aunque también dijimos anteriormente que los gestos de la audiencia son siempre muchísimo menos frecuentes que los de quienes hablan (McNeill, 1985). Otra es que los instructores no fuesen muy sensibles a las expresiones de aquéllos. Pero sabemos que codificadores no entrenados son capaces de extraer información de los gestos de otros (e.g., Alibali et al., 1997). Otra posibilidad más es que la interacción encubierta sea irrelevante. Al fin y al cabo, dada la revisión que hicimos sobre el impacto de la gesticulación en los oyentes, no parece razonable pensar que el paralingüaje de la audiencia sirva para que el hablante acomode su discurso a las necesidades del momento. ¿Descartamos entonces el valor de la interacción encubierta? Hay una última posibilidad de salvarlo.

La interacción encubierta se refería a tener acceso a cierta información sobre la audiencia. Esa información puede ser paralingüística, en cuyo caso no sirve de mucho, como acabamos de ver. Pero puede ser también de otro tipo, como también se anticipó. Si tenemos pistas suficientes sobre qué puede saber y qué no nuestra audiencia (a partir de su edad, nivel socio-económico, etc.), podremos adaptar el nivel de conocimiento exigido para entender lo que decimos y conseguir así ser más comprensibles. Y sabemos que adaptar nuestro discurso al nivel de conocimientos del otro redundará en una comunicación eficaz y eficiente. Nückles, Wittwer y Renkl (2005) reunieron a un grupo de expertos y novatos en informática. Se trataba de que los expertos enseñasen informática a los segundos a través de un sistema informático de mensajería instantánea tipo *MSN Messenger*. Los novatos tecleaban unas preguntas básicas (fijadas por los experimentadores) sobre informática y las enviaban a los expertos, quienes debían responderlas. Hubo dos condiciones. En una los expertos tenían a su disposición una ficha que detallaba el nivel de conocimiento que el novato que le habían asignado tenía sobre distintos aspectos (e.g., “sabe qué es un módem pero no cómo funciona exactamente”). En otra el experto no sabía nada sobre su novato asignado. Recibida la respuesta del experto a sus

preguntas, el novato podía hacer alguna pregunta adicional a voluntad. Los resultados mostraron que la condición con ficha fue más eficaz y eficiente que la otra. ¿Por qué? Los de la condición con ficha aprendieron más, tal como indicaban unos ensayos que elaboraron tras la fase de interacción. Éstos también plantearon muchas menos preguntas *follow-up* de las que se formulan para aclarar algún asunto confuso. Así que si la interacción encubierta nos brinda información sobre los conocimientos de la audiencia, esto permite un ajuste entre interlocutores que favorece la comunicación.

Resumiendo, la interacción encubierta como acceso a las pistas gestuales de la audiencia no parece ser muy útil. Sin embargo, si la interacción encubierta nos da acceso a información sobre los conocimientos que posee la audiencia sobre cierto tema, esto agiliza la comunicación.

2.6. Impacto en el pensamiento.

Varios psicólogos (Morais et al., 1986; Olson, 1977) había planteado que la modalidad escrita tiene una característica específica singular: modifica nuestro pensamiento. Lo hace porque los sonidos lingüísticos pasan a ser fonemas y porque los relatos e intuiciones se convierten en explicaciones y argumentaciones. Al margen de que pongamos esto en cuestión o no, lo cierto es que son cuestiones tangenciales a la que nos interesa. No son efectos sobre la interpretación del lenguaje sino efectos colaterales sobre una cultura. De modo que el impacto sobre el pensamiento queda fuera del catálogo por impertinente.

Por otra parte, es posible matizar el planteamiento de Olson. Olson mantenía que la escritura permite llevar a cabo un pensamiento sostenido y exhaustivo sobre un tema dado, lo que permite confeccionar argumentaciones sólidas. La escritura es, desde este punto de vista, una herramienta de pensamiento. Y es cierto que la escritura puede representar un instrumento de pensamiento muy potente. Al vernos forzados a expresar nuestro pensamiento de forma que sea inteligible y de que nuestras afirmaciones sean consistentes entre sí, matizamos, corregimos, desarrollamos nuestras ideas. Scardamalia y Bereiter (1987) llaman a esto mismo *transformar el conocimiento*. Ellos conciben el proceso (maduro) de composición de un texto como una interacción entre los espacios de contenido y retórico. El primero engloba todas las ideas que poseemos acerca de un tema concreto. El segundo tiene que ver con cómo expresar esas ideas para conseguir el efecto deseado. Por ejemplo, si el escritor desea plantear una afirmación de forma convincente (demanda que se produce dentro del espacio retórico) se verá abocado a hacer explícitas las razones que sustentan tal afirmación, buscar analogías para ilustrarla, rebajar la importancia de posibles contraargumentos, etc. (dentro del espacio de contenido). Como fruto de tal actividad su conocimiento se verá ampliado o transformado (utilizando su terminología). El problema es que no todo escritor es maduro. Los escritores inmaduros exhiben otra estrategia bien diferente de la anterior: *decir el conocimiento*. De acuerdo con esta estrategia, el escritor emplea el tópico del que desea contar algo (e.g., “los ordenadores portátiles son peores que los de sobremesa”) como clave para localizar una serie de ideas relacionadas (e.g., “capacidad del disco duro”, “calentamiento de la CPU”, “ampliación de memoria”). Estas ideas, a su vez, sirven para activar un proceso asociativo de búsqueda en la memoria. Las ideas activadas propagan su activación y toda una red de ideas asociadas pasa a estar disponible. La información activada puede ser también de tipo retórico.

En nuestro ejemplo, se activaría el esquema de tipo “comparación”, que permitiría que fórmulas como “en cambio” o “a diferencia de” estuviesen disponibles. Este proceso de propagación de la activación daría lugar a la generación de un fragmento de texto (“los portátiles se calientan rápidamente a diferencia de los de sobremesa que cuentan con ventiladores potentes”). El fragmento producido serviría nuevamente como clave de activación en la memoria. La composición final podría demostrar cierta coherencia, en virtud de la asociación que existe entre toda la información plasmada en el texto. No obstante, el escritor no se ha visto obligado a hacer explícito su pensamiento, a reconsiderarlo, ampliarlo o modificarlo. O sea, la escritura puede ser una herramienta del pensamiento pero no lo es necesariamente. Eso sí, si pretendemos pensar sostenida y exhaustivamente sobre un tema, lo escrito nos proporciona una ayuda inestimable.

2.7. Balance.

Recapitulemos. Intentábamos encontrar las características específicas de la modalidad oral y de la escrita. Para nosotros (y Ferreira & Anes, 1994) la característica específica de lo oral es la expresividad, siendo el control la de lo escrito. Para Gough y otros la única característica que diferencia las modalidades es que lo escrito exige decodificación y lo oral no. Otros tantos han señalado otras características muy distintas. Integrando las aportaciones de psicólogos y lingüistas confeccionamos un catálogo que sumaba 15 dimensiones de variación, o sea, cada modalidad tiene no una ni dos sino 15 características específicas que la diferencian de la otra. Tal discrepancia en las caracterizaciones de cada modalidad nos hizo pensar que era necesario cribar.

Lo que hemos hecho en esta segunda sección es examinar si cada característica identificada es o no una variable crítica en la comprensión semántica y/o pragmática del lenguaje. El plan era descartar las que no lo eran, las variables irrelevantes.

Si escrutamos la tabla 13 podremos apreciar que, en general, todas las supuestas diferencias entre la modalidad oral y la escrita son potenciales factores críticos. Sólo la gesticulación (que es la mitad de la expresividad) y el impacto en el pensamiento han demostrado ser irrelevantes. ¿Y el resto? Todas, de algún modo u otro, han demostrado afectar a la comprensión del lenguaje. En unos casos favorecen o perjudican que podamos integrar todos los contenidos que se nos proporcionan. En otros, favorecen o perjudican que podamos saber qué pretende el hablante con lo que dice (i.e., regular o declarar) o conocer la actitud que alberga para con lo que dice (e.g., no está seguro).

Tabla 13. Las supuestas diferencias oral/escrito a examen: ¿son relevantes?

Diferencias	Impacto en la comprensión del lenguaje
<i>Canal sensorial</i>	Relevante Si hay que procesar palabras e imágenes, el canal importa: perjudica o favorece el aprendizaje
<i>Control</i>	Relevante Quienes controlan el input lo comprenden mejor
<i>Decodificación</i>	Relevante Decodificar mejor favorece la comprensión
<i>Densidad léxica</i>	Relevante Obliga a disponer de vocabulario
<i>Elementos auxiliares</i>	Relevante La disfluencia "eer" afecta al procesamiento de palabras
<i>Integración</i>	Relevante Lo denso impone una carga adicional y perjudica la comprensión
<i>Densidad informativa</i>	Relevante Cuanto más complejo es un material peor se comprende y más esfuerzo se invierte
<i>Indicación idea central</i>	Relevante Si se hace explícito el meollo comprendemos mejor
<i>Expresividad, prosodia</i>	Relevante Facilita adquisición lengua, facilita desambiguación oraciones, anuncia fin de enunciado, refleja actitud del hablante
<i>Expresividad, gestos</i>	Irrelevante No hay evidencias suficientes de que los gestos, sean cuales sean, favorezcan la comprensión
<i>Entorno perceptivo</i>	Relevante No hay forma de entender ciertos enunciados si no es compartiendo el entorno perceptivo con el hablante
<i>Propiedades situación/hablante</i>	Relevante Si sabemos quién dice qué y en qué circunstancias comprendemos mejor el lenguaje
<i>Interacción abierta</i>	Relevante Que los dos interlocutores hablen libremente agiliza la comunicación y la hace más eficaz
<i>Interacción encubierta</i>	Relevante Si el hablante accede a información sobre la que la audiencia sabe la comunicación se vuelve más eficaz y eficiente
<i>Impacto en el pensamiento</i>	Irrelevante Se trata de efectos colaterales, no relacionados con la interpretación del lenguaje

Lo que pretendíamos era reducir la lista y no lo hemos conseguido ni de lejos. Cabe proponer una segunda criba. Vamos a someter las diferencias a un nuevo examen: ¿son realmente específicas las características identificadas? Nótese que habíamos sido explícitos al definir qué entendemos por específico. Inherente y exclusivo, dijimos. Ahora se trata de dilucidar si lo que algunos llaman específico es realmente inherente y exclusivo.

3. ¿Son realmente específicas las características identificadas?

Al señalar el control y la expresividad como características específicas de lo escrito y lo oral, respectivamente, lo hicimos con criterios. Por específico entendemos (a) inherente y (b) exclusivo de cada modalidad. Por inherente, a su vez, entendemos que cierta propiedad aparece en todas las variantes de una modalidad dada, al margen de factores coyunturales. Por exclusivo, que aparece en una sola modalidad y nunca en la otra. Apliquemos los criterios a las características del catálogo.

3.1. ¿Inherentes? Factores coyunturales.

En la primera sección confeccionamos un catálogo de características específicas o diferencias entre modalidades. Para ello nos hicimos eco de las aportaciones de un puñado de lingüistas y psicólogos. Todas las diferencias pueden existir, de hecho, entre las modalidades y, a tenor del resultado de la revisión crítica de la segunda sección, todas son relevantes. La cuestión es que quizá no todas dependen exclusivamente de la modalidad sino de la presencia de algunos factores coyunturales. Vemos acto seguido cuáles son esos factores y qué características se considerarán no inherentes a la luz de los factores.

3.1.1. Presencia de imágenes. El canal sensorial es inherente a cada modalidad. De eso no hay duda. Puede ser un post-it pegado en la puerta de la nevera, un e-mail o un anuncio en un periódico. Todo entra por vía visual. Puede tratarse de una conversación telefónica, una charla en una cafetería o una video-conferencia. El lenguaje penetra en nuestro cerebro vía auditiva. Ahora bien, si el canal sensorial era relevante era porque, cuando debíamos aprender de materiales con información verbal y pictórica, lo escrito agotaba (o inutilizaba parte de) los recursos de la memoria de trabajo y por eso aprendíamos menos. Resulta que no siempre aprendemos de ese tipo de materiales o, más generalmente, no siempre procesamos información verbal y pictórica al mismo tiempo.

¿Qué pasa si sólo información verbal es presentada? ¿Importa que entre por vía visual activando la agenda o que entre por la auditiva activando el bucle? Creemos que no. Esto es, si no hay imágenes consumiendo capacidad de procesamiento de la agenda visoespacial, no habría de haber ventaja ni desventaja por parte del bucle fonológico. Si antes nos interesaba utilizar el bucle era para evitar la sobrecarga de la agenda, pero no por un beneficio intrínseco al utilizarlo. En definitiva, sin imágenes el canal sensorial es irrelevante.

Hay, sin embargo, una posible objeción a lo que acabamos de plantear. Penney (1989) realizó una revisión de estudios en memoria humana que ponen de relieve una superioridad del bucle sobre la agenda. Dicho de otro modo, la gente recuerda mejor si la información les fue presentada en modalidad oral que si les fue presentada por escrito porque, defiende Penney, el bucle fonológico deja una huella más poderosa que la agenda. La interpretación puede ser discutible pero, nos tememos, la evidencia empírica es incontestable: la ventaja oral sobre la escrita en el recuerdo se ha replicado un número considerable de veces. Por consiguiente, cuando se trata de memorizar sin presencia de imágenes el canal sensorial

importa.

Ahora, la tarea de los participantes en los experimentos que repasa Penney es memorizar listas sin sentido. En los experimentos de memoria humana se presentan listas de palabras sin relación semántica alguna o listas de números al azar y los participantes disponen de un tiempo para memorizarlas. Más tarde se les pide que traten de recordarlas (por sorpresa o sabiéndolo de antemano, en el orden de presentación o en orden libre, con claves de recuerdo o sin ellas, etc.). Pero nada de eso es comprender el lenguaje. Interpretar qué nos quieren decir, qué pretenden con lo que dicen, etc.; todo eso queda fuera de las tareas de estos experimentos. No hay razón para suponer que una huella más poderosa asociada al canal auditivo vaya a redundar en un recuerdo mayor de un discurso de tres páginas o en un reconocimiento de la intención comunicativa más eficiente. Esto es, aunque al memorizar el canal importe eso no tiene nada que ver con la comprensión del lenguaje: seguimos manteniendo que sin imágenes el canal es irrelevante.

En conclusión, la importancia del canal sensorial depende de la presencia de imágenes o de la tarea a realizar. En ausencia de imágenes y cuando la tarea no es memorizar, el canal sensorial no es una variable relevante. Luego no es un problema inherente a lo escrito, sino que depende de factores coyunturales. (Aún así en el próximo capítulo emplearemos los términos auditivo y visual por las razones que se entenderán en su momento).

3.1.2. Cuando la decodificación se automatiza. La modalidad escrita exige la *decodificación* de las palabras impresas. Es decir, obliga a convertir patrones gráficos en representaciones ortográficas, fonológicas y léxicas. Nada de eso es necesario en la modalidad oral. Esta diferencia era para Gough y la simple view of reading (Gough et al., 1996; Hoover & Gough, 1990) central: es lo único que diferencia lo escrito de lo oral.

Lo cierto es que, sea o no la única, sí demostró ser una variable relevante. Los niños entrenados en decodificación son más capaces de comprender textos (Tan & Nicholson, 1997) y los malos comprendedores rinden mucho peor que los buenos en tareas que exigen decodificar (Bell & Perfetti, 1994). Luego la decodificación determina la comprensión lectora.

Es cierto. Pero no es menos cierto que su papel va decayendo a medida que uno gana experiencia con lo impreso. Hoover y Gough (1990) realizaron un pequeño meta-análisis de los resultados de un set de estudios. Los resultados de su análisis mostraron que, cuando se consideran los estudios con niños de escuela Primaria, hay alta correlación entre decodificación y comprensión lectora pero nula correlación entre comprensión lectora y comprensión oral; cuando se contemplan los estudios con estudiantes de Secundaria, la correlación comprensión lectora-comprensión oral es altísima. Esto significa que en los primeros años como lector la decodificación determina la comprensión lectora y a medida que uno gana experiencia la decodificación pierde peso.

También Bustos (en preparación) ha encontrado algo similar recientemente. La investigadora tomó varias medidas de un grupo de estudiantes de 12 años. Les pidió leer un texto de geografía para responder a varias preguntas sobre el mismo más tarde. Antes, les

midió en decodificación (velocidad y precisión en lectura de palabras y pseudopalabras), en conocimientos generales sobre geografía, en memoria de trabajo y en capacidad para servirse de los elementos retóricos del texto (e.g., “en primer lugar... en segundo lugar”). La comprensión del texto de geografía evaluada a través del rendimiento en las preguntas se utilizó como variable criterio en una regresión. El resto, como predictoras. La decodificación, considerando al mismo tiempo todas las medidas (velocidad y precisión de palabras y pseudopalabras), explicó menos del 1% de la varianza en comprensión. O sea, la comprensión de los estudiantes de 12 años (al menos los de la muestra de Bustos), no depende de su capacidad para decodificar. Depende de las otras variables (e.g., memoria de trabajo, conocimientos sobre geología) pero no de la habilidad para acceder al léxico de palabras impresas.

Esto significa que hay un factor coyuntural que modula el efecto de la decodificación: la automatización. Con suficiente experiencia con lo impreso el lector medio (no el lector con problemas específicos, como los de Bell & Perfetti (1994)) logra dominar la decodificación hasta el punto de que no repercute en su comprensión (e.g., el 1% de Bustos). En consecuencia, vamos a considerar que, despejado el factor automatización, el peso de la decodificación no es inherente.

3.1.3. Cuestión de género: lo expositivo. Los textos científicos, según Goldman y Bisanz (2002), desempeñan distintos roles y emplean distintos géneros. Entre los roles está el compartir información entre científicos, divulgar ciencia y preparar a los estudiantes para entrar en la comunidad científica. Los géneros, que es lo que nos interesa, son formas especializadas de comunicación.

Pues bien, en el campo de la comprensión de textos se suele distinguir entre los géneros narrativo y expositivo (e.g., Graesser, Millis & Zwaan, 1997). Los primeros son textos que relatan experiencias comunes (e.g., chica-conoce-chico); los segundos son textos alejados de la experiencia cotidiana (e.g., geología). Lo capital es que cada género se sirve de un estilo particular. El narrativo pone el acento en la acción, es algo ambiguo y su orden refleja el orden natural de los acontecimientos. Éste se corresponde con el lenguaje natural o cotidiano. El expositivo es más formal, resalta las causas, intenta ser totalmente explícito y su organización es taxonómica. Se corresponde con el lenguaje académico o científico (Bruner, 1986).

(Un inciso. Aprovechamos para advertir que en esta tesis empleamos el término “narración” para designar la información verbal presentada oralmente y “texto” para la presentada por escrito. Y ello al margen del género. Lo hacemos así siguiendo la nomenclatura de Mayer (2001). Hasta aquí esta distinción no ha sido central. En lo que sigue será importante no perderla de vista para evitar confusiones. Y será así porque a partir de ahora la distinción sí sera central).

Retomando la cuestión de los géneros, el problema es que género y modalidad suelen confundirse. Es decir, lo escrito y lo expositivo se equiparan y lo oral y lo narrativo también. Se identifica, pues, texto escrito con estilo formal, académico. ¿Qué implica esa confusión? Pues que muchos de los rasgos considerados propios de lo escrito dependen, en realidad, del

género y no de la modalidad. A lo escrito se le atribuía *densidad léxica*, ausencia de *elementos auxiliares*, *integración*, *carga informativa* y una *indicación explícita de la idea central*. Pero, ¿son realmente propiedades inherentes a lo escrito?

Chafe y Danielewicz (1987) reunieron 100 pequeños monólogos extraídos de conversaciones informales, 100 monólogos procedentes de discusiones en clases universitarias, 100 cartas personales y 100 artículos de investigación. Las dos primeras categorías son orales y las dos segundas, escritas. Al comparar lingüísticamente las conversaciones con los artículos los resultados son un calco de lo que presentamos líneas arriba. Esto es, lo escrito presenta las características mencionadas hace un instante y lo oral, las contrarias. Lo interesante surge del análisis de las categorías restantes: las discusiones en clase y las cartas personales. Al analizarlas Chafe y Danielewicz descubrieron un acercamiento de posiciones. Las cartas eran como conversaciones mientras las discusiones eran como artículos. Dicho llanamente, lo escrito era como lo oral y lo oral era como lo escrito. Si no depende de la modalidad, ¿de qué dependían las diferencias? Ya lo hemos advertido: del género. Obviamente, cartas y conversaciones pertenecen al terreno de lo no expositivo y discusiones y artículos al terreno de lo expositivo.

De lo que acabamos de exponer se desprende una conclusión. La densidad léxica, la ausencia de elementos auxiliares, la integración, la carga informativa y la indicación explícita de la idea central dependen del género (expositivo o no expositivo) y no de la modalidad. Luego no son características inherentes de lo escrito.

3.1.4. Conversación cara a cara. En la sección anterior comprobamos que compartir el *entorno perceptivo* con el hablante, conocer ciertas *propiedades del hablante y la situación comunicativa* y la posibilidad de mantener una *interacción* (abierta y encubierta) eran variables de peso. Todas ellas favorecen la interpretación del lenguaje. Y, más importante, todas pertenecían supuestamente a la modalidad oral. Pero ¿es de verdad de este modo?

Si hace un instante hemos dicho que modalidad y género se confunden, ahora lo que vamos a defender es que modalidad y *setting* (o condiciones particulares de una situación) se confunden también. La modalidad oral se utiliza en distintos settings. Lo oral está presente en una conversación telefónica, en una lección magistral, en una película de vídeo, en una conversación de sobremesa, en un meeting, etc. Todos son settings orales. El problema es que la modalidad oral explícita o tácitamente se reduce al setting de la conversación cara a cara.

Algunos lo hacen de manera explícita. Clark (1996), por ejemplo, estudia la conversación cara a cara porque la considera la forma de comunicación básica. Clark lo argumenta diciendo que (a) es el primer setting en el que uno participa (i.e., madre/padre e hijo), (b) está presente en todas las sociedades y (c) participamos en ella sin necesidad de poseer habilidades especiales (Lieberman (1992) argumenta de la misma forma). Pickering y Garrod (2004) estudian la conversación cara a cara porque están específicamente interesados en desentrañar los procesos que explican su funcionamiento.

Otros, sin embargo, asumen directamente que la modalidad oral y la conversación cara a cara son lo mismo. Por ejemplo, Sánchez, García y Bustos (en prensa) asumen

implícitamente que comprender lo escrito es más difícil que comprender lo oral porque en lo primero no se cuenta con la interacción abierta propia de lo segundo. La interacción abierta es propia de la conversación cara a cara pero no de la película de vídeo o de la lección magistral.

Si aceptamos que modalidad oral y el setting específico de la conversación cara a cara no son lo mismo, entonces debemos descartar más características. Compartir el entorno perceptivo con el hablante, conocer ciertas propiedades del hablante y la situación comunicativa y la posibilidad de mantener una interacción son todas variables relevantes que dependen del setting, no de la modalidad. Por eso no pueden considerarse inherentes.

3.1.5. La expresividad y el control en distintas coyunturas. Nos falta por saber qué sucede con dos características. Justo las que nosotros y Ferreira y Anes (1994) seleccionamos como específicas de lo oral y lo escrito: la expresividad oral y el control escrito.

Consideremos las siguientes situaciones comunicativas (con géneros y settings que varían). Una película, una grabación en el contestador automático, una charla en una cafetería, una lección magistral, una ponencia en un meeting, una conversación a través del portero automático, una conversación a través de *Skype*, la defensa de una tesis. Todas tienen algo en común. Todas emplean la modalidad oral. Y algo más. Todas conservan la expresividad (prosódica). Hay énfasis, fraseo, tono y ritmo en todas ellas. Puede que veas la cara de tu interlocutor o no, puede que se utilice un lenguaje informal o académico, puede que el oyente tenga permitido hablar o no y siempre queda la prosodia.

Consideremos estas otras situaciones comunicativas. Una nota en un tablón de anuncios, una conversación a través del *Messenger*, una carta íntima, un artículo en la prensa de hoy, un post-it en el marco del monitor del ordenador, un e-mail, una novela, el capítulo de una tesis. Todas presentan puntos en común. Para empezar, la modalidad escrita. Pero además en todas ellas es posible controlar el input. Uno puede fijar su mirada en el segmento de interés del texto, puede leerlo despacio o deprisa, puede volver a leer una vez más. Y eso aunque el género no sea expositivo.

Con esto queremos demostrar que no hay coyuntura que afecte al control y la expresividad. Están presentes siempre, luego son inherentes a cada modalidad.

3.1.6. Balance. Tras el repaso a todo un arsenal de propuestas configuramos un catálogo de características específicas de lo oral y lo escrito. Casi todas demostraron tener peso en la comprensión del lenguaje y decidimos plantearnos si eran realmente inherentes a cada modalidad. Acabamos de ver que algunas no son inherentes, en la medida en que dependen de factores coyunturales. Ahora hacemos balance de las descartadas por coyunturales.

La importancia del canal sensorial depende de si hemos de procesar imágenes junto con las palabras: no es inherente. La decodificación importa sólo si no se ha dominado: no es inherente. La densidad léxica, la ausencia de elementos auxiliares, la integración, la carga informativa y a indicación explícita de la idea central dependen de si el género es expositivo o

no: no son inherentes. El entorno compartido, las propiedades atribuibles al hablante y la situación y la interacción dependen de si el setting es o no la conversación cara a cara: no son inherentes. Con esto hemos descartado todas las características excepto dos: el control y la expresividad (prosodia).

Ahora queda por saber si, además de inherentes, estas características son exclusivas de cada modalidad. Es decir, ¿puede lo escrito ser expresivo y lo oral tener control?

3.2. *¿Exclusivos? La expresividad escrita y el control oral.*

¿Puede lo escrito ser expresivo? Hay que admitir que la modalidad escrita incluye una serie de marcas llamadas *signos de puntuación* que sirven para encarnar la expresividad (prosodia) oral. Así, lo escrito (en castellano) cuenta con tildes, signos de exclamación, de interrogación, comas, puntos y puntos suspensivos. Cada cual simula los fenómenos prosódicos que conocemos. Las tildes indican énfasis. La exclamación e interrogación nos dan pistas sobre el fraseo. Las comas y puntos reflejan pausas. La pregunta es si todos esos signos igualan la prosodia oral.

Hay varias razones para pensar que no. Para empezar, uno debe decodificar esos signos. Eso exige poseer la habilidad oportuna, que no todos han dominado. De hecho, la prueba estandarizada de comprensión lectora PROLEC-R (Cuetos, Rodríguez, Ruano & Arribas, 2007) incluye una tarea de interpretación de signos de puntuación. Los evaluados deben leer un texto que contiene fragmentos como éste: “¡Qué suerte! - exclamó Juan. ¡Cómo me gustaría volar!” y deben interpretarlo como corresponde. Los alumnos de 1º de Primaria, por ejemplo, tienen un porcentaje de error del 47%. Así, no todos saben cómo decodificar los signos de puntuación. En segundo lugar, la interpretación que uno hace de los signos no refleja la riqueza de la prosodia. Es decir, los signos dan pistas pero no agotan la interpretación. Unos puntos suspensivos señalan una pausa de tres, cuatro, cinco, ¿cuántos segundos? Tercero, hay fenómenos prosódicos específicos que no son codificados por los signos de puntuación. Por ejemplo, el rango tonal no puede representarse, tampoco el alargamiento de final de frase (del que ya hablamos; Ferreira, 1993). Por todo ello pensamos que lo escrito no puede ser nunca tan expresivo como lo oral. Luego la expresividad es propiedad exclusiva de la modalidad oral.

¿Puede lo oral ser controlable? Hay settings en los que la modalidad oral permite la repetición del mensaje. Por ejemplo, si lo que tenemos entre manos es una grabación en cassette, siempre podemos rebobinar y escuchar de nuevo un pedazo (e.g., Mayer (1983) permitió a sus participantes escuchar de nuevo los cassettes para favorecer la comprensión). Si estamos inmersos en una conversación cara a cara podemos, haciendo uso de la interacción abierta, solicitar que nos repitan algo. Circunstancias como éstas hacen dudar de la exclusividad del control.

Sin embargo, como sabemos el control es algo más. El control permite regular el ritmo de presentación y enfocar en cada momento el segmento de interés (sea éste un grafema, una palabra o una frase). Como sabemos, los patrones de inspección de un texto son intrincados, nuestra mirada se desplaza a través de lo impreso de una forma muy peculiar. Dedicamos más

tiempo a las palabras contenido (Just et al., 1982), nos detenemos en las frases y oraciones importantes (Hyönä et al., 2002), destinamos tiempo extra a las oraciones densas (Kintsch & Keenan, 1973). Y más todavía, sabemos que mientras leemos una palabra extraemos información útil de la siguiente, gracias al procesamiento parafoveal, y eso facilita el procesamiento de la esa palabra cuando le llega el turno (e.g., Rayner et al., 2005). O sea, autoadministramos la consulta de un texto de forma tan microscópica que nunca la modalidad oral podría otorgar tal nivel de control. Luego el control es exclusivo de lo escrito.

En conclusión, lo escrito no puede ser expresivo y lo oral no puede ser controlable. Lo escrito no puede ser expresivo porque sus signos de puntuación no igualan la prosodia oral. Lo oral no otorga nunca tanto control como lo escrito porque la autoadministración que éste proporciona es microscópica.

3.3. ¿Descartamos definitivamente las otras variables identificadas?

De la recta final de nuestra criba podría inferirse (erróneamente) que ciertas variables no son de interés. Nosotros buscábamos las características que fuesen inherentes y exclusivas de cada modalidad. Aunque muchas otras son relevantes, sólo éstas cumplen los criterios de inherencia y exclusividad. Por eso descartamos las demás. Ahora bien, eso no significa que las demás no puedan ser de interés para otros investigadores.

Si alguien está interesado, por ejemplo, en la diferencia entre la conversación cara a cara y la lectura de un texto expositivo sobre el mismo tema, habrá de jugar con las variables oportunas. Variables que son más que el control y la expresividad. Aquí el contraste sería entre expresividad, contexto compartido e interacción por un lado versus control y densidad por el otro. Si alguien está interesado en comparar la comprensión de un relato presentado oralmente con una conversación telefónica sobre el mismo tema, está en su pleno derecho. Ahí el interés recaerá probablemente en la presencia/ausencia de interacción.

Lo que nosotros hemos hecho es asegurarnos de cuáles son las características específicas de cada modalidad, que no dependen de coyunturas particulares y son exclusivas. De paso, hemos hecho un análisis de las potenciales diferencias entre lo oral y lo escrito. Si algo debe sacarse en limpio de este capítulo, además de cuáles son las características específicas de cada modalidad, es que ahora sabemos exactamente qué estamos comparando si jugamos con las modalidades. Es decir, el catálogo puede ayudar a determinar dónde radica la diferencia entre dos situaciones comunicativas dadas.

4. Resumen, predicciones, consecuencias.

El marco de investigación que proponemos, la aproximación funcional, descansa sobre varios supuestos. El segundo supuesto es que cada modalidad del lenguaje tiene características específicas que, además, pueden resultar ventajosas. Dijimos que estas características son el control escrito y la expresividad oral. En el capítulo tuvimos que argumentar por qué esas y no otras son las características específicas.

Al comienzo descubrimos que no todos señalan las mismas características como específicas de cada modalidad. Así, Gough y la simple view of reading (Gough et al., 1996; Hoover & Gough, 1990) defendían que lo único que distingue lo oral de lo escrito es que lo segundo exige decodificar y lo primero no. Un grupo de lingüistas y psicólogos había propuesto unas cuantas características más. Integrando todas las aportaciones elaboramos un catálogo de diferencias.

El catálogo constaba de 14 dimensiones de variación. El canal sensorial (visual en lo escrito, auditivo en lo oral), la decodificación (necesaria en lo escrito, no en lo oral), el control (presente en lo escrito, no en lo oral), la densidad léxica (presente en lo escrito, no en lo oral), los elementos auxiliares (presente en lo oral, no en lo escrito), la integración (presente en lo escrito, no en lo oral), la carga informativa (presente en lo escrito, no en lo oral), la indicación de la idea central (explícita en lo escrito, implícita en lo oral), la expresividad (presente en lo oral, no en lo escrito), el entorno perceptivo común (presente en lo oral, no en lo escrito), el acceso a las propiedades del hablante y la situación (presente en lo oral, no en lo escrito), la interacción abierta (presente en lo oral, no en lo escrito), la interacción encubierta (presente en lo oral, no en lo escrito) y el impacto en el pensamiento (como consecuencia de la invención de la escritura). Dada la discrepancia entre planteamientos (dos, una, varias dimensiones de variación) resolvimos deshacernos de unas cuantas características.

Una primera criba consistió en descartar las que no tuviesen un impacto en la comprensión semántica y/o pragmática del lenguaje. ¿Qué descubrimos? Que casi todas son potencialmente relevantes. El canal sensorial determina el sub-procesador de la memoria de trabajo que es utilizado. Eso puede suponer un obstáculo cuando se trata de aprender conjuntamente de palabras e imágenes. La decodificación determina en algún grado la comprensión que se alcanza. El control significa que los lectores pueden releer, regular el ritmo de lectura y focalizar el segmento de interés, algo que, de hecho, hacen. Todo eso redundaría en una mejor comprensión y, si algo de eso falla, la comprensión sufre. La densidad léxica hace que haya más items léxicos a reconocer lo que obliga a tener un amplio vocabulario (que no todos tienen). Algunos elementos auxiliares tienen un impacto en el procesamiento de palabras. La integración obliga a invertir mayores esfuerzos y penaliza la comprensión. Lo mismo con la carga informativa. Señalar explícitamente la idea central contribuye a comprender mejor. La prosodia tiene múltiples funciones. Ayuda en la adquisición de una lengua, clarifica la interpretación de oraciones ambiguas, anuncia el fin de un enunciado, refleja las actitudes del hablante. No está claro que los gestos tengan un impacto. Compartir el acceso a un mismo entorno en el que ciertos eventos tienen lugar favorece la comunicación. Tener acceso a cierta información sobre la situación y el hablante beneficia a la interpretación del lenguaje. Mantener una interacción abierta o encubierta puede hacer más eficiente y eficaz la comunicación. El impacto en el pensamiento se produce, pero no es relevante para la comprensión. Balance: dos bajas.

En una tercera sección realizamos una criba adicional. Para empezar, nos preguntamos si las características eran realmente inherentes a cada modalidad o si, por el contrario, dependían de cuestiones coyunturales. El canal sensorial se vuelve importante sólo cuando hemos de aprender conjuntamente de palabras e imágenes. Sin esa circunstancia, no importa. La decodificación importa cuando uno está aprendiendo a leer. Una vez dominada la habilidad

para traducir palabras impresas, la decodificación importa poco. La densidad depende no de la modalidad sino del registro. La densidad aparece si el registro expositivo, no si no lo es. El acceso a cierta información sobre la situación comunicativa y la interacción están disponibles no en la modalidad oral sino en un setting particular. La conversación cara a cara es la que presenta dichas características, no otras formas de comunicación oral. El control y la expresividad aparecen en todas las variantes de la modalidad correspondiente. Balance: nos quedamos con control y expresividad.

Para reforzar el análisis nos preguntamos si el control y la expresividad son exclusivos de lo escrito y lo oral, respectivamente. Lo escrito incorpora signos de puntuación que capturan en cierto modo la prosodia oral. Los problemas son que exigen habilidad para decodificarlos, que sólo dan pistas y no instrucciones completas, que no dan cuenta de todos los fenómenos prosódicos posibles. Lo oral puede permitir la consulta repetida del input lingüístico. No obstante, nunca proporciona la posibilidad de una inspección minuciosa. De modo que control y expresividad son características exclusivas de lo escrito y lo oral, respectivamente.

Volvemos al principio. Asumimos entonces que son el control y la expresividad las características específicas de cada modalidad. Ni más ni menos. Sabemos, además, que ambas pueden resultar ventajosas en tanto que hay evidencias de que favorecen la comprensión del lenguaje.

4.1. Predicciones y consecuencias.

Gough y sus asociados (Gough et al., 1996; Hoover & Gough, 1990) planteaban, desde la simple view of reading, que sólo hay una diferencia entre las modalidades. Lo escrito exige decodificar y lo oral no. Por lo demás procesar el lenguaje escrito y hacer lo propio con el oral no presenta diferencias.

La aproximación funcional predice que las ayudas regulatorias orales son mejores que las escritas y que las ayudas explicativas escritas son mejores que las orales. Se asume que hay dos tipos de ayudas, unas que solicitan al aprendiz hacer algo y otras que le proporcionan una elaboración aclaratoria. Al recibir las primeras el aprendiz debe decidir si hacer o no lo que se le pide; al recibir las segundas debe procesar e integrar la información. La modalidad escrita proporciona control, como hemos visto en este capítulo. Eso hace posible releer, regular el ritmo de lectura y enfocar el segmento de interés. La modalidad oral, por su parte, proporciona expresividad, lo que aporta información sobre la actitud del hablante para con lo que dice -como sabemos. La interacción prevista por la aproximación funcional se explica como sigue. Las ayudas regulatorias orales informarán sobre la actitud del hablante lo que le hará al aprendiz entender lo importante de hacer lo que se le pide. Las ayudas explicativas escritas otorgarán al aprendiz el control que necesita para integrar la compleja elaboración que se le ofrece.

¿Qué diría la simple view of reading? A igualdad de habilidad para decodificar, no debería haber diferencia entre modalidades ni al regular ni al explicar. Recordemos que para ellos, más allá de la frontera de la decodificación no hay diferencias entre lo oral y lo escrito.

Por tanto, si logramos confirmar las dos predicciones de la aproximación funcional, la simple view of reading podrá ser revisada. ¿Cómo explicar que lo oral sea mejor que lo escrito unas veces y lo escrito mejor que lo oral otras? La simple view of reading diría que lo oral puede ser mejor que lo escrito si la habilidad para decodificar no se ha desarrollado plenamente. Pero no podría explicar un beneficio de lo oral cuando la decodificación está dominada o, aunque no lo esté, que aparezca unas veces sí (ante la regulación) y otras no (ante la explicación). La simple view of reading nunca diría que lo escrito puede ser mejor que lo oral. Si existiese un dominio absoluto de la decodificación Gough y los suyos sólo aspirarían a la equiparación entre comprensión oral y escrita. Pero nunca podrían encajar el beneficio de lo escrito sobre lo oral. O sea que confirmar nuestras predicciones permitiría revisar el planteamiento de Gough.

Capítulo 3.

La interacción función-modalidad en su contexto: diseñando ayudas para los computer-based learning environments.

0. Introducción.

Aprender de los Computer-Based Learning Environments (CBLEs, a partir de ahora) plantea serias dificultades a los aprendices. Entre ellas están el tener que (a) supervisar su aprendizaje para detectar problemas de comprensión y (b) generar algún tipo de reparación para tales problemas. Ni detectar ni reparar resulta fácil para los aprendices. Por eso los CBLEs incorporan habitualmente algún tipo de asistencia o *ayuda*. Dado que los CBLEs son muy flexibles, estas ayudas pueden diseñarse de formas muy distintas. Una de las cuestiones que los diseñadores deben responder es qué modalidad, la auditiva o la visual²², es la mejor para la presentación de ayudas verbales. Como descubriremos no hay respuesta para este interrogante. En consecuencia, nuestro objetivo será proponer un marco de investigación capaz de proporcionar pautas sobre cómo presentar las ayudas verbales.

Comenzaremos el capítulo definiendo qué son los CBLEs. Arrancaremos con una escueta definición de material instructivo. Continuaremos describiendo las posibilidades de los CBLEs como materiales instructivos. Diremos que los CBLEs son materiales que emplean un soporte digital para enseñar sobre algún tópico. Veremos que son materiales versátiles, en tanto que pueden incluir información verbal y/o pictórica, en modalidad auditiva y/o visual, todo ello presentado en forma lineal o hipertextual.

En una segunda sección lo que haremos será especificar qué procesos y representaciones entran en juego al aprender de los CBLE. Habrá de quedarnos claro que aprender a partir de un CBLE es crear un modelo mental. Después plantearemos un listado de dificultades u obstáculos que hacen que la construcción de dicho modelo mental no sea una tarea fácil. Comprobaremos que el volumen de procesamiento que debe soportar quien

22 En el capítulo anterior concluimos que el canal sensorial no era un factor relevante en la comprensión del lenguaje. Ahora, sin embargo, hablamos de información verbal/ayudas verbales visuales y auditivas. Lo hacemos porque en el campo de los CBLE es más apropiado usar estos términos en lugar de escrito y oral, ya que los primeros se pueden aplicar también a lo pictórico. Se gana así en parsimonia.

aprende a partir de un CBLE es considerable y que, paradójicamente, su capacidad de procesamiento es limitada por distintas razones. Ello explicará por qué es común que el aprendizaje que alcanzan en muchas ocasiones los aprendices es superficial, pobre. Se pondrá sobre el tapete la necesidad de llevar a cabo procesos de detección y reparación.

En la sección siguiente trataremos la cuestión de cómo ayudar a los aprendices a realizar tales procesos. Primero habremos de definir qué entendemos por detectar y reparar. Después, haremos acopio de evidencias que muestran que estos procesos son cruciales en el aprendizaje. Acto seguido recabaremos otras evidencias; éstas tendrán que ver con el desafortunado hecho de que ni la detección ni la reparación resultan fáciles para los aprendices. Finalmente, propondremos una relación de ayudas posibles a estos procesos y daremos cuenta de su probada efectividad.

En la sección cuarta nos adentraremos de lleno en el problema que motiva estas páginas: no sabemos aún qué modalidad emplear para presentar las mencionadas ayudas. Y si no lo sabemos es porque, tal como se argumentará en su momento, el principio de modalidad no es aplicable para las ayudas verbales y la poca investigación disponible sobre el impacto de la modalidad de las ayudas arroja resultados confusos. Lo que habrá que hacer entonces es presentar el principio de modalidad y razonar por qué no es susceptible de ser aplicado y revisar la investigación empírica sobre el tema en cuestión para comprobar que no revela un patrón claro.

En la última sección del capítulo plantaremos el marco de investigación que pretende ser una respuesta al problema: la aproximación funcional. La aproximación funcional proporciona pautas de diseño de ayudas y permite reinterpretar los resultados anteriores. Descansa sobre tres supuestos, dos de los cuales han sido objeto de discusión en los capítulos precedentes, de modo que aquí bastará con mencionarlos. El tercero plantea dos predicciones que son las que serán testadas en la parte empírica. Al terminar la sección ofreceremos una panorámica de los experimentos que siguen.

1. Qué son los computer-based learning environments.

1.1. Los materiales instructivos.

Entendemos por *material instructivo* todo aquel material que pretende enseñar algo a un aprendiz. Enseñar algo a un potencial aprendiz, a su vez, tiene como fin último el que éste pueda desempeñar su papel en la sociedad o, más específicamente, que pueda integrarse en una comunidad particular (e.g., la comunidad científica). Ejemplos de materiales instructivos son un libro de texto, un CD-Rom que incluye presentaciones multimedia o la entrada de una enciclopedia virtual.

1.2. Las posibilidades de los computer-based learning environments.

Los CBLEs son un tipo particular de material instructivo. Son los que utilizan un soporte digital, o dicho de otro modo, están basados en el ordenador. Los CBLEs brindan

múltiples posibilidades, lo que hace de ellos herramientas realmente flexibles. Enseguida definimos todas esas posibilidades.

Desde un punto de vista semiótico, los CBLEs pueden presentar información de dos tipos: *verbal* y *pictórica*. La información verbal está basada en palabras y oraciones. De este modo, una explicación, una taxonomía, un comentario de cualquier tipo (e.g., “lee esto con atención”), etc. son informaciones de tipo verbal. El otro tipo de información es pictórico y se refiere a información basada en imágenes y sonidos. Así, una ilustración, una fotografía, una animación, un vídeo, un mapa conceptual, una melodía o un sonido (e.g., el rugido de un león, un aplauso) son informaciones de tipo pictórico.

Desde un punto de vista sensorial, los CBLEs pueden hacer uso de dos modalidades, la *visual* y la *auditiva*, para presentar la información verbal y pictórica. Así, lo verbal puede proveerse de forma auditiva u oral (i.e., narración) y visual o escrita (i.e., texto); lo pictórico, por su parte, puede consistir en imágenes visuales o auditivas. Las primeras se corresponden con ilustraciones (e.g., diagrama, fotografía), animaciones (e.g., simulaciones, vídeos), etc.; las segundas, con melodías o sonidos.

Desde un punto de vista estructural, la información puede presentarse de forma *lineal* o *hipertextual*. Una estructura lineal es la que sigue un texto convencional: un párrafo sigue al anterior dentro de un apartado, los apartados se van sucediendo hasta completar secciones, las secciones se suceden para formar capítulo, etc. Una estructura hipertextual es la que hace posible saltar de un bloque o módulo a otro (e.g., de párrafo a párrafo, de animación a animación) de forma libre. Es decir, cada aprendiz diseña su propio recorrido a través de los bloques o módulos del material. Esta estructura es la que podemos encontrar en la world wide web.

Por último, desde un punto de vista instructivo, los CBLEs incluyen dos clases de componentes. Por un lado están los *contenidos*, que son los bloques que servirán al aprendiz para conformar una representación mental del tópico tratado en el material (e.g., geología, biología, ingeniería). Por otro lado están las *ayudas*, que son dispositivos que asisten al aprendiz en la ejecución de los procesos mentales necesarios para construir la representación mental ya mencionada. Un ejemplo (28) que permite entender la diferencia entre una clase de componente y otra es el que sigue.

- (37) “El Mediterráneo se muere porque sufre graves problemas. *En primer lugar*, su contaminación es cada vez mayor. *En segundo lugar*, el canal por el cual deben renovarse sus aguas es muy estrecho. *En resumen*, hay serios problemas que ponen en peligro la vida del Mediterráneo.”

Puede apreciarse que los elementos señalados en cursiva no dicen nada sobre el Mediterráneo. O sea, los elementos en cursiva no aportan contenidos sobre el tópico en cuestión (en este caso, la muerte del mar Mediterráneo). Sin embargo, no son elementos accesorios: desempeñan un papel muy particular. Los elementos en negrita indican al aprendiz qué operaciones mentales debe llegar a cabo, le indican cómo conjugar las piezas de información para configurar la representación mental deseada. De esa forma le asisten en la ejecución de

los procesos mentales oportunos.

De este análisis somero que acabamos de hacer cabe deducir que los CBLEs son materiales versátiles. Versátiles porque todas esas posibilidades descritas pueden explotarse para dar lugar a materiales diversos (ver Figura 1). Un CBLE puede ser una animación sobre la formación de tormentas acompañada de una narración que describe los eventos representados en tal animación (1). O puede ser la misma animación acompañada de texto (2). Puede ser también una ilustración (en forma de diagrama) acompañada de un texto que explica los pasos a seguir para verificar la seguridad de un circuito eléctrico (3). Puede constar de ilustraciones sobre el vuelo de los aviones en las que se han insertado ayudas pictóricas (4) en forma de flechas que dirigen la atención sobre un aspecto concreto de la ilustración. Así mismo, todos estos materiales podrían ser bloques ordenados lineal o hipertextualmente.

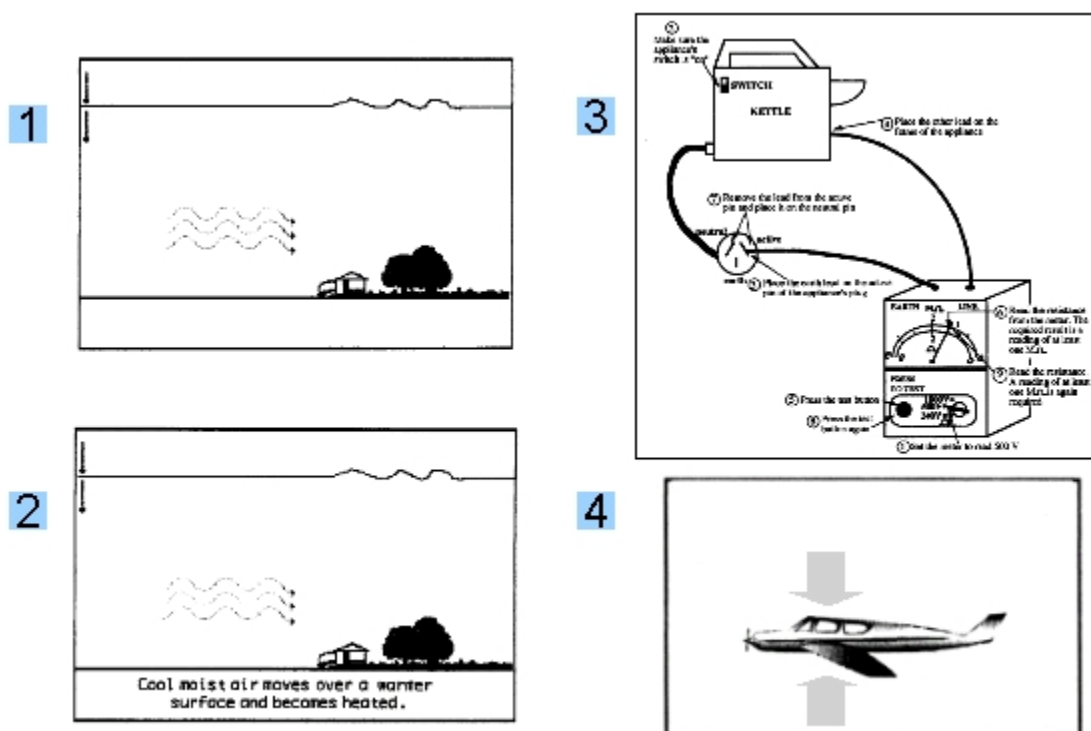


Figura 1. Ejemplos de CBLE. (De izquierda a derecha, de arriba a abajo) Tomados de Mayer y Moreno (1998); Tindall-Ford, Chandler y Sweller (1997); Mautone y Mayer (2001).

Las posibilidades, decíamos, son múltiples. La tabla 14 recoge todas las fuentes de variación (excepto la estructural).

Tabla 14. Las posibilidades de los CBLEs.

	Verbal		Pictórico	
	Auditivo	Visual	Auditivo	Visual
Contenido	Narración	Texto	Música, sonidos	Animaciones, ilustraciones
Ayuda	Ayuda verbal auditiva	Ayuda verbal visual	Ayuda pictórica auditiva (e.g., aplauso)	Ayuda pictórica visual (e.g., flechas, elementos coloreados)

2. Aprender a partir de los CBLE es difícil.

En esta sección vamos a comprobar que si está justificado introducir ayudas en los CBLE es porque al aprender mediante ellos los aprendices han de sortear varios obstáculos. Primero vamos a ver qué es aprender de ellos. A continuación repasamos cuáles son esos obstáculos y exponemos evidencias que demuestran que resultan ser tal cosa para los aprendices. Finalmente veremos que, tomados conjuntamente, los obstáculos aumentan las probabilidades de que los aprendices construyan representaciones mentales pobres a partir de los CBLEs.

2.1. Aprendizaje a partir de CBLEs: construir modelos mentales.

Aprender a partir de un CBLE supone configurar una representación mental coherente del tópico tratado en él que queda fijada en nuestra memoria y que puede ser invocada posteriormente para resolver tareas de distinta índole. Para construir dicha representación es necesario poner en marcha una serie de procesos mentales que tratan de interpretar lo que el material presenta y que se apoyan en ciertos conocimientos previos. En lo que sigue vamos a describir esas representaciones, procesos y conocimientos con detalle. En primer lugar, lo haremos para el caso del aprendizaje a partir de narraciones y textos, es decir, de información exclusivamente verbal. Seguidamente, lo haremos para el caso del aprendizaje a partir de información verbal y pictórica (i.e., aprendizaje multimedia).

2.1.1. Aprender a partir de textos. Tal como se ha dicho hace un instante, este primer subapartado estará dedicado a las representaciones, procesos y conocimientos que entran en juego cuando se aprende a partir de un texto. Más adelante veremos qué ocurre cuando las imágenes entran a escena.

2.1.1.1. Las representaciones: modelos mentales. Cuando leemos un texto construimos una recreación o escenario mental que da cuenta del estado de cosas descrito por aquél. Tal representación mental está compuesta de una serie de unidades básicas, unidades cuya naturaleza no está clara aún. Además, la representación puede diferir en el grado en que se despega de lo presentado originalmente en el texto, distinguiéndose al menos dos niveles. Enseguida diferenciamos esos niveles de representación para después discutir sobre la

naturaleza de las unidades básicas.

Venimos diciendo que comprender un texto es construir un escenario o representación mental que refleja el mundo referido en el texto. Así, si el texto nos habla de un viaje en barco nosotros “visualizamos” de algún modo una embarcación, el mar, los tripulantes y otros elementos pertinentes para reconstruir mentalmente la situación que se nos relata; si el texto habla sobre el cocinado de una tortilla de patata, hacemos tres cuartos de lo mismo: imaginamos una cocina, con sus cazuelas y fuegos y a un actor ejerciendo de cocinero. Una cuestión intrigante es que, dado que un texto no aporta nunca suficiente información como para que el lector pueda recrear el estado de cosas descrito en él, ¿cómo consigue ese escenario mental ser finalmente una simulación del mundo descrito en el texto? Pues, como comprobaremos ahora, porque toma la información que el texto no proporciona del fondo de conocimiento almacenado en la memoria a largo plazo.

- (38) “Llovía a cántaros. Decidí volver a casa a por mi paraguas. No sirvió de mucho: terminé empapado”

El lector habrá podido inferir que si el protagonista de este pequeño relato (38) vuelve a casa a por su paraguas es porque éste le puede proteger de la lluvia. También habrá interpretado que el protagonista acabó empapado por culpa de la intensa lluvia y la ineficacia de su paraguas²³. Quizá el lector también se ha imaginado el cielo repleto de nubes o el disgusto del protagonista; tal vez ha visualizado a éste dirigiéndose de nuevo a casa, etc. Así, la representación mental que el lector ha configurado integra información procedente del relato original con otra aportada por él mismo (i.e., las inferencias). El resultado es una representación despegada del texto: en nuestra mente no pervive la descripción proporcionada por el texto sino una simulación del pequeño mundo referido por aquél.

Pero el tipo de comprensión descrito no siempre tiene lugar. ¿Qué sucede si el lector no genera inferencias? En esos casos la representación mental construida está más apegada al texto.

- (39) “El magma de la parte inferior del manto se encuentra cerca del núcleo, el cual presenta unas temperaturas muy elevadas. El núcleo calienta el magma cercano a él y, como consecuencia, este magma asciende a la superficie. El magma de la superficie se encuentra lejos del núcleo caliente y, por tanto, se enfría y desciende.”

En este caso (39), muchos lectores que no saben a qué fenómeno se está refiriendo el texto (las corrientes de subducción en el manto de la Tierra) o que no han puesto mucho empeño al comprender no habrán podido más que retener algunas ideas. Ideas que, además, estarán deslavazadas en su mente. Estos lectores quizá hayan entendido que hay una entidad llamada “magma” que asciende cuando es calentada por el “núcleo” y que puede también encontrarse lejos de éste, lo que provoca que descienda. Por supuesto, estos lectores no tienen por qué haber entendido que el hecho de que el magma que ascienda y descienda da lugar a una corriente continua. Y que si asciende y desciende es porque el calor reduce la densidad de los fluidos y,

23 Todo esto puede parecer evidente pero si nos ceñimos al texto apreciamos que, en realidad, no se dice explícitamente. Es nuestra concepción del mundo la que nos ayuda a completar las “fisuras” del texto.

por tanto, los hace ascender cuando se encuentran inmersos en un fluido más denso; el frío, sin embargo, aumenta su densidad y los hace descender. Para eso, claro está, hay que saber que el manto está compuesto de magma, una sustancia semilíquida formada por rocas fundidas. La representación de esos hipotéticos lectores profanos en geología o lectores perezosos recogería parte de lo que dice el texto pero ni de lejos conseguiría simular mentalmente el fenómeno descrito (como sucedía cuando uno configuraba un escenario mental, como en el ejemplo del día lluvioso (38)). Si es así es porque no han podido o no han querido aportar, gracias al fondo de conocimiento correspondiente, la información que el texto omite. La representación de estos lectores hipotéticos estará consecuentemente más apegada al texto: en su mente perviven algunas ideas presentadas por el texto pero no el mundo al que aquél hizo referencia.

A tenor de las situaciones que hemos ilustrado, cabe plantear que existen distintos niveles representacionales. Éste es precisamente uno de los supuestos básicos para muchos estudiosos de la comprensión del texto (e.g., Graesser et al., 1997; Kintsch, 1998; van Dijk & Kintsch, 1983). Las representaciones que incluyen cierto volumen de información inferida son *modelos mentales* o modelos de situación (dependiendo de si dan cuenta de fenómenos alejados de la experiencia cotidiana o no, respectivamente). Las que no contienen tantas aportaciones por parte del lector y, por tanto, están más apegadas al texto original, se conocen con el nombre de *base del texto*. Más técnicamente cabe decir que el modelo mental es una representación que integra la información proporcionada por el texto con los conocimientos del lector; tal representación refleja el mundo descrito en el texto. La base del texto es una representación mental que recoge las ideas presentes en el texto y algunas conexiones entre ellas; estas conexiones sirven para ensamblar ideas adyacentes o para sintetizar varias ideas en una de orden superior.

Tal como dijimos en el capítulo 1, los modelos mentales son multidimensionales (Zwaan & Radvansky, 1998). Lo son porque dan cuenta de las distintas circunstancias de la situación referida en el texto. Recogen así el lugar y el momento en los que tiene lugar la acción, las causas de los eventos que ocurren, las intenciones de los personajes que emprenden acciones y las propiedades de personajes y eventos.

¿Qué evidencias tenemos de que existen estos niveles? En primer lugar, hay estudios que indican que la base del texto se desvanece de nuestra memoria más rápidamente que el modelo mental. En varios experimentos se ha examinado el modo particular como el modelo mental y la base del texto van decayendo de la memoria con el tiempo (e.g., Kintsch, Welsch, Schmalhofer & Zimny, 1990; Schmalhofer & Glavanov, 1986). Por ejemplo, en el artículo de Kintsch et al. (1990) se describe un estudio de esta clase. Después de leer un texto los participantes debían decidir si oraciones de distintos tipos figuraban o no en el texto original: oraciones literales extraídas del texto, paráfrasis u oraciones que recogían ideas no presentes en el texto pero inferibles de él. Algunos sujetos resolvieron la tarea inmediatamente tras la lectura del texto, otros con una demora de 40 minutos, otros con demora de 2 días y, por último, otros tantos lo hicieron con una demora de 4 días. Lo que se observa es que la probabilidad de responder “sí” (i.e., considerar que la oración pertenecía al texto original) tiene una evolución en el tiempo distinta en función del tipo de oración que consideremos. La clave es que las paráfrasis están asociadas a un nivel de base del texto (i.e., el lector es capaz

de decir lo mismo que el texto pero con sus propias palabras) mientras que las inferencias lo están a un nivel de modelo mental (i.e., el lector no retiene lo que decía el texto sino el mundo que aquél describía). Tras aplicar ciertos cálculos sobre las curvas de probabilidad, los investigadores estimaron la intensidad con la que cada nivel representacional permanecía vivo en la memoria de los participantes. El modelo mental no experimentaba cambios después de 4 días de demora: los sujetos podían recuperar esta representación sin problema. La base del texto, no obstante, decayó a un ritmo notablemente más rápido.

En segundo lugar, existen estudios que ponen de manifiesto que muchos lectores son capaces de construir la base del texto pero no el modelo mental (e.g., McNamara, Kintsch, Songer & Kintsch, 1996; McNamara & Kintsch, 1996; Perrig & Kintsch, 1985; Vidal-Abarca, Martínez, & Gilabert, 2000). Por ejemplo, en el estudio de Perrig y Kintsch (1985) los sujetos debían leer un texto que describía la ubicación de los lugares de interés de una ciudad imaginaria. Tras leerlo tuvieron que resolver tareas diferentes. Recordaron todo lo posible acerca del texto, juzgaron como verdaderas o falsas algunas oraciones y dibujaron un mapa de la ciudad imaginaria. Los participantes recordaron razonablemente bien lo que el texto decía y pudieron identificar con éxito las oraciones verdaderas/falsas, siempre y cuando éstas hiciesen referencia a alguna idea planteada explícitamente por el texto. El rendimiento fue significativamente peor en las otras tareas: no pudieron dibujar el mapa de la ciudad ni pudieron distinguir las oraciones verdaderas de las falsas cuando éstas hacían referencia a información implícita en el texto.

Es de interés que, de acuerdo con el tipo de representación mental elaborada, los lectores son capaces de resolver unas tareas u otras. Así, quienes han elaborado una base del texto son capaces de resumir lo que han leído (e.g., Kintsch & Kozminsky, 1977), responder a preguntas sobre información presente en el texto (e.g., Kintsch & Yarbrough, 1982) y pueden recuperar la información si se les pide que recuerden todo lo que puedan sobre el texto (e.g., Kintsch, Kozminsky, Streby, McKoon & Keenan, 1975). Quienes han elaborado un modelo mental son capaces de realizar todas las tareas ya mencionadas y alguna más. En concreto, pueden responder a preguntas de inferencia y de solución de problemas (e.g., Gilabert, Martínez & Vidal-Abarca, 2005; Mayer et al., 1984). Esto, obviamente, es de interés en un contexto educativo. Si algún tipo de representación permite al lector resolver tareas complejas inasequibles para el otro nivel (como es el caso), entonces un objetivo educativo es lograr que todo lector alcance tal nivel de representación.

De lo que acabamos de decir se desprende que si un investigador está interesado en evaluar un nivel u otro deberá plantear tareas específicas. Si pretende saber si cierto participante ha sido capaz de construir una base del texto bastará con pedirle, por ejemplo, que recuerde todo lo que pueda de lo que leyó. Si, en cambio, pretende determinar si el participante ha elaborado un modelo mental será necesario plantear otras tareas, como por ejemplo preguntas de inferencia.

Tabla 14. Las representaciones mentales que resultan al aprender a partir de textos.

<i>Nivel representacional</i>	<i>Descripción</i>	<i>Posibilidades</i>	<i>Evidencias</i>
Base del texto	Representación apegada al texto, que recoge las ideas presentes en el mismo	Recordar la información presente en el texto, resumirla, responder a preguntas sobre ella	Primero, el recuerdo de la base del texto decae antes que el modelo mental (e.g., Kintsch et al., 1990)
Modelo mental	Representación despegada del texto, que recrea el estado de cosas descrito en el texto incluyendo diferentes dimensiones	Resolver tareas que implican <i>utilizar</i> lo aprendido	Segundo, puede construirse la base sin haberse construido el modelo (e.g., Perrig & Kintsch, 1985)

Aprender a partir de un texto, hemos afirmado, supone conformar una recreación mental de los contenidos presentados por el texto. ¿De qué está hecha dicha recreación? ¿Cuáles son las unidades básicas del significado? Muchos investigadores asumen que el significado es representado en la memoria en forma de estructuras formadas por símbolos mentales o *proposiciones* (e.g., Ratcliff & McKoon, 1978; van Dijk & Kintsch, 1983). Las proposiciones capturan las propiedades esenciales de una afirmación: argumento (tema) y predicado (comentario). Pongamos que se nos presenta esta oración: “Luis suele ir al gimnasio”; otro modo de presentarla sería: “Luis tiene costumbre de dirigirse al gimnasio”; o incluso: “Habitualmente Luis acude al gimnasio”. Pues bien, en todos los casos la representación mental que generamos es:

- I. (LUIS, IR, GIMNASIO)
- II. (TIEMPO: (I), A MENUDO)

El tema o argumento es “Luis” y el comentario o predicado es “ir, gimnasio” con el complemento “a menudo”. La proposición pues abstrae la esencia del enunciado, al margen de su formato superficial. Nótese que algo semejante a esa estructura sería lo que quedaría fijado en nuestra memoria, de acuerdo con esta postura.

No todos están de acuerdo con esto. Otros estudiosos proponen que la unidad básica del significado son pequeños fragmentos de experiencia o representaciones sensorio-motoras (e.g., de Vega, 2002, 2005; Glenberg, 1997; Zwaan, 2004). Así, al comprender una oración como la anterior no elaboramos símbolos mentales que codifican su significado (algo que equivaldría a traducirla del español al inglés: *Luis is used to going to gym*; arguyen estos autores), sino que revivimos de algún modo la experiencia que el texto describe. En este caso y, suponiendo que Luis sea mi hermano, se activan percepciones (imágenes, sonidos, representaciones hápticas, etc.) y *affordances* (o posibilidades de acción sobre el objeto, cosas que se pueden hacer con él) que tengo asociadas a Luis: la imagen de su cara, el timbre de su voz, la consistencia de su cuerpo, etc. Lo mismo con gimnasio: imágenes, sonidos y otras percepciones y *affordances* vinculados a este concepto son activados. Leer o escuchar una historia equivale entonces a experimentarla, aunque de una forma atenuada: al leer el micro-texto anterior yo podría imaginar que adopto la perspectiva del protagonista (Zwaan, 2004), es

decir, me voy de camino a un gimnasio montado en un coche (y se activan en algún grado representaciones sensorio-motoras asociadas a conducir) y, una vez allí, realizo varios ejercicios (y se activan las representaciones sensorio-motoras pertinentes), etc.

Los defensores de cada postura se hallan hoy día implicados en una estimulante polémica (más bien los segundos siguen reuniendo evidencias a favor del *embodiment* o corporeidad del significado; evidencias que no son interpretables desde una postura simbólica). No vamos a detenernos en ella. Baste decir dos cosas. Primero, lo cierto es que hay un buen número de trabajos que aportan datos sólo explicables desde una perspectiva corpórea. Por ejemplo, Glenberg, Meyer y Lindem (1987) encontraron que el concepto “sudadera” estaba más activado tras la lectura de “John estaba preparando una maratón, *se puso* su sudadera y se fue a correr” que tras la lectura de “John estaba preparando una maratón, *se quitó* su sudadera y se fue a correr”. Ello indica que, en efecto, uno revivie la experiencia del protagonista en lugar de codificar el lenguaje en símbolos (los cuales asociarían “sudadera” a “John” de la misma manera en las dos condiciones). Hay también datos de la investigación neurológica que (Tettamanti et al., 2005) indican que ciertas áreas motoras cerebrales se activan al leer oraciones como “agarro la pelota” pero no ante otras como “valoro la amistad”. Sabemos también (lo vimos en el capítulo 2) que incluso los conceptos abstractos parecen tener cierta base corpórea, en tanto que producimos más gestos cuando tenemos que explicar la relación entre unos objetos que cuando tenemos que describirlos (Alibali et al., 2000). Ello indica que el conocimiento tiene un formato multimodal.

Segundo, cada postura está (más) ligada a un campo de investigación específico. Quienes defienden la corporeidad del significado están (más) interesados en el procesamiento del lenguaje. Quienes asumen una postura simbólica están (más) interesados en el aprendizaje a partir de textos. Como de lo que se trata en este capítulo es de averiguar en qué consiste aprender de un texto, la mayoría de los trabajos consultados serán simbólicos.

En suma, tras la lectura de un texto cabe recrear el estado de cosas descrito en él o retener las ideas y relaciones presentes en él. Estos niveles se conocen con el nombre de modelo mental y base del texto, respectivamente. Hay evidencias que indican que son niveles diferentes. Alcanzar un nivel u otro permite resolver determinadas tareas en cada caso. El investigador de turno deberá emplear unas u otras dependiendo de su foco de interés. Un asunto abierto al debate es cuáles son las unidades básicas del significado, ¿de qué está hecho el modelo mental?

2.1.1.2. Los procesos: decodificación, sintaxis e inferencias. Para construir las representaciones que ya hemos descrito los lectores deben poner en marcha procesos. Estos procesos sirven para entresacar las ideas del texto, establecer conexiones entre ellas e ir agregando otras ideas procedentes de la memoria a largo plazo.

Lo primero es descifrar las palabras o *acceder al léxico*. Cuando las palabras se presentan por escrito, el lector debe ser capaz de llegar a su significado a partir de un patrón gráfico. El modelo de doble ruta de Coltheart (e.g., Coltheart, 2005) es tal vez el más

invocado a la hora de explicar el citado proceso²⁴. Según el modelo de doble ruta, para entender una palabra (escrita) seguimos dos caminos alternativos: (a) le asignamos el correspondiente sonido (fonema) a cada una de las letras o combinación de letras (grafemas) re-ensamblando el sonido completo de la palabra y activando así una representación fonológica que nos lleva a su significado asociado o (b) reconocemos una secuencia ortográfica familiar que activa directamente su fonología y significado asociado. La primera (a) se conoce con el nombre de ruta fonológica y obliga a reconstruir la fonología de la palabra para llegar a su significado y la segunda (b) con el de ruta léxica, la cual no exige más que acceder al significado (y el resto de representaciones asociadas a la palabra) a partir de un patrón ortográfico conocido. Según Coltheart, la fonológica se pone en funcionamiento ante palabras nuevas y la léxica ante palabras conocidas. Una buena manera de ganar apoyos para este modelo es demostrar que, tal como se predice del mismo, la longitud de las palabras sea una variable crítica en el procesamiento de palabras nuevas y no tanto en el procesamiento de palabras conocidas. Las primeras han de re-ensamblarse fonológicamente, trabajo tanto más arduo cuanto más larga es la palabra; las segundas se reconocen y descifran “de golpe”, luego la longitud no debería ser relevante en esos casos. Así es como sucede, ya que hay una interacción longitud x frecuencia de aparición de la palabra (Haberlandt & Graesser, 1985).

Sabemos, a partir del capítulo 2, que la decodificación tiene cierto peso en la comprensión de un texto. De acuerdo con el meta-análisis de Hoover y Gough (1990), en los primeros años de escolarización la decodificación explica gran parte de la varianza en comprensión lectora. Conforme ascendemos los peldaños de la escalera educativa, no obstante, el peso se va reduciendo. Supuestamente, la decodificación llega a automatizarse.

Una vez entendidas las palabras mediante esos mecanismos, un dispositivo mental establece relaciones entre los elementos de la oración. Concretamente, se identifica el rol que desempeña cada palabra en la oración, se trazan las relaciones entre ellos y se conforma una estructura. Es decir, hay un proceso de *análisis sintáctico* (que, en principio, es instantáneo; no como el que hacíamos en la escuela...). ¿Por qué es preciso construir tal estructura y no basta con considerar al mismo tiempo los significados aislados de cada palabra? Una razón es que no son suficientes los significados aislados para entender esto:

- (a) “Rubén golpeó a David”
- (b) “Rubén fue golpeado por David”

¿Cómo saber quién pegó a quién en cada caso sin la sintaxis? Los significados aislados son los mismos en (a) y (b) e incluso su orden de presentación: “Rubén”, “golpear” y “David”. Sólo si asignamos roles y ordenamos los elementos en una estructura puede entenderse quién golpeó a quién. Un indicador más contundente de que debemos realizar una suerte de análisis sintáctico antes de jugar con los significados aislados de cada palabra es éste: ciertas lesiones cerebrales (i.e., afasia de Broca) impiden entender oraciones como “el gato es perseguido por el ratón” puesto que los conceptos aislados se combinan de acuerdo con lo que uno sabe sobre gatos y ratones y no con arreglo al esquema sintáctico presentado (el afásico entenderá que “el gato persigue al ratón” y no al revés). Así lo pusieron en evidencia Caramazza y Zurif (1976;

24 Si bien, no es en absoluto el único. De hecho, hay otros modelos que han recibido considerable apoyo empírico también (e.g., Plaut, 2005).

citados en deVega, Carreiras, Gutiérrez-Calvo, & Alonso-Quecuty, 1990) en su investigación.

Una vez conocido el significado de cada palabra y ordenados todos ellos de acuerdo con su sintaxis, una o varias proposiciones son construidas. A medida que se van construyendo más y más proposiciones se hace necesario ir conectándolas de alguna forma. Hay por lo menos dos formas de conectar unas proposiciones y otras: una es local, porque pone en relación ideas adyacentes; otra es global, dado que pone en relación ideas dispares. Veamos en qué consisten.

Para conformar una representación coherente a nivel local es preciso generar *inferencias puente* (cf. León & Pérez, 2003). Hay dos procesos que actúan bajo esta denominación: la resolución de anáforas y las conexiones (inferencias) entre frases consecutivas. Una anáfora es todo elemento del texto que remite a otro anterior. Hay varios elementos que pueden actuar como anáfora: el pronombre (e.g., “Andrea es mi amiga. *Ella* es muy vehemente”), también un hiperónimo (e.g., “La guitarra eléctrica tiene seis cuerdas. Gibson inventó *este instrumento* en el siglo XX”) o un sintagma nominal (e.g., “José tuvo un accidente con su coche. Daba pena ver su *Seat Panda siniestrado*”). Para resolver la anáfora el lector ha de reactivar un argumento anterior (Andrea, guitarra eléctrica, coche), asociarlo con la anáfora presente (élla, este instrumento, Seat Panda siniestrado) y aplicar el nuevo predicado al argumento rescatado (Andrea es muy vehemente, etc.).

El otro tipo de operaciones locales son las conexiones (inferencias) causales, temporales, espaciales, motivacionales y lógicas. Un ejemplo de las primeras es el vínculo que establecemos al leer este fragmento: “El coche chocó contra la pared. El coche quedó destrozado”. Aquí el lector asume que el coche quedó tan maltrecho “a consecuencia de” la colisión contra la pared. Sin esta operación las dos frases quedan deslavazadas²⁵. Un ejemplo de conexión espacial es el que adjuntamos a continuación: “Marina atravesó el porche de la casa. En la calle se encontró con Paloma”. Sabemos que Marina llegó a la calle cruzando el porche de la casa “que comunica ambas estancias”. Sin realizar esta conexión no habría trabazón entre las dos frases.

Está documentado que los lectores generan inferencias puente. McKoon y Ratcliff (1992), quienes llevaron a cabo varios experimentos sobre inferencias puente, nos ofrecen evidencias en este sentido. Por ejemplo, en su primer experimento la palabra “rifle” estaba más activada en la mente de quienes habían leído “el asesino quería matar al presidente con un potente rifle, colocó el rifle sobre su hombro, la mira se cayó, buscó la manera de disparar sin ella, el sol golpeó sus ojos” que en la de quienes leyeron “el asesino quería matar al presidente con un potente rifle, colocó el rifle sobre su hombro, la mira se cayó, buscó sus granadas para lanzarlas, el sol golpeó sus ojos”. El resultado indica que los lectores conectan oración a oración el texto vía inferencias. En la primera versión el asesino dispara sin mira *pero con rifle*, por lo que es preciso mantener activo tal concepto; en la segunda, el asesino abandona el rifle para lanzar sus granadas: no es preciso mantener activo tal concepto para entender las oraciones finales.

25 Obviamente, el lector puede no realizar la conexión, formando en su mente una representación aislada de las dos proposiciones.

Conectando ideas dispares cabe inferir ideas de orden superior que aglutinan otras: las macroproposiciones. Así, las inferencias a nivel global buscan reducir la información a un número manejable de ideas principales o macroproposiciones. Para formar macroproposiciones es necesario aplicar distintos procedimientos o *macro-reglas* (Kintsch & van Dijk, 1978): borrado, generalización y construcción. Borrar consiste en eliminar las proposiciones que no son condición necesaria para interpretar otra proposición. Suprimiendo éstas, obtenemos las ideas fundamentales. La generalización, por su parte, consiste en sustituir mediante una proposición general un set de proposiciones denotadas por élla. Por último, cabe construir una proposición que recoge otras que sirven como condiciones, componentes o consecuencias de la primera. En todos los casos lo que conseguimos es operar con un número manejable de ideas que dan unidad al texto. Valga el siguiente ejemplo para ilustrar la macro-regla de construcción: “Ricardo entró en el establecimiento. Encargó una sopa al camarero. Después de comerla solicitó la cuenta.”, la macroproposición es “Ricardo comió en un restaurante”.

¿Con qué evidencias de que estos procesos tienen lugar contamos? Graesser, Singer y Trabasso (1994) revisan unos cuantos experimentos en su conocido artículo. Se trataba de probar que, además de las inferencias puente, los lectores producen inferencias para conectar elementos dispersos (como una macro-regla de construcción). Uno de los experimentos que describen puso de manifiesto que quienes leían que “Valerie salió pronto de casa para ir al cumpleaños” lo conectaban con “pasó una hora comprando en el centro comercial” que apareció cinco oraciones más tarde. La prueba es que éstos tenían más activado el concepto cumpleaños que quienes leyeron inicialmente “Valerie salió pronto de la casa del cumpleaños”.

Una vez desentrañadas las ideas del texto y establecidas las relaciones locales y globales pertinentes cabe enriquecer la representación resultante agregando conocimientos desde nuestra memoria a largo plazo. Estas aportaciones se conocen con el nombre de *inferencias elaborativas* (cf. León & Pérez, 2003). Las elaborativas son inferencias innecesarias para mantener la coherencia local o global, a diferencia de las puente. Son útiles, en cambio, para ampliar el modelo de la situación referida en el texto. Pongamos que hemos leído el texto sobre Ricardo en el restaurante. Inferencias como “al salir del restaurante Ricardo tenía menos dinero en su cartera”, “la sopa estaba caliente” o “usó un plato y una cuchara para comérsela”, son elementos que amplían la representación que nos hemos construido sobre la comida en el restaurante. Amplían, enriquecen, otorgan mayor credibilidad a la narración pero no son estrictamente necesarias para comprender el texto (como lo son las inferencias puente).

O'Brien y Albrecht (1992) llevaron a cabo varios experimentos que permiten sostener que los lectores generan inferencias elaborativas durante la lectura de un texto. O'Brien y Albrecht presentaron historias como ésta (40):

- i. Carmen estaba dentro/fuera del club deportivo sintiéndose un poco perezosa.
- ii. Claro que una sesión de ejercicio físico siempre le hacía sentir con más energía. [...]
- iii. Decidió salir para estirar las piernas.

Los resultados mostraron que cuando (i) contenía la palabra “fuera” los sujetos experimentaban una incongruencia al leer (iii), según reflejaban las diferencias en sus tiempos de lectura de la oración (iii) entre las condiciones “dentro” y “fuera”. Esto significa que los lectores recrean mentalmente la escena descrita incluyendo aspectos que van más allá de la trama principal. En este caso, la dimensión espacial.

En resumidas cuentas, para construir las representaciones a las que nos hemos referido antes los lectores deben ejecutar ciertos procesos. Concretamente, deben acceder al léxico de las palabras impresas, deben conectar las oraciones consecutivas, deben establecer vínculos entre oraciones dispares para crear ideas que sintetizan otras y deben generar inferencias elaborativas para enriquecer el modelo mental construido. Hay pruebas de que tales procesos tienen lugar.

2.1.1.3. Los conocimientos. Sin ciertos conocimientos no es posible llevar a término ninguno de los procesos a los que nos hemos referido y, por extensión, es imposible configurar las representaciones descritas más arriba. Más tarde dedicaremos un apartado a ello, baste ahora un ejemplo sencillo. Si no sé que los paraguas protegen de la lluvia no podré realizar la inferencia puente que conecta estas dos oraciones: “llovía a cántaros; decidí volver a casa a por el paraguas”. Y si no las conecto no podré construir un pequeño modelo mental a partir de ellas.

Hay distintos tipos de conocimiento. Uno es conocimiento general sobre el mundo. En el ejemplo, saber que “los paraguas pueden proteger de la lluvia” forma parte de nuestros conocimientos generales, que hemos aprendido a partir de nuestra experiencia con el mundo. Para conectar estas otras dos oraciones: “la bujía se gastó; el coche se detuvo”; uno debe conocer específicamente el motor de cuatro tiempos. Debe, por tanto, poseer conocimientos específicos sobre una materia determinada. Los conocimientos del primer tipo son relativamente comunes pero los del segundo no lo son para nada. Como entenderemos más abajo, esta circunstancia supone un importante escollo.

2.1.2. Aprender a partir de palabras e imágenes: aprendizaje multimedia. Aprender conjuntamente de palabras e imágenes, el aprendizaje multimedia, significa construir un modelo mental sobre un tópico en cuestión (Mayer, 2001; Schnotz, 2002, 2005). Es decir, el resultado de la comprensión conjunta de palabras e imágenes es el mismo que se produce como consecuencia de la comprensión de un texto.

2.1.2.1. Ventajas de la inclusión de imágenes. Como hemos dicho hace un momento, para construir un modelo mental son necesarios ciertos conocimientos, que son los que permiten realizar las inferencias pertinentes. No disponer de esos conocimientos, por fuerza, impide que un modelo mental sea confeccionado.

La ventaja de la incorporación de imágenes a la información verbal es que ahorran la necesidad de (a) disponer de conocimientos y de (b) aplicarlos (realizar inferencias). Si el texto de hace un instante sobre el motor de cuatro tiempos incluyese imágenes, la necesidad de tener ciertos conocimientos y de ponerlos en juego vía inferencias quedaría satisfecha.

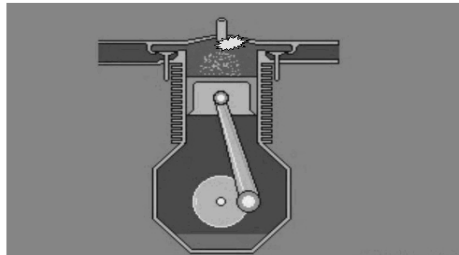


Figura 2. Imagen que se inserta en material con información verbal.

Fijémonos en la figura 2. Pongamos que cierto material ha hecho ya las presentaciones pertinentes. Nos ha dicho que el palito de la parte superior es una bujía, que la gran cámara es el cilindro y que la rueda inferior está conectada con las ruedas del coche de manera que si la primera gira, la segunda también. Tras eso vemos una pequeña secuencia en la que pequeños chispazos de la bujía provocan una explosión en la cámara que termina con la rueda girando sin parar. Entonces leemos que “la bujía se gastó; el coche se detuvo”. En ese caso las oraciones se entenderían perfectamente: el conocimiento suplido por las imágenes nos permite conectar las dos oraciones para configurar un pequeño modelo mental.

Mayer y Gallini (1990) explican esto misma de la siguiente manera. Sostienen que las imágenes ayudan a construir modelos mentales en virtud de sus dos características centrales. Primera, la *system topology* (o topología del sistema), es decir, las imágenes retratan los componentes principales en la estructura del sistema. Segunda, el *component behavior* (o comportamiento de los componentes), según la cual los estados de los componentes y los cambios en los estados de los componentes que afectan a otros componentes son retratados.

De acuerdo con lo planteado hasta aquí, los materiales que incorporen tanto imágenes como palabras deberían propiciar un aprendizaje mayor, con respecto a los que sólo incluyen palabras. Lo cierto es que esta predicción se ha confirmado en numerosas ocasiones (e.g., Glenberg & Langston, 1992; Mayer & Anderson, 1992; Mayer & Gallini, 1990). Por ejemplo, Mayer y Anderson (1992) solicitaron a un grupo de participantes que aprendiese cómo funciona una bomba de hinchar (Exp. 1) y el freno de tambor de un coche (Exp. 2) bien a partir de una animación con narración simultánea o bien a partir de la narración a secas. Tras consultar el material los participantes resolvieron tareas de transferencia (e.g., “Imagina que pulsas arriba y abajo el mango de la bomba de hinchar y no sale aire, ¿qué podría haber ocurrido?”). En dos experimentos, los de la primera condición rindieron significativamente mejor que los de la segunda en estas tareas. El efecto ha sido replicado en tantas ocasiones que se conoce, al menos en los términos de Richard Mayer, como *principio multimedia* (Mayer, 2001).

Que los materiales con imágenes y palabras conduzcan a un mayor aprendizaje que los que sólo incluyen palabras prueba que, en efecto, las imágenes contribuyen a la construcción de un modelo mental. No obstante, no nos dan pistas sobre el mecanismo por el cual esto

sucede. Habíamos asumido que si las imágenes ayudan es porque ahorran la necesidad de disponer de conocimientos y de aplicarlos. Consecuentemente, sólo quienes dispongan de pocos conocimientos son los que deberían beneficiarse de la presencia de imágenes. Y eso es exactamente lo que ocurre (Mayer, 2001, Capítulo 10).

En definitiva, la inclusión de imágenes a un material con información verbal tiene ventajas. En concreto, hace innecesario disponer de los conocimientos requeridos para construir un modelo mental. Hay evidencias de que imagen más palabras es mejor que sólo palabras; aunque, como es lógico, el efecto se limita a los aprendices con poco conocimiento previo.

2.1.2.2. Desventajas. Con respecto a la inclusión de imágenes no todo es coser y cantar, desafortunadamente. Para empezar, la imagen necesita ser procesada, lo cual es exigente. La imagen, nos recuerda Lowe (2004) muestran tres tipos de cambios que han de ser procesados: (a) de transformación (los atributos de la imagen, como el color o la forma, cambian), (b) de traslación (la posición de las imágenes cambia), (c) de transición (las entidades aparecen y desaparecen).

En segundo lugar, imágenes y palabras deben ser integradas mentalmente. Deben establecerse las conexiones correspondientes entre imágenes y texto y eso, como veremos un poco más abajo, tampoco es fácil. Remitimos al lector al apartado 2.3.

En suma, la inclusión de imágenes tiene también su contrapartida. Las imágenes deben ser (a) procesadas, lo que no es fácil dados los cambios múltiples que las imágenes sufren, y (b) deben ser integradas con la información verbal, lo que tampoco es fácil -como luego explicaremos con detalle.

2.2. Obstáculo 1: tópicos complejos.

Vamos a atender ahora a los obstáculos que un aprendiz debe salvar para alcanzar el aprendizaje del que hemos hablado hasta hace un momento. Comenzaremos con el primero obstáculo: los tópicos de los CBLEs son complejos. Basta un repaso somero para advertir que los CBLEs enseñan sobre tópicos complejos. Complejos, porque entrañan un buen número de ideas y relaciones entre ellas conformando intrincados sistemas conceptuales. Los aprendices, consecuentemente, han de identificar un número nada despreciable de ideas y han de ser capaces de establecer relaciones entre ellas, lo cual no es en absoluto fácil. De modo que (a) los temas son complejos y eso (b) supone un escollo para los aprendices. En lo que sigue argumentamos estas dos afirmaciones.

Por lo que respecta a los tópicos tratados en los CBLEs, cabe hacer un listado de tópicos a partir de una muestra (no estrictamente) aleatoria de estudios. Ello sirve para ilustrar la complejidad habitual de los tópicos en los CBLEs.

Tabla 16. Tópicos complejos en los CBLEs.

Tópico tratado en el CBLE	Referencia
Geometría, Teoremas	Aleven & Koedinger, 2002
Matemáticas, Probabilidad	Atkinson, Renkl, & Merrill, 2003
Biología, Sistema circulatorio humano	Azevedo, Moos, Greene, Winters, & Cromley, 2008
Psicología, Aprendizaje asociativo	Bannert & Mengelkamp, 2008
Matemáticas, Probabilidad	Berthold, Eysinck, & Renkl, 2008
Estadística, Conceptos sobre el análisis de varianza	Bodemer, Ploetzner, Feuerlein, & Spada, 2004
Física, Leyes de Newton	Conati & VanLehn, 2000
Geología, Erupciones volcánicas	Graesser, Wiley, Goldman, O'Reilly, Jeon, & McDaniel, 2007
Biología, Sistema circulatorio humano	Hausmann & Chi, 2002
Geografía, Sedimentación y erosión	Mautone & Mayer, 2007
Ingeniería, Mecanismo de freno de tambor y de bomba de hinchar	Mayer & Anderson, 1992
Matemáticas, Probabilidad	Renkl, 2002
Psicología, Anorexia; Geografía, Aspectos físicos, políticos, económicos y sociales de Perú	Rouet, 2003
Biología, Mitosis celular	Scheiter, Gerjets, Huk, Imhof, & Kammerer, 2008
Bioquímica, Vitamina C y hierro en el metabolismo	Seufert, 2003

Otra forma de mostrar lo mismo es incluir un análisis más pormenorizado de uno de estos tópicos. Ello permite apreciar mejor la complejidad del mismo. Ofrecemos un análisis del tópico sobre el que el CBLE que empleamos en nuestro experimentos versaba. En concreto, se trata de un diagrama que captura el modelo mental que cabe construir a partir de nuestro material: ver Appendix (B).

Por lo que se refiere al modo como la complejidad supone un obstáculo para el aprendizaje cabe invocar las siguientes evidencias. Por un lado, hay evidencias que indican que la complejidad afecta al procesamiento (medidas on-line) de los contenidos a aprender. Por ejemplo, Coté, Goldman y Saul (1998) pidieron a un grupo de estudiantes de Primaria que aprendiesen geografía a partir de unos textos. Los textos fueron clasificados como fáciles o difíciles en función de varios criterios. En concreto, se aplicaron fórmulas de legibilidad (que computan la frecuencia léxica y la complejidad sintáctica de los textos) y se pidió a jueces expertos (profesores) e inexpertos (estudiantes semejantes a los utilizados en el experimento) que evaluaran la complejidad de los textos. La tarea de los participantes era leer los textos mientras pensaban en voz alta sobre ellos; una vez hecho eso, los participantes debían tratar de recordar todo lo posible sobre los textos. Los resultados mostraron que el número de elaboraciones se reducía significativamente ante los textos difíciles mientras que el número de problemas identificados se incrementaba drásticamente ante esos textos. Kintsch y Keenan (1973) solicitaron a un grupo de estudiantes que leyese un set de oraciones y párrafos para tratar de recordarlos justo después (como ya sabemos a partir del capítulo 2). Las oraciones/párrafos diferían en su número de proposiciones (manteniendo el número de

palabras bajo control). Los resultados mostraron que cuanto mayor complejidad (mayor número de proposiciones) mayor tiempo se requiere para procesar adecuadamente la oración/párrafo (es decir, para poder recordarlo después). De estos trabajos se deduce que la complejidad supone un obstáculo para el procesamiento: demanda un tiempo mayor de procesamiento y restringe la producción de elaboraciones.

Por otro lado, hay evidencias de que el aprendizaje (medidas off-line) también se ve afectado por la complejidad. Linderholm et al. (2000), a quien ya citamos, pusieron a un grupo de estudiantes universitarios a aprender historia a partir de varios textos. Se distinguió entre textos complejos y simples, en función (entre otras cosas) del número de conexiones causales entre los eventos descritos. Los resultados del estudio permiten apreciar que ante los textos complejos el rendimiento en tareas de recuerdo y de contestación a preguntas es significativamente peor, en comparación con lo que sucede ante textos simples. En los trabajos de Coté et al. (1998) y Kintsch y Keenan (1973), ya descritos, se observó algo parecido. En ambos trabajos el recuerdo de los contenidos a aprender fue significativamente peor en los textos complejos. De todo esto cabe deducir que la complejidad afecta negativamente al aprendizaje.

En resumen, en este apartado hemos comprobado que la complejidad es uno de los obstáculos al aprendizaje a partir de CBLEs. Hemos visto que un somero repaso a la literatura permite constatar que los tópicos sobre los que versan habitualmente los CBLEs son complejos. Hemos visto que hay evidencias de que dicha complejidad tiene un impacto negativo tanto el procesamiento como en el aprendizaje.

2.3. Obstáculo 2: habilidades especiales.

Los CBLEs son versátiles, como sabemos. Hacen posible la presentación de información no sólo verbal sino también pictórica; adicionalmente, permiten presentar los contenidos de forma no secuencial sino hipertextual. Ambas cualidades son armas de doble filo: pueden promover un mayor aprendizaje pero exigen la puesta en funcionamiento de habilidades especiales. Éstas son la integración de la información verbal y pictórica y la navegación en espacios hipertextuales.

2.3.1. Integración verbal-pictórica. Por lo que respecta a la integración de información verbal y pictórica cabe (a) analizar en qué consiste y (b) recabar evidencias de la dificultad que entraña. Y eso es justo lo que haremos a continuación.

¿Qué es integrar información verbal y pictórica? La integración se traduce en la puesta en funcionamiento de procesos físicos y mentales. Los procesos físicos son unos movimientos oculares específicos, que revelan una coordinación de la información verbal y pictórica. Así, si lo que debe hacer un aprendiz es aprender de un texto que contiene un diagrama, sus ojos han de concentrarse alternativamente en el texto y en el diagrama. Cabe esperar al menos dos patrones de alternancia, tal como describen Hegarty y Just (1993). Uno es local porque el salto del texto al diagrama se produce pieza a pieza: una vez leída una oración los ojos se desplazan al diagrama tratando de encontrar los referentes de la oración, tras eso los ojos regresan al texto para volver a empezar (i.e., texto – diagrama – texto – diagrama...). Otro es

global porque el salto se produce bloque a bloque: una vez leído un párrafo los ojos se desplazan al diagrama tratando de integrar las piezas del texto con ayuda del diagrama, después los ojos regresan al texto y comienza un nuevo ciclo (i.e., texto – texto – texto – diagrama – diagrama – diagrama...). Este segundo patrón supone que los aprendices construyen un pequeño modelo mental a partir del texto y lo complementan y ajustan examinando el diagrama completo: se trata pues de un patrón más sofisticado que el anterior.

Los procesos mentales que la integración verbal-pictórica exige son los de mapeado estructural. De acuerdo con Schnotz (2002), este proceso consiste en identificar correspondencias estructurales entre las representaciones verbal y pictórica; dicho de otro modo, se trata de encontrar relaciones de analogía entre ellas. El resultado de la aplicación de ambas clases de procesos (físicos y mentales) es la construcción de una representación mental integrada a partir de palabras e imágenes.

¿Es difícil llevar dichos procesos a cabo? Resulta que sí, como veremos enseguida. Para empezar, los movimientos oculares de alternancia texto-diagrama son locales. Hegarty y Just (1993) solicitaron a un grupo de estudiantes que aprendiesen cómo funcionan distintos tipos de poleas a partir de un texto que incluía un diagrama. La peculiaridad es que lo hicieron con un casco montado en su cabeza capaz de registrar los movimientos de sus ojos. Los análisis de sus movimientos oculares evidenciaron que había un alto número de inspecciones del diagrama por párrafo de texto y que a mayor número de oraciones en el texto mayor número de inspecciones del diagrama. Ello parece corresponderse con el patrón de alternancia local antes descrito. De modo que los aprendices habían de buscar los referentes de cada oración en el diagrama antes de seguir leyendo pero no conformaban una representación mental del texto antes de inspeccionar el diagrama para ajustar tal representación. Por otro lado, la integración verbal-pictórica no siempre tiene final feliz: no siempre se produce ese mapeado estructural que mencionábamos. Ainsworth, Bibby y Wood (2002) enseñaron a un grupo de estudiantes de Primaria a resolver problemas aritméticos con ayuda de un CBLE. Los investigadores diseñaron distintas versiones del CBLE. En la versión pictórica se presentaban distintas imágenes de un mismo fenómeno; en la versión matemática había distintas expresiones matemáticas de un mismo fenómeno; en la versión mixta había tanto imágenes como expresiones matemáticas. Así, aunque los aprendices debían coordinar distintas representaciones en todas las condiciones, el proceso de mapeado estructural era más complicado en la versión mixta (en tanto que la distancia entre representaciones es mayor). Los investigadores registraron el rendimiento de los aprendices bajo cada tipo de instrucción, no sólo para cada versión (pictórica/matemática/mixta) sino para cada tipo de información dentro de cada versión. Eso hizo posible calcular en cada una de las condiciones experimentales la correlación entre el rendimiento bajo un tipo de información y el rendimiento bajo el otro tipo de información. Si las correlaciones son altas, cabe interpretar que los aprendices establecen conexiones entre unos y otros tipos de información; una correlación baja indicaría lo contrario. En dos experimentos los resultados mostraron que, así como había una alta correlación entre rendimientos en las condiciones pictórica y matemática, la correlación era mínima en la condición mixta. Según Ainsworth et al., esto se explica porque los participantes en la condición mixta encontraron dificultades para integrar representaciones.

Otras evidencias de que no siempre es fácil integrar mentalmente palabras e imágenes proceden de los trabajos en los que imagen más palabras no ha sido mejor que sólo palabras. Es decir, hay un buen número de trabajos que proporcionan pruebas indirectas de la dificultad para integrar. La revisión de Tversky, Morrison y Betrancourt (2002) repasa varios de ellos.

En síntesis, integrar información verbal y pictórica implica la puesta en funcionamiento de procesos específicos. Estos son los patrones de inspección visual globales (físicos) y el mapeado estructural (mentales). El problema es que hay evidencias de que los aprendices encuentran difícil la ejecución de tales procesos.

2.3.2. Navegación en hipertexto. Una vez más debemos (a) especificar en qué consiste navegar a través de un espacio hipertextual y (b) mostrar pruebas de que no resulta fácil para los aprendices. Es justo lo que vamos a hacer ahora.

Navegar en un espacio hipertextual es buscar y utilizar información pertinente para una meta. Si, por ejemplo, uno surfea en internet buscando información sobre insecticidas habrá de tener claro qué es exactamente lo que le interesa (e.g., sprays mata-cucarachas), buscar páginas que aporten información fiable al respecto, seleccionar las que se ajusten a ese criterio (e.g., descartar las que hablen sobre polvos en lugar de sprays, rechazar los blogs y foros en pro de las páginas oficiales) y procesar e integrar la información que contengan (e.g., el spray mata-cucarachas actúa de modo X, puedo adquirirlo en X, a un precio de X). De manera que navegar implica varias cosas. Lazonder y Rouet (2008) concluyen, a partir de un análisis de los modelos de navegación más citados, que navegar incluye tres pasos básicos. Primero, definir la meta; segundo, buscar y evaluar información relevante; tercero, usar la información. Cada uno de los pasos, a su vez, conlleva la implementación de un buen número de procesos. Por ejemplo, evaluar la información exige leerla, comprenderla y compararla con un estándar; procesos que, además, pueden dividirse en sub-procesos, como son el reconocimiento de palabras, procesamiento sintáctico, generación de inferencias, etc. Buscar información, por su parte, implica servirse de los hipervínculos disponibles y llevar un registro mental de los lugares visitados y el modo como uno llegó a ellos.

La cuestión es que nada de eso es tarea fácil para los aprendices. Por ejemplo, muchos aprendices tienen dificultades para seleccionar las fuentes oportunas de información. Shapiro (1998) solicitó a un grupo de participantes que aprendiese sobre un período de la historia estadounidense (i.e., The Gilded Age, entre los años 1865 y 1900) a partir de un hipertexto para después elaborar un resumen y contestar a varias preguntas. El hipertexto contenía en torno a 20 nodos (documentos individuales) que versaban sobre los deportes, la vida familiar o la cultura del citado período de la Historia de los EEUU. Una versión los presentaba de manera jerárquica, es decir, los participantes podían acceder de un nodo a otro saltando de categoría en categoría (e.g., de deportes a cultura) o de nivel en nivel (e.g., del boxeo a las primeras críticas a la novela del período). Otra versión era lineal, es decir, presentaba los nodos de uno en uno igual que lo hace un texto convencional. Los análisis de las conductas de navegación de los participantes revelaron que los asignados a la condición jerárquica rechazaron consultar muchos nodos que eran claves para elaborar buenos resúmenes y contestar a las preguntas: muchos nodos quedaron sin ser visitados y sus resúmenes y rendimiento en las preguntas fueron peores que los de sus compañeros. Cataldo y Oakhill

(2000) pidieron a un grupo de participantes que trataran de comprender un par de relatos (uno sobre un asesino y otro sobre un soldado), cada uno de cuatro páginas. Al terminar de leerlos, los participantes tuvieron que resolver varias preguntas sobre los mismos. Posteriormente, los investigadores presentaron a los participantes las mismas preguntas y les invitaron a resolverlas de nuevo; pero esta vez les dejaron consultar los textos originales y registraron sus movimientos a lo largo de las páginas. Al analizar sus movimientos, Cataldo y Oakhill observaron que los comprendedores pobres (según una prueba estandarizada de comprensión) demostraban un patrón de búsqueda menos eficiente que el de los comprendedores buenos. Concretamente, los pobres comenzaban a inspeccionar los textos desde la primera línea, fuera cual fuera la pregunta; los buenos acudían al punto adecuado. Last, O'Donnell y Kelly (2001) hicieron que una muestra de estudiantes aprendiese sobre psicometría a partir de un hipertexto para contestar a unas preguntas. Muchos de los participantes (específicamente, los que tenían poco conocimiento previo sobre el tópico en cuestión) manifestaron no saber qué nodos habían visitado y cuáles no ni cuáles debían visitar para obtener la información buscada: estaban desorientados. Tomados en conjunto, estos resultados permiten constatar que al tener que navegar a través de un hipertexto no siempre es fácil buscar y seleccionar las fuentes relevantes de información para satisfacer las metas iniciales.

2.4. Obstáculo 3: conocimiento previo específico de dominio.

El conocimiento previo es imprescindible para comprender un material instructivo. Una afirmación de estas características requiere explicación adicional. Veamos (a) a qué nos referimos al hablar de conocimiento previo para después (b) comprobar cómo éste afecta sensiblemente al aprendizaje a partir de materiales instructivos.

Existen distintos tipos de conocimiento previo. Kintsch (1994b), por ejemplo, distingue conocimientos lingüísticos, conocimientos generales sobre el mundo y conocimientos específicos de dominio. Los primeros sirven, entre otras cosas, para acceder al significado de las palabras (“casa” es un “edificio para habitar”) y para asignar el referente a las anáforas (en “Carol-Anne se fue de compras. Ella compró unas zapatillas rosas” la expresión “Ella” se refiere a “Carol-Anne”). Los segundos son necesarios para realizar inferencias que conectan piezas consecutivas de información (en “Marina atravesó el porche de la casa; en la calle se encontró con Paloma” es necesario saber que un “porche” comunica la “casa” y la “calle” para entender la oración) o para sintetizar cierto volumen de información (el texto “Ricardo entró en el establecimiento; encargó una sopa al camarero; después de comerla solicitó la cuenta” puede resumirse a “Ricardo comió en un restaurante”, para lo cual es preciso saber qué suele ocurrir cuando acudimos a un restaurante). Los conocimientos previos específicos de dominio, que son los que nos interesan ahora, son conocimientos sobre el tópico particular sobre el que versa el material. Así, para entender que “en el Himalaya no hay volcanes porque ninguna placa subduce” uno debe saber que la formación de volcanes es el resultado del choque de placas tectónicas diferentes, siendo una más pesada, lo cual provoca que presione fuertemente a la menos pesada de forma tal que ésta se agrieta dejando paso al magma del interior terrestre, el cual aflorará a la superficie en forma de volcanes. Más concretamente, disponer de conocimiento previo (específico de dominio) implica contar con (a) más piezas de información, (b) que son correctas y (c) que están bien conectadas entre sí.

Por supuesto, contamos con pruebas empíricas del impacto de los conocimientos previos en el aprendizaje a partir de materiales instructivos. Por un lado, está documentado que disponer de conocimiento correcto e integrado tiene un efecto positivo sobre el aprendizaje. Una forma de confirmar dicho patrón es proporcionar a los aprendices el conocimiento previo necesario para aprender de un material dado antes de consultarlo y comparar su rendimiento con el de aprendices que no han recibido tal instrucción. Gilabert et al. (2005) solicitaron a un grupo de estudiantes que aprendiese sobre la Revolución Rusa a partir de un texto. La mitad de los participantes recibió una lección sobre movimientos socio-políticos en los siglos XVIII y XIX antes de leer el texto; el resto, leyeron el texto sin lección previa. Tras la lectura los participantes recordaron todo lo posible acerca de lo leído y respondieron a preguntas que exigían haber realizado inferencias. Los resultados de dos experimentos mostraron que quienes recibieron la lección previa rindieron mejor en la tarea de recuerdo y en las preguntas. McNamara y Kintsch (1996) encontraron un efecto semejante en uno de sus experimentos usando una lección previa más apegada al contenido del texto posterior. Otra forma de explorar el impacto del conocimiento previo es emplear materiales poco amables y evaluar si disponer de conocimiento previo puede mitigar tal circunstancia. Así, Voss y Silfies (1996) encontraron que el conocimiento previo puede compensar las carencias de un texto poco cohesivo. Los investigadores presentaron textos de historia con y sin lagunas (o fisuras en la línea causal) a lectores con alto y bajo conocimiento previo sobre historia. Después los participantes tuvieron que resolver varios sets de preguntas sobre lo que habían leído. Bajo la condición texto sin lagunas, el conocimiento previo sobre historia tenía un impacto insignificante; bajo la condición texto con lagunas, el conocimiento previo sobre historia se volvió crucial para responder bien a las preguntas.

Por otro lado, está documentado también que disponer de conocimiento incorrecto afecta negativamente al aprendizaje. Por ejemplo, Kendeou y van den Broek (2005) exploraron el efecto de las *misconceptions* sobre el aprendizaje. Por *misconceptions*, los investigadores entendían ideas (a) erróneas y (b) preconcebidas que los aprendices sostienen sobre un tópico. Erróneas porque no se corresponden con el saber normativo sobre el tema en cuestión. Preconcebidas porque no son distorsiones que surgen como una simplificación o confusión de los contenidos del material, sino que están presentes antes de estudiarlo (son fruto del pensamiento intuitivo del aprendiz). El caso es que Kendeou y van den Broek se entrevistaron con unos cuantos estudiantes para interrogarles sobre electricidad. Siguiendo una entrevista estructurada que incluía preguntas cerradas, preguntas abiertas y varias pruebas prácticas (usando materiales como una pila, una bombilla y unos cables) los investigadores pudieron identificar a los aprendices con y sin *misconceptions*. Una vez hecho eso, se les planteó la siguiente tarea a todos: debían leer un texto sobre electricidad oración a oración y revelar los pensamientos que acudiesen a su mente durante la lectura. Al terminar, los participantes trataron de recordar todo lo posible sobre el texto. Dicho método hizo posible examinar el efecto que la posesión de *misconceptions* tiene sobre el procesamiento y aprendizaje a partir de materiales instructivos. Kendeou y van den Broek observaron que los participantes con *misconceptions* generaron un mayor número de inferencias erróneas durante la lectura del texto y sus protocolos de recuerdo incluyeron un número menor de ideas correctas, en comparación con los aprendices sin *misconceptions*.

El problema, claro está, es que no todos poseen conocimiento previo suficiente. Más

bien, lo que predominan son novatos. Novato entendido en dos sentidos alternativos: por una parte, aprendices que cuentan con pocas piezas y/o poco interconectadas; por otra, aprendices que cuentan con conocimiento más o menos coherente pero erróneo. De acuerdo con Kintsch (1998) los primeros son los que construyen una representación mental “dominada por la base del texto” (mientras que si se ponen en juego los conocimientos previos la representación final está “dominada por el modelo de la situación”). Con esto Kintsch quiere decir que lo que pueden hacer los novatos de esta categoría es entresacar las ideas del texto y sus relaciones para formar una suerte de copia mental de los contenidos del texto; lo que no podrán hacer es agregar información adicional para enriquecer la representación resultante²⁶. La segunda clase de novatos es bien distinta, tal como describen deLeeuw y Chi (2002). En estos casos el aprendizaje también está dominado por lo que los aprendices ya saben (y no por lo contenido en el texto) pero lo que saben no se corresponde con un modelo científico o normativo. El novato posee un cuerpo de conocimiento sólido, en el sentido de que hay conexión entre sus piezas y que lo aplica consistentemente para resolver tareas; pero es conocimiento erróneo.

En resumen, disponer de conocimiento previo específico de dominio es clave. Lo es porque cuando se dispone del mismo se aprende más y mejor. Además, si el conocimiento del que se dispone es alternativo al normativo se aprende menos y peor. El asunto es que, por lo general, los aprendices no cuentan con vastos conocimientos y, si los tienen, presentan misconceptions.

2.5. Obstáculo 4: memoria de trabajo.

La memoria de trabajo, como es ya bien sabido, es el sistema que permite mantener y manipular la información necesaria para realizar tareas complejas como la comprensión o el razonamiento (e.g., Baddeley, 1997, 1998, 2003). Es, si se prefiere, el espacio donde la actividad mental de un momento dado tiene lugar. La cuestión es que su capacidad es limitada, restringiendo nuestras posibilidades de procesamiento. La memoria de trabajo es entonces crucial en el aprendizaje, puesto que es el sistema que permite procesar información, y lo restringe, puesto que la capacidad del sistema es limitada. Enseguida presentamos (a) pruebas de la relación entre la capacidad de la memoria de trabajo y el aprendizaje a partir de materiales instructivos y (b) el modo como la primera puede condicionar lo segundo.

Los datos indican que existe una correlación notable entre la capacidad de la memoria de trabajo y el aprendizaje a partir de textos (i.e., la comprensión lectora). El trabajo clásico a este respecto es el de Daneman y Carpenter (1980). Las autoras idearon una prueba que mide la capacidad de la memoria de trabajo en tareas de lectura: el “Reading Span Test”. Al aplicarla, el experimentador presenta al sujeto series de dos a seis oraciones que el sujeto debe leer en voz alta. Al mismo tiempo se le pide que retenga la última palabra de cada oración. De modo que el sujeto debe procesar las oraciones y retener las palabras clave simultáneamente. Tras la presentación de las series de oraciones se le solicita que recuerde las palabras que ha retenido y se le especifica que lo haga en el orden en que fueron presentadas. Por supuesto, las

²⁶ Lo cierto es que aunque los conocimientos no sean vastos pueden aplicarse ciertas estrategias para enriquecer la representación resultante. Por ejemplo, Chi, Bassok, Lewis, Reimann y Glaser (1989) encontraron que algunos participantes aplicaban sus conocimientos generales sobre el mundo y derivaban información de la definición de algunos términos aparecidos en el texto para aprender más sobre física.

palabras a recordar no guardan relación semántica alguna. Las series de oraciones están ordenadas en secuencia creciente: cada nueva serie es más larga que la anterior. Existen cinco niveles. Cada serie contiene grupos de tres oraciones. La serie en la que el sujeto exhibe un rendimiento nulo da por finalizada la prueba. Se considera que la capacidad de su memoria de trabajo se corresponde con el nivel en el que el sujeto es capaz de reproducir al menos dos secuencias de palabras clave de las tres contenidas en una serie. Las autoras aplicaron esta prueba a un conjunto de lectores. Después solicitaron a los participantes que leyesen un texto para tratar de contestar a unas preguntas facilitadas posteriormente. El rendimiento en las preguntas reflejaba su nivel de comprensión. La correlación entre las puntuaciones en una y otra prueba fue significativa (de más de .50, o sea, explica en torno al 30% de la varianza). El resultado del estudio de Daneman y Carpenter ha sido replicado en trabajos posteriores. Por ejemplo, Elosúa, Gutiérrez, García Madruga, Luque y Gárate adaptaron la prueba al castellano en 1996. Los españoles registraron una correlación significativa entre memoria de trabajo y comprensión de .31 (es decir, explica aproximadamente un 10% de la varianza). La medida de comprensión empleada en este caso fue el recuerdo de las ideas principales del texto.

La memoria de trabajo tiene relación con el aprendizaje, como hemos visto, pero podemos ir más allá. Cabe afirmar que la contribución de la memoria de trabajo al aprendizaje es específica, en tanto que explica una porción de su varianza *por encima de* otras variables. Cain, Oakhill y Bryant (2004) aplicaron un set de pruebas a una muestra de estudiantes de primaria. En concreto, los participantes tuvieron que resolver una prueba de comprensión lectora, una de lectura de palabras, una de vocabulario, una de inteligencia verbal, varias de componentes específicos del aprendizaje a partir de materiales instructivos (inferencia, supervisión, conocimiento estructural) y dos de memoria de trabajo. Las pruebas de memoria eran similares a la de Daneman y Carpenter. Una de ellas empleaba material verbal y la otra numérico. Las mismas pruebas se aplicaron en tres ocasiones: cuando los niños tenían ocho años, a sus nueve años y a sus once años. Un resultado muy interesante fue que las pruebas de memoria tomadas conjuntamente predecían una porción de varianza en comprensión lectora, una vez aislado el efecto de la lectura de palabras, el vocabulario y la inteligencia verbal. Lo mismo sucedió en los tres momentos de medición. Así que la contribución de la memoria de trabajo es genuina, no está mezclada con la de otras variables.

La memoria de trabajo no sólo está relacionada con el aprendizaje a partir de materiales instructivos en general sino también con sus componentes. Un componente esencial del aprendizaje es la capacidad para supervisar si uno está entendiendo bien o no el material que tiene entre manos. Cain et al. (2004) midieron esta capacidad presentando textos inconsistentes a sus participantes a fin de que éstos pudiesen identificar el problema (“un momento... esto no tiene sentido”). Pues bien, la memoria de trabajo estuvo correlacionada con la capacidad para supervisar en los tres momentos de medición.

En virtud de lo expuesto hasta aquí parece probado que la memoria de trabajo está relacionada con el aprendizaje. Relación, claro está, que debe significar crecimiento/decrecimiento directamente proporcional. Ahora bien, hay pruebas más contundentes de que el aprendizaje puede verse condicionado si la capacidad de la memoria de trabajo es particularmente limitada. Resulta que una de las diferencias entre los que

aprenden a partir de textos y los que no encuentran tan fácil dicha empresa radica en la capacidad de su memoria de trabajo. Yuill, Oakhill y Parkin (1989) distinguieron entre buenos y malos comprendedores a partir de una prueba estándar en comprensión lectora. A ambos grupos les aplicaron una prueba de memoria de trabajo basada en números. Al comparar el rendimiento de unos y otros ante la prueba se hizo evidente que los malos cuentan con una capacidad de memoria de trabajo significativamente menor que la de sus compañeros. En una segunda fase de su estudio, Yuill y sus colegas plantearon la siguiente prueba a los mismos buenos y malos comprendedores. Les presentaron un set de textos a comprender. Algunos de los textos contenían una afirmación inconsistente con las incluidas anteriormente (“Tommy no compartió los caramelos con su hermano. Al ver esto, su madre agradeció el gesto de Tommy”). Parte de éstos incorporaba la resolución (“El hermano de Tommy estaba bastante grueso y seguía una dieta”) justo después de la incongruencia o un poco más tarde. La tarea de los participantes era responder a una pregunta que reflejaba si habían comprendido la aparente inconsistencia del relato. Los resultados mostraron que buenos y malos comprendedores rendían igual cuando la resolución aparecía enseguida pero los buenos eran significativamente mejores que los malos cuando había demora en la resolución. Esto sugiere que los malos comprendedores tienen dificultades a la hora de resolver tareas que demandan mantener activa cierta información, i.e., tareas que ponen en juego la capacidad de la memoria de trabajo.

Otra manera de mostrar que la capacidad limitada de la memoria de trabajo puede perjudicar el aprendizaje es la siguiente. Aprender supone mantener y manipular muchos elementos (e.g., ideas, expresiones numéricas) y sus relaciones al mismo tiempo; si manipulamos artificialmente el número de elementos/relaciones podemos mostrar que cierto volumen puede obstaculizar el aprendizaje. Algo como eso haría evidentes las limitaciones de nuestro sistema de mantenimiento y manipulación de información. Pues eso es precisamente lo que hicieron Pollock, Chandler y Sweller en 2002. Los investigadores reclutaron a un grupo de estudiantes profanos en materia de electricidad. Pollock et al. pretendían que sus participantes aprendiesen cómo llevar a cabo tests de seguridad eléctrica y cuál es el razonamiento que subyace a los mismos. Un conocimiento que, al menos para los profanos, exige procesar al mismo tiempo un buen número de elementos y relaciones. Lo que manipularon Pollock et al. fue el libreto a partir del cual los participantes habrían de aprender sobre electricidad. Un libreto presentaba sintéticamente un diagrama mostrando los materiales relevantes (un medidor de amperios, una kettle, etc.) y los pasos a seguir para realizar el test de seguridad eléctrica. El otro, más completo, presentaba todo eso y explicaciones de cada uno de los pasos: no sólo qué hay que hacer sino por qué y qué cosas pueden suceder si algún parámetro cambia. Los participantes, en función de la condición experimental, recibieron primero el libreto sintético y luego el completo o el completo primero y después (i.e., sintético-completo versus completo-completo). Tras las dos fases de consulta de los libretos los participantes resolvieron varias tareas: contestaron a preguntas y realizaron pruebas prácticas. Los resultados de dos experimentos revelaron que los de la condición sintético-completo rindieron significativamente mejor que los de la condición completo-completo en las tareas. Pero hay más todavía. Pollock et al. realizaron el mismo experimento también con aprendices versados en la materia. Los investigadores se aseguraron de que los participantes tenían conocimientos sobre electrónica y que los distintos niveles de conocimiento quedaban igualados en las dos condiciones. La cuestión es que para los expertos el volumen de

información a procesar es menor, con respecto a los novatos. Es menor porque ellos pueden formar esquemas, chunks o unidades significativas que aglutinan elementos individuales. Así, para una persona que no sabe nada de ajedrez memorizar una configuración de fichas sobre un tablero puede resultar un auténtico desafío (e.g., “una ficha con forma de torre junto a otra con una crucecilla encima... ”); mientras que para el mítico Kasparov (hoy metido en política) dicha tarea no reviste dificultad alguna (e.g., “jaque mate”). Si esto es cierto, entonces el formato del libreto sobre tests de seguridad eléctricos no debería marcar diferencias como lo hacía en el caso de los profanos. Así fue. En dos experimentos Pollock et al. no encontraron diferencias significativas entre los rendimientos de los participantes de la condición sintético-completo y los de la completo-completo. Esto pone de relieve las restricciones de la memoria de trabajo: no se puede procesar mucho a la vez; si se intenta hacer tal cosa se produce una sobrecarga que perjudica el aprendizaje. Cuando el volumen de información no es grande (como en el caso de los expertos) la memoria de trabajo cumple sin problema.

En resumidas cuentas, la memoria de trabajo guarda relación con el aprendizaje. Es más, explica una parte de la varianza en tareas de aprendizaje por encima de otras variables que sabemos importantes. Más específicamente, hay pruebas directas de que las limitaciones de la memoria de trabajo pueden obstaculizar el aprendizaje. Así, los malos comprendedores rinden peor en tareas de memoria de trabajo y cuando se incrementa el número de elementos a procesar el rendimiento decae.

2.6. Interrelación entre las dificultades: la paradoja.

Habíamos asumido que los CBLEs deben incorporar ayudas dadas las dificultades que los aprendices han de sortear cuando aprenden de ellos. En este apartado lo que hemos hecho es describir esas dificultades con cierto detalle y hemos revisado estudios en los que se pone de manifiesto que tales dificultades entorpecen, de hecho, el aprendizaje.

Los temas sobre los que versan los CBLEs son complejos. Eso lo hemos comprobado examinando los CBLEs empleados por una muestra de estudios y constatando que versan sobre temas complejos (psicología, geología, ingeniería, etc.). Y la complejidad tiene un impacto sobre el aprendizaje. Eso lo hemos apreciado al revisar estudios en los que se constata que los materiales complejos afectan al procesamiento (e.g., Coté et al., 1998) y al aprendizaje (e.g., Linderholm et al., 2000): cuanto más complejo, menos elaboraciones se generan, más tiempo de procesamiento es requerido, menos es recordado y aprendido.

Los CBLEs permiten presentar información verbal y pictórica en secuencia no lineal. Estas posibilidades demandan a los aprendices la puesta en funcionamiento de procesos específicos. Éstos son la integración verbal-pictórica y la navegación. La integración requiere una inspección visual de los materiales particular y un proceso de mapeado estructural o búsqueda de analogías. El problema, tal como hemos visto, es que la inspección que se lleva a cabo es local (Hegarty & Just, 1993) y que el mapeado no siempre tiene lugar (Ainsworth et al., 2002). La navegación supone fijar un objetivo de búsqueda, buscar y seleccionar las fuentes de información adecuadas y, finalmente, utilizarlas. El problema es que sabemos que muchos experimentan dificultades al tener que buscar y evaluar las fuentes y terminan desorientados (e.g., Last et al., 2001; Shapiro, 1998).

Comprender y aprender de un material instructivo precisa disponer de conocimientos específicos. De hecho, cuando proporcionamos artificialmente conocimientos necesarios para entender un material, los aprendices aprenden más (e.g., Gilabert et al., 2005). Y esto funciona en sentido contrario: cuando se poseen conocimientos pero son erróneos, el procesamiento de la información se ve contaminado y se termina aprendiendo menos y peor (e.g., Kendeou & van den Broek, 2005). Y lo problemático es que, por lo general, lo que predominan son novatos en los dos sentidos: o no tienen conocimiento o el que tienen es incorrecto.

Las ideas y relaciones que uno aprende deben procesarse en el espacio limitado de la memoria de trabajo. Esto implica que el volumen de información que uno puede procesar de una vez es acotado. Por eso la capacidad de la memoria de trabajo guarda relación con la capacidad para aprender de materiales instructivos (e.g., Daneman & Carpenter, 1980) y por eso cuando la memoria de trabajo se ve especialmente constreñida el aprendizaje se ve seriamente perjudicado (e.g., Pollock et al., 2002).

Tabla 17. Los obstáculos que deben salvar quienes aprenden a partir de un CBLE.

Dificultad		Descripción	Evidencias
Tópicos complejos		Los temas tratados en los CBLE incluyen muchas ideas y relaciones	Cuanto más complejo es un material menos elaboraciones se generan y peor se comprende (e.g., Coté et al., 1998)
Habilidades especiales	Integración verbal-pictórica	Establecer conexiones entre lo verbal y pictórico a partir de procesos conductuales y mentales	Cuando hay que manejar representaciones distintas de un mismo concepto experimentamos dificultades (e.g., Ainsworth et al., 2002)
	Navegación	Saber qué buscas, buscar y seleccionar, usar	Las estructuras hipertextuales nos desorientan (e.g., Shapiro, 1998)
Conocimiento previo escaso		Los conocimientos son imprescindibles para comprender pero o no se tienen o se tienen misconceptions	Las misconceptions provocan la generación de inferencias erróneas y perjudican el aprendizaje (e.g., Kendeou & van den Broek, 2005)
Memoria de trabajo		La capacidad de procesamiento con que contamos es limitada	Quienes disponen de menos capacidad aprenden peor (e.g., Yuill et al., 1989)

Cabe ahora establecer relaciones entre estos cuatro obstáculos. Tenemos, por un lado, un volumen de información considerable que debe ser procesado (muchas ideas y relaciones de contenidos verbales y pictóricos presentados no linealmente) y, por otro lado, un sistema de procesamiento seriamente limitado (capacidad restringida de la memoria de trabajo y aún más acotada por la presencia de conocimientos erróneos y ausencia de chunks). O sea, tenemos unos materiales muy demandantes para unos aprendices poco preparados. El resultado, como es obvio, es que aprender de los CBLEs no es fácil. Pero podemos, con lo que sabemos ahora, precisar algo más esta afirmación. Lo precisamos acto seguido.

La complejidad de los tópicos a aprender unida a las conexiones que deben establecerse entre la información verbal y pictórica y entre los nodos de un hipertexto dan como resultado un CBLE realmente demandante. Esto es, un CBLE que demanda al aprendiz realizar muchas operaciones mentales. El sistema de procesamiento es reducido, de modo que el aprendiz no podrá realizar todas las operaciones que se esperan de él. Pero, además, la capacidad de procesamiento se ve más acotada aún por la presencia de *misconceptions* (que actúan como elementos extraños) y la ausencia de *chunks* o esquemas (que podrían reducir el volumen de información a procesar). Así que tenemos un volumen de información y operaciones enorme (temas complejos presentados en múltiples formatos e hipervinculados) para un sistema procesador doblemente restringido (por su propia naturaleza y por la condición novata de los aprendices). El resultado de tal paradoja es la construcción de representaciones mentales pobres. Pobres porque faltan ideas importantes, porque faltan conexiones entre lo verbal y pictórico, porque sobran ideas erróneas... Pobres, al fin y al cabo.

En el apartado siguiente veremos que la única forma que tiene el aprendiz para no sucumbir ante esta paradoja es llevando a cabo un aprendizaje estratégico. Un aprendizaje estratégico es aquel en el que se supervisa el proceso de aprendizaje en busca de problemas que puedan surgir (e.g., simplificaciones, distorsiones) y en el que se realizan ajustes para reparar los problemas. Ser estratégico es, pues, detectar y reparar malentendidos. Veremos también que, aunque críticos, estos procesos no son ejecutados con facilidad; ello hará evidente la necesidad de proporcionar ayudas que asistan a los aprendices en la detección y reparación y que, por extensión, permitan soslayar la paradoja antes descrita.

En conclusión, dadas las demandas de los CBLE y las limitaciones de procesamiento de los aprendices, lo más probable es que la paradoja se salde con un aprendizaje pobre. Para evitar que esto suceda el aprendiz no puede simplemente dejarse llevar: ha de poner en juego un procesamiento estratégico con el que detectar y reparar malos entendidos. A continuación describimos en qué consisten estos procesos y argumentamos por qué es preciso apoyarlos ofreciendo ayudas.

3. Las ayudas para detectar y reparar.

Acabamos de plantear una seria paradoja: materiales demandantes para aprendices poco preparados. El resultado más probable de dicha paradoja es la construcción de representaciones mentales pobres, como sabemos. Hemos anticipado, además, que para que tal resultado indeseable no tenga lugar los aprendices deben desplegar un procesamiento estratégico. En esta sección describimos qué entendemos nosotros por este tipo de procesamiento. Después presentaremos distintos grupos de evidencias que avalan el peso que el aprendizaje estratégico tiene sobre el aprendizaje. Presentaremos también indicadores de que es muy común que los aprendices no lleven a cabo un aprendizaje estratégico, lo que justificará la necesidad de propiciar y apoyar un procesamiento de tales características. Haremos un pequeño repaso de los tipos de ayuda que se han propuesto para tal fin, atendiendo además a los estudios empíricos que avalan su eficacia.

3.1. Aprendizaje estratégico: detectar y reparar.

Los procesos de comprensión de textos y de integración de palabras e imágenes que describimos más arriba podrían llevarse a cabo de maneras distintas. Una posibilidad es que se produzcan sin que seamos conscientes de ello, sin que lo hayamos decidido y sin que nos cueste trabajo. Esto es, los procesos podrían ejecutarse de manera *automática*. Otra posibilidad es que tengamos que activarlos consciente y deliberadamente y que requieran esfuerzo. Esto es, los procesos podrían operar de manera *estratégica*. ¿Son automáticos o estratégicos?

Ni sí, ni no, sino todo lo contrario. Kintsch (1998) resuelve la pregunta sentenciando que la comprensión de textos es algo “entre la percepción y la resolución de problemas”. O sea, la comprensión está a medio camino entre lo automático y lo estratégico, desplazándose a un polo u otro en función de las circunstancias²⁷ (Gernsbacher (1995) o van den Broek, Rapp y Kendeou (2005) plantean lo mismo).

3.1.1. Procesamiento automático. ¿Cómo la comprensión podría ser igual a la percepción? Kintsch plantea un modelo de comprensión en dos fases: Construcción e Integración (modelo CI). La fase de *construcción* arranca con la construcción de proposiciones o ideas (previo procesamiento léxico y sintáctico). Estas proposiciones activan otras almacenadas en nuestra memoria con las que existe una asociación. La activación es insensible al contexto: toda proposición de nuestra memoria vinculada a las que actualmente están siendo procesadas es susceptible de ser activada (e.g., “se acercó a un banco” podría activar “no tengo dinero” y/o “los bancos son para sentarse”). Todo depende de la probabilidad. Cuanto mayor es la conexión entre la proposición de la memoria y las que se han construido, más probable es que aquella sea activada. Una vez conformada una red por este procedimiento asociativo, los niveles de activación de las proposiciones interactúan: reforzándose y suprimiéndose mutuamente. Es la fase de *integración*. Un mecanismo de propagación de la activación se inicia para estabilizar la red. En el momento inicial cada proposición posee un valor de activación. Además, las proposiciones están interconectadas de forma positiva -cuando hay asociación entre ellas- o negativa -cuando son dos ideas incompatibles. A través de las conexiones, las proposiciones se refuerzan y suprimen mutuamente. Las proposiciones que reciben mayor activación por parte de otras resultan fortalecidas cuando la activación se propaga. Las que reciben conexiones negativas se atenúan y desaparecen de la red. Aquí es donde el contexto ejerce su influencia: las ideas irrelevantes no logran sobrevivir. El proceso de propagación continúa hasta que la red se estabiliza y así concluye lo que el autor llama *constraint-satisfaction*. La red resultante, que es una representación mental provisional, pasa a un segundo plano. Una o dos ideas, las que han recibido mayor activación, podrán mantenerse activas para el ciclo de construcción-integración siguiente. Terminado el siguiente ciclo ocurrirá lo mismo: la nueva red pasará a segundo plano quedando un par de proposiciones en el foco de atención. Y así sucesivamente. El resultado final: una representación que incluye las ideas más activadas y conexiones entre ellas.

En la secuencia que hemos descrito parece no haber lugar para un proceso consciente

²⁷ Recuérdese que lo mismo se planteó en el capítulo 1 a propósito de la comprensión del lenguaje.

y deliberado como el que ocupa al sujeto que resuelve problemas. Pero eso no impide que los procesos inferenciales sean ejecutados. Es perfectamente factible que al leer “llovía a cántaros” se activen proposiciones como “el paraguas protege de la lluvia” o “me molesta quedar empapado”. Cuando poco después leemos “decidí volver a casa a por el paraguas” entendemos instantáneamente la relación entre la primera y la segunda frase: “volvió a por el paraguas *para* protegerse de la lluvia”. Lo que el lector habría hecho en ese caso es una inferencia, pero siguiendo el proceso asociativo que plantea Kintsch. Algo parecido tendría lugar al desentrañar de un material sus ideas principales. En el ejemplo anterior del restaurante (i.e., “Ricardo entró en el establecimiento; encargó una sopa al camarero; después de comerla solicitó la cuenta”) todas y cada una de las frases podrían evocar la proposición “comer en un restaurante” de forma que al terminar la lectura ésta fuese una proposición que habría recibido mucha activación de las otras proposiciones. Por eso esa idea que sintetiza otras formaría parte de la representación final. En estas situaciones la comprensión fluye como la percepción: de forma inconsciente, involuntaria y sin esfuerzo; y ello no significa que se comprenda peor.

3.1.2. Procesamiento estratégico. ¿Cuándo la comprensión puede convertirse en un problema a resolver? Fiarnos de un proceso asociativo como el descrito no tiene por qué dar siempre buenos resultados. Hay circunstancias que lo pueden menoscabar. Si el material es muy complejo, no disponemos de conocimientos previos, poseemos ideas erróneas sobre el contenido a aprender o la capacidad de nuestra memoria de trabajo está limitada por alguna razón, la cosa puede torcerse: un procesamiento automático con seguridad nos conducirá a un aprendizaje pobre. Son justo los obstáculos de los que ya hemos hablado largo y tendido. Hay entonces una oportunidad para mejorar las cosas: el procesamiento estratégico. Ser estratégico, ahora sí, es como resolver un problema: se define el espacio del problema y se trata de encontrar el camino a la solución; para ello se seleccionan y sopesan distintas alternativas, se las pone a prueba y se evalúan los resultados. Sin duda una actividad de gran consumo intelectual. Algo así nos exige conciencia, deliberación y esfuerzo.

Por ejemplo, si nos topamos con un texto como éste “la bujía se gastó; el coche se detuvo” sin los conocimientos necesarios para entenderlo podemos iniciar un proceso estratégico como el que sigue.

- (41) “No entiendo qué relación, si existe alguna, hay entre estos dos eventos. Podría ser que la bujía sea responsable de alguna parte del proceso de automoción de un coche... Un momento: líneas atrás me dijeron que el movimiento del coche es resultado de la explosión en el motor, explosión detonada por la chispa de una bujía... Entonces ya sé la relación entre los eventos.”

La verdad es que plasmado así (41) resulta rocambolesco. Pero un proceso semejante es el que tiene lugar en nuestra mente cuando consciente y deliberadamente nos esforzamos por comprender. El proceso estratégico garantizaría que, aun en presencia de las circunstancias u obstáculos que ya hemos listado, la comprensión fuese exitosa.

3.1.3. Detectar y reparar. Hay varios modelos sobre cómo aprendemos estratégicamente (Chi, 2000; deLeeuw & Chi, 2002; Hacker, 1998; Otero, 2002; Thiede &

Dunlosky, 1999; Winne, 1995). Todos ellos pueden resumirse en dos componentes o procesos: uno que sirve para percatarse de que algo no va bien y otro que sirve para tomar las medidas oportunas para enmendarlo.

Otero (2002), por ejemplo, propone su particular modelo de comprensión estratégica de textos. Para empezar, según Otero fijamos un *minimum acceptance level* (nivel de aceptación mínima o MAL) al estudiar un texto. El MAL establece el nivel de comprensión que vamos a considerar apropiado. En segundo lugar, podemos verificar periódicamente el nivel de *goodness* (bondad o G) que corresponde al grado de ajuste/discrepancia entre el nivel de comprensión alcanzado en un momento dado y el MAL. Tercero, si G es menor que MAL, esto es, si el nivel de comprensión alcanzado no se ajusta al considerado apropiado, iniciamos un proceso de regulación. La regulación puede ser nula, si no hacemos nada por aproximar G y MAL (e.g., “vaya, no veo la conexión bujía-coche detenido... ...”). Puede ser adecuada si tomamos las medidas oportunas (e.g., consultar una fuente externa de información, realizar inferencias basadas en lo que uno ha leído). Y puede ser inadecuada si realizamos inferencias no derivables de la información provista por el texto (e.g., “la bujía es otro nombre para llamar a la gasolina, por eso el coche se paró”). En suma, hay un proceso de identificación de problemas (cuando G es menor que MAL) y otro de resolución. La correcta ejecución de ambos permitiría garantizar la comprensión en los momentos en que se ve amenazada.

Chi (2000; deLeeuw & Chi, 2002) proponen un modelo similar. Se apoya en varios supuestos. Primero, los aprendices construyen modelos antes (basados en concepciones previas) o durante la consulta de un material instructivo. Segundo, estos modelos *suelen* ser *flawed* (defectuosos) en el sentido de que, en lugar de tener lagunas y débiles conexiones entre sus elementos, son completos y coherentes pero compuestos por elementos erróneos. Un modelo *flawed* permite resolver tareas de transferencia pero erróneamente (e.g., “¿qué podrías hacer si la bujía de tu coche se gasta?” - “llenar el depósito de gasolina o buscar un combustible alternativo”). Si son frecuentes los modelos *flawed* es porque (a) es común sostener ideas ingenuas sobre fenómenos físicos (e.g., “un objeto arrojado desde un avión en marcha caería al suelo dibujando una línea recta”) y (b) aprender a partir de un material instructivo plantea serios obstáculos, como sabemos. Tercero, cuando el aprendiz detecta una discrepancia entre lo que él piensa y lo que el material dice, arranca un proceso de reparación. La reparación consiste en la generación de inferencias que resuelven las discrepancias encontradas. En suma, hay un proceso de *detección* y *reparación* de problemas de comprensión. Ésta es la nomenclatura que utilizamos aquí.

Debemos advertir que los modelos de Otero y Chi describen un comportamiento ideal. Es obvio que si uno detecta adecuadamente, parece natural que comience a reparar y, una vez hecho eso, que la comprensión se reestablezca. De hecho, como se verá en el apartado próximo, realizar estos procesos es clave para aprender. Pero ¿y si no detectamos? Puede ocurrir que no supervisemos nuestra comprensión para identificar problemas o que, teniéndolos delante, no sepamos verlos. ¿Y si no reparamos? Puede ocurrir también que una vez detectados iniciemos un proceso de reparación pero que éste no sea el más apropiado o incluso que nuestra reparación consista en no hacer nada. Estos problemas son más comunes de lo que uno quisiera y es lo que comprobaremos en el apartado 3.3.

3.2. La importancia de detectar y reparar.

Son muchas las evidencias que indican que detectar y reparar son procesos clave. En concreto, hay estudios que comparan buenos y malos aprendices, estudios de regresión y estudios de entrenamiento y todos ellos aportan datos que apuntan en la misma dirección. ¿Y que revelan los resultados de tales estudios? Los aprendices buenos se diferencian de los malos porque los primeros detectan y reparan problemas de comprensión; la capacidad para detectar y reparar explica una porción significativa de la varianza en aprendizaje; cuando se entrena a los aprendices para que lleven a cabo procesos de detección y reparación su aprendizaje se incrementa significativamente. O sea, detectar y reparar son procesos críticos para el aprendizaje. Lo que vamos a hacer a continuación es repasar algunos de estos trabajos.

3.2.1. Estudios de comparación buenos/malos. Si reunimos una muestra de aprendices, distinguimos (atendiendo a un criterio oportuno) entre buenos y malos y medimos la capacidad para detectar y reparar de unos y otros, ¿qué es lo que encontramos? Resulta que los primeros rinden mejor que los segundos en tareas que exigen detectar y reparar. Así que estos procesos son algunas de las armas en el arsenal de los buenos aprendices; dicho de otro modo, si quieres ser buen aprendiz éstos son dos de los procesos que deben formar parte de tu repertorio.

En cuanto al proceso de detección vamos a repasar un par de trabajos. En ambos se pondrá de manifiesto que los malos aprendices supervisan peor su aprendizaje, en comparación con los buenos. Maki y Berry (1984) pidieron a un grupo de estudiantes que leyese varios textos sobre psicología para responder a preguntas de alternativa-múltiple que serían planteadas después. Basándose en la comprensión que los participantes exhibieron del primero de los textos, Maki y Berry distinguieron entre buenos y malos aprendices. La tarea de los participantes era predecir qué tal rendirían en las preguntas de alternativa-múltiple que más tarde habrían de responder. Una vez hecho todo eso, los investigadores calcularon la correlación entre las predicciones de los participantes y el nivel de comprensión real demostrado en las preguntas. Eso permitía evaluar la precisión que los aprendices demuestran al evaluar su propio aprendizaje. Los resultados mostraron que los buenos eran significativamente más precisos que los malos al estimar su rendimiento. Rubman y Waters (2000) reunieron una muestra de estudiantes. A partir de sus puntuaciones en una prueba estándar de comprensión de textos (California Achievement Test) y de las evaluaciones de sus profesores, Rubman y Waters dividieron la muestra en dos: buenos y malos aprendices. A todos ellos se les planteó la tarea de leer varios relatos y evaluar si tenían o no sentido. Los investigadores habían introducido varios textos modificados entre los relatos. Los textos modificados presentaban inconsistencias del tipo “...había estado nevando pero era demasiado pronto para que el lago estuviese helado... después de hacer un muñeco de nieve, los niños estuvieron patinando sobre el lago mientras un pez saltaba fuera del agua” que los participantes tendrían que identificar. Los resultados revelaron que los malos aprendices eran menos capaces de detectar las inconsistencias, con respecto a los buenos. En conjunto, los malos aprendices encuentran más dificultades cuando se trata de supervisar si uno entiende o no el material a aprender.

Respecto del proceso de reparación, veremos que sucede algo análogo. También aquí

los malos aprendices son menos capaces de revisar y reparar sus representaciones mentales. Hyönä et al. (2002) solicitaron a un grupo de participantes que leyese varios textos sobre medio ambiente para elaborar un resumen. Hyönä et al. registraron los movimientos oculares de los participantes durante la fase de estudio de los textos. Sobre la base de sus resúmenes posteriores, pudo comprobarse que había diferencias notables entre unos participantes y otros. Lo interesante es que cada cual mostró un patrón de movimientos oculares específico. Los autores de los resúmenes de menor calidad eran justamente quienes habían leído de forma lineal, es decir, sin retroceder y sin asignar mayor tiempo a los segmentos importantes. Lo cierto es que no hay medida directa de su capacidad para detectar malentendidos; pero si suponemos que en algún ocasión los malos aprendices notaron que algo iba mal, sus patrones de movimiento ocular indican que no son capaces de solventarlo. Imaginemos que uno se da cuenta de que ha malentendido algo en la línea 15 del texto. Lo que hace entonces es volver a atrás y se detiene en el lugar oportuno para aclarar el malentendido. Un comportamiento de esas características se hubiese manifestado en forma de regresiones y diferencias en los tiempos asignados a cada segmento, cosas ambas que brillan por su ausencia en los patrones de los malos. Este razonamiento se ve apoyado por los datos registrados por Cataldo y Oakhill (2000). Los investigadores pidieron a un grupo de estudiantes que leyese varios relatos para responder a unas preguntas más tarde. Se distinguió entre buenos y malos en función de sus puntuaciones en una prueba estandarizada de comprensión lectora (Neale Analysis of Reading Ability). Una vez contestadas a las preguntas sobre los relatos, se les volvieron a plantear las mismas preguntas pasado un rato. En esa ocasión, los participantes tenían que buscar la respuesta en el texto; sus patrones de búsqueda fueron registrados. Los malos aprendices usaron con mucha más frecuencia la estrategia de empezar a buscar por el principio; los buenos, por su parte, se sirvieron más frecuentemente de la estrategia de ir directamente al punto adecuado del texto. Así, una vez identificado el problema (o pregunta, en este caso) a resolver, los buenos se las ingenian mejor que los malos para hacerlo. Y eso que la memoria espacial del texto de buenos y malos se mantuvo bajo control: en una prueba que exigía localizar ciertas palabras de los textos en unas copias codificadas de los mismos (“había una vez” se cambiaba por “----- --- ---”), buenos y malos rindieron igual. Datos como los de Hyönä et al. o Cataldo y Oakhill sugieren que, una vez detectados los problemas, los malos siguen siendo peores a la hora de desplegar comportamientos útiles para repararlos.

También hay estudios en los que se exploran los procesos de detección y reparación simultáneamente. En los trabajos que vamos a revisar enseguida el método es sencillo. Se les proporciona a los participantes un material instructivo del cual deben aprender. Durante la fase de estudio del material se les demanda expresar en voz alta los pensamientos que acuden a su mente. De esa manera, se espera, los participantes hacen visibles sus procesos mentales y representaciones provisionales. Pasada la fase de estudio, los participantes resuelven tareas que miden el aprendizaje alcanzado en la fase anterior. Analizando los procesos revelados en voz alta cabe establecer relaciones entre ellos y el aprendizaje alcanzado. Esto es lo que hicieron Chi et al. (1989). Una muestra de participantes hubo de aprender física de Newton a partir de un libro y unos problemas resueltos. Por supuesto, mientras estudiaban los materiales los participantes verbalizaron todo lo que acudía a su mente. Tras la fase de estudio resolvieron problemas similares a los que habían estudiado. Sobre la base de su puntuación en estos problemas se diferenció entre buenos y malos aprendices. Los análisis de sus verbalizaciones durante el estudio pusieron en evidencia que los buenos generaban más

detecciones y reparaciones. Concretamente, la media de comentarios de supervisión (e.g., “no entiendo qué significa X”) por parte de los buenos fue significativamente superior a la de los malos. Además, el porcentaje de comentarios de este tipo que fue seguido de explicaciones que intentaban despejar el problema identificado fue significativamente menor en el caso de los malos. Azevedo, Guthrie y Seibert (2004) llevaron a cabo un estudio semejante, en su caso con un CBLE sobre el sistema circulatorio humano. Los participantes fueron asignados al grupo de malos o buenos en función del tamaño de su “salto” de modelo mental. Es decir, los modelos mentales previos de los participantes sobre el funcionamiento del sistema circulatorio fueron analizados. También los que construyeron tras consultar el CBLE sobre el tópico en cuestión. Dada una taxonomía de modelos del sistema circulatorio que distingue niveles, los cambios o saltos de modelo pudieron calcularse. Durante la fase de estudio a partir del CBLE se instó a los participantes a revelar su pensamiento. Los análisis de las verbalizaciones permitieron comprobar, una vez más, que los buenos detectan y reparan con mayor frecuencia. Las detecciones manifestadas por los buenos fueron el doble que las de los malos. Los buenos, además, pasaron más tiempo aplicando estrategias de reparación efectivas (e.g., releer el segmento oportuno, buscar la fuente de información pertinente) que inefectivas; los malos, en cambio, dedicaron el mismo tiempo a unas y otras. Pirolli y Recker (1994) también implementaron un estudio de estas características. Aquí los participantes debían aprender programación en LISP a partir de un libro de texto (haciendo explícito su pensamiento, claro está) y varios ejercicios de programación resueltos. Tras la fase de estudio siguió una fase de evaluación en la que los participantes resolvieron ejercicios de programación. Con arreglo al número de fallos en estos ejercicios, Pirolli y Recker diferenciaron buenos y malos aprendices. De nuevo, la frecuencia de comentarios de supervisión (o detecciones) fue mucho mayor en los protocolos de los buenos aprendices. Eso sí, la frecuencia de explicaciones reparadoras subsiguientes a las detecciones no fue diferente entre buenos y malos; lo que sí les diferenció fue la frecuencia de producción de explicaciones adicionales. Adicionales porque no respondían a la necesidad impuesta por la detección de un error sino por el afán de profundización: los buenos iban más allá en sus reparaciones que los malos. Sintetizando, cuando registramos los pensamientos que se producen durante la consulta de materiales instructivos descubrimos que los buenos aprendices identifican problemas de comprensión y despliegan medios para resolverlos con más frecuencia que los malos.

En este subapartado hemos repasado estudios que distinguen buenos y malos aprendices y evalúan sus procesos de detección y reparación. En unos casos se plantean tareas que exigen detectar o reparar malos entendidos y en otros se deja a los aprendices consultar libremente un material instructivo registrando su modo de abordar la tarea. Sea cual sea el método empleado, buenos y malos muestran diferencias significativas en detección y reparación a favor de los primeros. En tanto que es una de las diferencias entre los buenos y los malos aprendices, detectar y reparar puede considerarse clave.

3.2.2. Estudios de regresión. Para estimar el peso de un proceso sobre el aprendizaje es posible también llevar a cabo análisis de regresión. Lo que se necesita es tomar datos de un número suficiente de participantes midiendo su nivel en un repertorio de procesos. En ese repertorio habrá procesos considerados básicos, los cuales habrán de predecir razonablemente bien el comportamiento de otra variable criterio (el aprendizaje), y otros

menos básicos, que son los que uno somete a análisis para determinar si tienen o no poder predictivo por encima de los básicos.

Podría utilizarse la comprensión lectora como variable criterio, por ejemplo. Las variables que habrían de predecir el comportamiento de esta variable razonablemente bien serían la capacidad para leer palabras (conocidas y desconocidas), el vocabulario del que uno dispone, la capacidad para realizar inferencias o la capacidad de la memoria de trabajo. Todas esas deberían predecir o explicar una porción de la varianza en comprensión lectora considerable. Después de comprobado eso, uno puede averiguar si otra variable puede aumentar significativamente la cantidad de varianza explicada. Este método es el que Cain, Oakhill y Bryant (2004) siguieron. Los investigadores reclutaron a más de cien niños y les aplicaron un set de pruebas evaluando todos los procesos mencionados y más. Una de las pruebas evaluaba la capacidad para detectar fallos en la comprensión. Cain et al. presentaban textos con inconsistencias (como los de Rubman y Waters de hace un momento) a los niños. Se esperaba que los niños notasen que algo no tenía sentido en los textos. Si lo notaban y hacían esto en todas las ocasiones oportunas, obtenían la máxima puntuación. Lo que hicieron Cain et al. fue introducir las variables ya mencionadas en el análisis de regresión para después hacer lo propio con la puntuación en detección y observar si ésta aportaba poder predictivo. Así fue. Esto significa que una parte de la comprensión lectora depende *exclusivamente* de la capacidad de los aprendices para notar que algo va mal.

Maña, Vidal-Abarca, Domínguez, Gil y Cerdán (2008) hicieron algo parecido, esta vez incluyendo una medida de reparación. Los participantes debían resolver un conjunto de pruebas. Para empezar, debían leer un texto y responder a unas preguntas sobre el mismo. Su rendimiento en esta prueba reflejaba su capacidad para llevar a cabo las operaciones básicas implicadas en comprensión (e.g., acceso léxico, análisis sintáctico, inferencias puente). Después, debían afrontar una tarea algo más compleja. Debían utilizar un CBLE que incluía varios textos y un set de problemas a resolver empleando los textos. El CBLE registraba su comportamiento (tiempo dedicado a leer cierto segmento, secuencia de segmentos consultados, etc.). Los investigadores consideraron la identificación de oraciones inconsistentes que habían sido deliberadamente insertadas en el material como medida de detección. El uso de información pertinente para resolver los problemas planteados el CBLE sirvió como medida de la reparación. Una vez los participantes sabían qué problema habían de resolver podían emplear los materiales contenidos en el CBLE a su antojo. Un aprendiz estratégico dedicará más tiempo a aquellos segmentos que le ayudan a reparar el problema y tratará de resolver el problema inmediatamente después de haber obtenido la información que necesita para ello. Esos dos indicadores son los que se tomaron aquí como medida de reparación. Terminada la consulta del CBLE, los participantes respondieron a preguntas que revelaban cuánto habían aprendido de los tópicos tratados en el CBLE (vacunas contra la gripe y calzado deportivo). La puntuación en estas preguntas servía como variable criterio; variable que las otras deberían predecir. Las operaciones básicas en comprensión explicaban una buena parte de la varianza (más del 30%), pero lo interesante es que las medidas de detección y reparación contribuían a explicar la varianza por encima de las operaciones básicas. De modo que tanto la detección como la reparación son responsables, por encima de otras variables, de un parte significativa del aprendizaje.

En virtud de lo expuesto cabe decir que tanto detectar como reparar representan una porción del aprendizaje. Y lo hacen con derecho propio: explican una parte de la varianza en aprendizaje *por encima* de otras variables básicas.

3.2.3. Estudios de entrenamiento. Hay un último conjunto de evidencias al que vamos a recurrir para demostrar la importancia de los procesos de detección y reparación en el aprendizaje. Se trata de los estudios de entrenamiento. En ellos los participantes son instruidos en los procesos clave y su capacidad para aprender a partir de materiales instructivos es evaluada tras el periodo de entrenamiento. Por lógica, si los aprendices entrenados aprenden mejor que otros aprendices no entrenados pero semejantes en otras variables, tal ventaja debe atribuirse al efecto de los procesos entrenados (y a la efectividad del entrenamiento).

Disponemos de evidencias que indican que propiciar el uso de la detección y la reparación a través de entrenamiento repercute en el aprendizaje. Bielaczyc, Pirolli y Brown (1995) entrenaron a un grupo de estudiantes en varias habilidades para aprender más profundamente de materiales instructivos. Entre ellas se encontraban las de (a) detectar problemas de comprensión y de (b) abordar los problemas identificados. En las sesiones de entrenamiento los participantes observaron a un modelo llevar a cabo los procesos clave (enfrentándose a problemas hipotéticos: “no entiendo esto, uhmmm, ¿qué puedo hacer?; ahá, volveré atrás...”) y, más tarde, ellos mismos practicaron lo que habían visto demostrar al modelo. Al tiempo que estos participantes se dedicaban a estas tareas, otros (el grupo control) leían un manual sobre programación LISP (el modelo del grupo entrenado empleaba el mismo manual en sus demostraciones). Pasada la fase de entrenamiento todos, grupo entrenado y grupo control, debían estudiar unos ejercicios de programación LISP a partir de un libro de texto para resolver unos problemas más tarde. Los participantes pensaron en voz alta (manifestaron lo que acudía a su mente, como antes explicamos) durante la fase de consulta del libro de texto. Los análisis de los protocolos de pensamiento en voz alta revelaron que quienes habían disfrutado del entrenamiento emitían un número significativamente mayor de detecciones y reparaciones. Eso sirvió como indicación de que el programa de entrenamiento había logrado el efecto deseado. Al analizar el rendimiento de unos y otros participantes Bielaczyc et al. observaron que los entrenados incurrían en un número sensiblemente menor de errores al resolver los ejercicios de programación en comparación con los no entrenados. Esto significa que quienes fueron inducidos a detectar y reparar más, aprendieron más.

Por lo que se refiere al componente de detección en solitario hay también evidencias de su impacto sobre el aprendizaje, una vez inducido el proceso vía entrenamiento. McNamara (2004) desarrolló y puso a prueba un entrenamiento para aprender mejor de los textos. El entrenamiento instruía en varias habilidades entre las que se encontraba la de detectar problemas de comprensión. En su estudio participaron estudiantes que fueron repartidos al azar entre dos condiciones: entrenamiento y control. Los del entrenamiento asistieron a una sesión en la cual se les describían y ejemplificaban las habilidades a incorporar, se les ofrecía una demostración de un modelo en el uso de tales habilidades y se les instaba a practicar lo que habían observado. Los del grupo control leían en voz alta los mismos textos que el modelo del grupo entrenamiento había utilizado en sus demostraciones. En una sesión posterior todos los participantes debían leer un nuevo texto (sobre medio ambiente) mientras revelaban su pensamiento concomitante para resolver unas preguntas

después. Los análisis de los protocolos de pensamiento mostraron que los entrenados supervisaban más su comprensión en comparación con los no entrenados. El rendimiento de los entrenados en las preguntas sobre el texto estudiado fue significativamente superior al de los participantes en el grupo control. Un par de años más tarde, McNamara implementó un estudio similar (McNamara, O'Reilly, Best & Ozuru, 2006), esta vez con un programa de entrenamiento informatizado. Los resultados son un calco de los que encontró en 2004. Azevedo y Cromley (2004) llevaron a cabo un estudio de entrenamiento también. En su caso, los participantes no leían textos sino que utilizaban un CBLE sobre el sistema circulatorio humano. Una vez más hubo dos condiciones, entrenamiento y control, y dos fases, de intervención y de evaluación. Una vez más, también, los resultados revelaron que los participantes entrenados supervisaban más su aprendizaje y lograban aprender más (aprendizaje, en esta ocasión, medido como magnitud del salto del modelo mental previo y posterior al estudio). Tomados en conjunto, en estos estudios se hace patente que se puede entrenar a los aprendices para supervisar mejor su comprensión (detectando más frecuentemente los escollos) y que tal cambio en su modo de obrar repercute en el aprendizaje. Es importante señalar que en todos ellos los participantes entrenados aumentaron significativamente también el número de elaboraciones que realizaban durante el estudio de los materiales instructivos. Ello podría ser un indicador de que no sólo detectaban más sino que reparaban más (mediante elaboraciones de los contenidos del material). El problema es que no se distinguió en los análisis entre elaboraciones surgidas espontáneamente y las surgidas como solución a un problema detectado.

En conclusión, puede instruirse a los aprendices para que detecten y reparen más sus problemas de comprensión y ello tiene un impacto significativo sobre el aprendizaje. De nuevo, queda constatado que los procesos de detección y reparación juegan un rol determinante en el aprendizaje a partir de CBLEs y textos.

3.2.4. Resumen. En la tabla figura un resumen de las evidencias consultadas y que revelan que detectar y reparar son procesos clave.

Tabla 18. Indicadores de la importancia de detectar y reparar.

Fuente	Evidencias
Estudios comparación	Los buenos aprendices detectan y reparan (e.g., Chi et al., 1989)
Estudios regresión	Los procesos de detección y reparación explican varianza en aprendizaje por encima de otras variables (e.g., Mañá et al., 2008)
Estudios entrenamiento	Entrenar a los aprendices en detección y reparación redundante en un mayor aprendizaje (e.g., Bielaczyc et al., 1995)

3.3. Un nuevo problema: los aprendices tienen dificultades para detectar y reparar.

Acabamos de pasar revista a un buen número de trabajos en los que se pone en

evidencia que detectar y reparar son procesos clave en el aprendizaje. Lo que sabemos ahora es que si quieres aprender debes detectar y reparar. Hasta aquí no hay problema. La cosa se complica cuando atendemos a los siguientes datos: los aprendices suelen tener dificultades al detectar malentendidos y repararlos no les resulta más fácil. Desafortunadamente está bien documentado que los aprendices ni suelen detectar ni suelen reparar, como comprobaremos a continuación. Concretamente, vamos a ver que hay un alto porcentaje de aprendices incapaz de percatarse de la existencia de incongruencias en ciertos materiales y que, en general, demuestran poca habilidad para calibrar cuán bien han aprendido. Vamos a ver también que aún cuando los problemas son detectados, los aprendices generan pobres o incluso nulas explicaciones para repararlos; también que hacen mal uso de material útil para la reparación.

3.3.1. Los aprendices no detectan. Para explorar la capacidad de los aprendices de detectar malentendidos es posible recurrir a dos métodos alternativos. Ambos han sido descritos someramente ya; ahora los explicamos más detenidamente.

Un método consiste en introducir deliberadamente inconsistencias en el material y ver si los aprendices las advierten. Tales inconsistencias pueden ser internas o externas. Las internas son contradicciones entre dos afirmaciones del propio material: “la superconductividad se obtiene enfriando ciertos materiales... la superconductividad se obtiene aumentando la temperatura de ciertos materiales”. Las externas son contradicciones entre lo que es comúnmente sabido y lo que dice el material: “los peces saltaban por encima del agua en el lago helado”. Se trata de que los aprendices informen de la existencia de tales problemas una vez consultado el material. Se considera ésta una medida de detección porque no cabe construir una representación mental coherente de un material que no lo es: si uno examina su propio entendimiento del material debe concluir, si se ha examinado bien, que hay una disonancia en su representación.

El otro método consiste en solicitar a los aprendices que estimen o calibren cuánto han aprendido (cf. Lin & Zabrocky (1998) para una revisión). Más específicamente, se trata de que los aprendices predigan su rendimiento posterior en una tarea que demandará demostrar el aprendizaje que han alcanzado a partir del estudio de cierto material. O sea, se pide a los aprendices que estudien cierto material y se les invita a estimar cuánto han aprendido, lo que les induce a examinar sus propias representaciones mentales. Más tarde se compara el aprendizaje estimado con el demostrado. La correlación entre ambas cantidades indica cuán bien los aprendices examinan sus propias representaciones mentales para evaluar su completud y corrección. Si, pongamos, un aprendiz ha construido una representación que presenta distorsiones o simplificaciones y, sin embargo, predice que rendirá bien en la tarea posterior, entonces cabe deducir que no ha sido capaz de identificar las limitaciones de su modelo mental.

¿Qué sucede cuando se pide a los aprendices que intenten aprender de materiales con inconsistencias? Pues que éstas pasan inadvertidas para buena parte de los aprendices. Otero y Campanario (1990) reunieron a un grupo de estudiantes de Secundaria (16 y 18 años) para que evaluaran la inteligibilidad de varios textos. Los textos consistían en párrafos de unas seis oraciones cada uno. Del total de seis textos, más de la mitad contenían contradicciones internas como las descritas antes (e.g., “la superconductividad se obtiene enfriando ciertos

materiales... la superconductividad se obtiene aumentando la temperatura de ciertos materiales”). Los resultados de los participantes de 16 y 18 años se consideraron independientemente. Sorprendentemente, el 34% de los estudiantes de 16 años identificó cero inconsistencias, siendo el porcentaje de los que identificaron sólo una del 25. Esto es, el 59% de los estudiantes de 16 rindió realmente mal en la tarea de detectar contradicciones. Los resultados de los estudiantes de 18 años son algo más halagüeños, aunque no todo lo buenos que uno esperaría. El 43% de ellos fue incapaz de detectar todas las contradicciones presentadas, dejando alguna en el tintero. Otero y Kintsch (1992) repitieron el estudio un par de años después. ¿Con qué resultado? Lo cierto es que fue algo peor, si cabe. Esta vez, el 70% de los estudiantes de 16 años encontró cero contradicciones; el 24% de los de 18 años encontraron cero contradicciones. Un repaso a resultados como éstos permite extraer una conclusión: un alto porcentaje de los aprendices rinde pobremente al tratar de identificar problemas en materiales en los que se han insertado inconsistencias.

¿Qué ocurre cuando se pide a los aprendices que calibren cuánto han aprendido? Tal como veremos enseguida, lo que sucede es que no hay relación entre las estimaciones de los aprendices y su nivel real. Maki, Foley, Kajer, Thompson y Willer (1990) solicitaron a un grupo de participantes que leyese varios textos sobre medicina para responder a unas preguntas más tarde. Antes de responderlas, los aprendices calibraron (en una escala 1-6) lo bien que habían aprendido a partir de los textos. Maki et al. calcularon la correlación entre las estimaciones y los rendimientos reales exhibidos en las preguntas. En el primer estudio la correlación fue de aproximadamente .12; en el segundo estudio de .14. Correlaciones, ambas, bastante bajas. Schommer y Surber (1986) llevaron a término un estudio similar. Sus participantes estudiaron textos de psicología y juzgaron su confianza (en escala 1-4) en el aprendizaje que habían ganado a partir de ellos. Después resolvieron cuatro preguntas que exigían poner en juego lo realmente aprendido. Los investigadores consideraron que allí donde un participante había mostrado una alta confianza (3-4) pero había contestado correctamente a pocas preguntas (1-2) había incurrido en una *illusion of knowing* o creencia errónea de saber. Globalmente, el porcentaje de *illusion of knowing* fue del 25%: uno de cada cuatro, lo que no es para nada desdeñable. Ahora bien, si consideramos sólo los resultados en una condición en la que se pidió a los participantes que leyesen el texto “para saber si era legible” (en lugar de “para dar una lección sobre el tema a unos compañeros”) y los correspondientes al texto difícil (porque había uno fácil y otro difícil), el porcentaje de ilusos fue del 51. Más recientemente, Commander y Stanwyck (1997) emplearon el método con los mismos fines. Sus textos fueron también sobre psicología pero la tarea posterior fue de recuerdo (en lugar de contestación a preguntas). Según los investigadores un 42% de los participantes mostró mala precisión calibrando su aprendizaje, en tanto que hubo gran divergencia entre sus predicciones y su rendimiento real. De nuevo, la conclusión es clara: un alto porcentaje de aprendices calibran con poca precisión su aprendizaje a juzgar por la discrepancia entre sus predicciones y su rendimiento efectivo.

En resumidas cuentas, distintos métodos coinciden al reflejar un mismo panorama: los aprendices encuentran difícil el examinar sus propias representaciones mentales. Esto, claro está, es un serio obstáculo para identificar distorsiones o simplificaciones en su comprensión de un material instructivo.

3.3.2. Los aprendices no reparan. A la hora de explorar la capacidad de los aprendices para reparar los problemas de entendimiento identificados cabe servirse de al menos dos métodos alternativos. ¿Cuáles? Sigamos leyendo y encontraremos una descripción de ambos.

Uno consiste en solicitar a los aprendices que, al tiempo que aprenden de un material, expresen todo aquello que pase por su mente. Un método es, pues, el que se sirve de los protocolos de pensamiento en voz alta de los que ya hemos hablado. Al expresar lo que acude a su mente mientras utilizan los materiales, los aprendices nos brindan una ventana a sus procesos mentales (no sin un arduo trabajo de codificación y análisis previo, pero ventana al fin y al cabo). Dicha ventana puede permitirnos apreciar cómo los aprendices identifican problemas en sus representaciones mentales (e.g., “no termino de ver claro eso de una placa se hunda...”) y, más interesante aún, cómo generan explicaciones para repararlos (e.g., “ahá, si la placa oceánica se hunde es porque es más pesada que la continental”). Obviamente, sólo si consideramos la producción verbal que precede a la reparación podemos calificarla como tal. Una elaboración puede surgir como resultado de un intento de enriquecer lo que uno está leyendo o como reacción ante la identificación de una incompreensión. Por tanto hay que estar atento al contexto de producción de las elaboraciones: las que siguen a una detección pueden considerarse reparaciones.

Otro método es plantear a los aprendices problemas (algo así como forzar una detección) y observar cómo los resuelven. Por ejemplo, uno puede diseñar un CBLE que contiene textos y problemas a resolver y registrar de algún modo cómo los aprendices se sirven de los textos para resolver los problemas planteados. Definido el problema y facilitados los bloques para construir la respuesta al mismo, si los aprendices no saben cómo armar los bloques para componer la solución, entonces cabe deducir que no reparan adecuadamente.

¿Qué pasa cuando les pedimos a los aprendices pensar en voz alta mientras consultan materiales? A veces lo que ocurre es que tras haber detectado malentendidos los aprendices no hacen nada para resolverlos. Coté et al. (1998), como ya explicamos, hicieron que varios estudiantes de Primaria estudiaran textos de geografía mientras pensaban en voz alta. Al analizar los protocolos de pensamiento Coté y sus colegas pudieron apreciar que los participantes supervisaban su comprensión. La mayoría de sus comentarios eran del tipo “entiendo bien esto” o “esto me resulta fácil” pero hubo también unos cuantos de tipo “no logro entender X”. La cuestión es como sortearon estos escollos. Coté et al. nos informan de que el 50% de sus participantes emplearon al menos una vez la estrategia de ignorar el problema. Hacer caso omiso al problema fue una estrategia con tanto éxito como releer pasajes anteriores. Kintsch y Kintsch (1995) solicitaron a una muestra de estudiantes de Secundaria que aprendiese sobre el sistema circulatorio humano mediante un texto. Al hacerlo pensaron en voz alta. Un dato que llamó nuestra atención es que Kintsch y Kintsch informan de una correlación negativa entre los comentarios de supervisión (detecciones) y las elaboraciones. A más detección, menos elaboración. Si para reparar un problema detectado es necesario generar elaboraciones que permitan resolverlo y aquí el aumento de lo uno está asociado al descenso de lo otro, cabe aventurar que no hubo reparaciones siguiendo a las detecciones. O sea, de nuevo la solución a los problemas es no hacer nada.

Otras veces lo que ocurre es que las explicaciones que los aprendices generan no sirven para reparar los malentendidos que han identificado. Otero y Campanario (1990), de los que ya hemos hablado, aportan datos en este sentido. Resulta que además de pedir a sus participantes que evaluaran si los textos facilitados eran o no inteligibles, los investigadores entrevistaron a todos los que juzgaron alguno de los textos contradictorios como inteligible o que fueron incapaces de justificar su juicio. En la entrevista pudieron comprobar qué les motivo a emitir su juicio. Así, Otero y Campanario identificaron malas reparaciones. En unos casos (en torno al 4%) los aprendices no hacían absolutamente nada con el problema identificado. Éstos habían juzgado el texto como ininteligible pero no sabían decir por qué, dónde residía el problema, o cómo resolverlo. Éstos recurrían a la táctica de los niños de Coté y compañía: ignorar el problema. Pero Otero y Campanario describen otro patrón más interesante aún. Otros tantos (4% aproximadamente) habían identificado la contradicción y habían resuelto hacer algo para despejarla: habían inventado una justificación para hacer sostenible la aparente contradicción. Por ejemplo, si el texto decía “la superconductividad se obtiene enfriando ciertos materiales... la superconductividad se obtiene aumentando la temperatura de ciertos materiales” los participantes razonaban que “antes se obtenía con frío pero ahora se obtiene con calor, gracias al avance de la ciencia”. Renkl (1997) puso a sus participantes a aprender cálculo de probabilidades a partir de ejemplos resueltos para más tarde resolver unos problemas de probabilidad. Instó a los participantes a pensar en voz alta mientras lo hacían, lo que debería inducirles a explicar el razonamiento subyacente a los ejercicios resueltos. En efecto, los participantes generaron un buen número de explicaciones para los ejercicios resueltos. Renkl distinguió varias categorías entre las explicaciones: había explicaciones basadas en los principios de probabilidad detrás del ejercicio (e.g., “los eventos son independientes y por eso se aplica la multiplicación”), había elaboraciones de la situación en torno al problema (“si ha sacado un melón de la cesta, entonces el número de melones en la cesta se ha reducido en uno”), imposiciones de metas para ciertas acciones (“ha multiplicado para saber la probabilidad de encontrar un melón pequeño”), etc. Aunque las explicaciones no surgieron necesariamente como respuesta a la detección de un problema (más bien surgieron de la necesidad impuesta de buscar las razones detrás de las cosas) es interesante saber que no todas las explicaciones condujeron a un mayor aprendizaje. Así, elaborar la situación del problema no estuvo correlacionado con el aprendizaje. En un análisis de regresión Renkl descubrió que imponer metas para las acciones no explicaba una porción significativa de varianza por encima de la habilidad matemática y el conocimiento previo. De modo que no toda explicación sirve: no sirve inventar un razonamiento para sostener una contradicción ni sirve que las explicaciones de uno se concentren en determinados aspectos; sin embargo, eso es lo que hacen muchos aprendices.

¿Qué se observa cuando registramos el comportamiento de los aprendices al resolver problemas? Pues que su comportamiento dista de ser eficiente, como ahora ilustraremos. Renkl (2002), de nuevo, solicitó a una muestra de participantes que aprendiese a calcular probabilidad, esta vez a partir de un CBLE. El CBLE contenía ejercicios resueltos y la mitad de los participantes podía solicitar explicaciones adicionales (revelando el razonamiento detrás del ejercicio) a voluntad. Tras utilizar el CBLE todos tuvieron que resolver varios ejercicios más o menos parecidos a los consultados. En general, quienes disfrutaron de la posibilidad de solicitar explicaciones adicionales rindieron mejor que sus compañeros. Ahora, un examen detenido del uso de estas explicaciones pone de relieve que muchos no las

emplearon con frecuencia o lo hicieron inadecuadamente. Justamente quienes rindieron peor dentro del grupo explicación-adicional-a-voluntad fueron quienes menos se apoyaron en las explicaciones adicionales. Esto se puede interpretar alegando que no fueron capaces de detectar lagunas en su comprensión y por eso pensaron no necesitar apoyo adicional. O sea, como un fallo en la detección. Pero dentro de los que las utilizaron hubo también quienes las emplearon mal. Muchos de los que las emplearon, dedicaron un tiempo mínimo a revisarlas y ello les condujo a un mal rendimiento. Cataldo y Oakhill (2000), cuyo estudio ya hemos comentado, nos proporcionan datos en esta línea también. Recordemos que sus participantes debían consultar unos textos y resolver problemas apoyándose en ellos. Vimos que muchos participantes leían desde el principio los textos para resolver los problemas planteados en lugar de acudir al punto donde se encontraba la información pertinente. Resultados como los de Renkl o Cataldo y Oakhill dibujan un panorama en el que los aprendices, aún contando con las piezas para reparar problemas, son poco capaces de operar con ellos para conformar una solución.

Resumiendo, aunque los problemas sean identificados, los aprendices no son siempre capaces de repararlos. A veces lo que ocurre es que no hacen nada para resolverlos; otras, que sus reparaciones no son buenas; otras, que no saben cómo utilizar las piezas para armar el puzzle que los saque del atolladero.

3.3.3. Resumen. En la tabla que adjuntamos a renglón seguido figura un resumen de cuanto hemos revisado en este apartado. En concreto, lo difícil que es para muchos aprendices detectar y reparar.

Tabla 19. Indicadores de que los aprendices tienen dificultades para detectar y reparar.

Problema	Tipo de evidencias
Los aprendices no detectan	<p>Cuando se insertan inconsistencias en los textos muchos aprendices son incapaces de advertir su presencia (e.g., Otero & Kintsch, 1992)</p> <p>Cuando se pide a los aprendices calibrar cuánto han aprendido la correlación con el aprendizaje real es baja (e.g., Maki et al., 1990)</p>
Los aprendices no reparan	<p>Aunque advierten la presencia de problemas emprenden malas medidas para resolverlos (e.g., Otero & Campanario, 1990)</p> <p>Aunque se les brinda la posibilidad de apoyarse en materiales para reparar los problemas, no saben bien cómo emplearlos (e.g., Cataldo & Oakhill, 2000)</p>

3.4. Ayudando a detectar y reparar.

Hasta el momento hemos podido constatar tres cosas. Una, aprender de los CBLEs plantea una paradoja (grandes exigencias de procesamiento para un sistema limitado) que hace probable la construcción de representaciones mentales pobres. Dos, que para soslayar la paradoja (es decir, para aprender profundamente) detectar y reparar son procesos clave. Tres,

que los aprendices experimentan dificultades a la hora de detectar y reparar. A tenor de tales circunstancias sólo cabe proponer asistir a los aprendices en los procesos clave: hemos de brindar ayudas para detectar y reparar. Por suerte, muchos investigadores han ideado formas de hacer esto. No sólo eso: sus ayudas son eficaces, a juzgar por los resultados experimentales hasta el momento. Lo que vamos a hacer en este apartado es precisamente repasar algunas de esas ayudas y revisar el apoyo empírico que han recibido.

3.4.1. Ayudas para detectar. Existen diversas ayudas para detectar. Se trata de dispositivos que asisten a los aprendices en la identificación de problemas (lagunas, simplificaciones, distorsiones) en sus representaciones mentales. Cabe distinguir niveles de apoyo a la detección; nosotros vamos a presentar unas cuantas ayudas ordenadas en función de su grado de apoyo: de menos a más.

Primero, algunos de estos dispositivos piden a los aprendices que se aseguren de estar comprendiendo a fondo el material pero sin ofrecer más pistas. Un ejemplo son los *open prompts* o claves abiertas. Por ejemplo, Schworm y Renkl (2006) hicieron que sus participantes aprendieran a diseñar materiales instructivos a partir de un CBLE que incluía *open prompts*. El CBLE describía distintos materiales instructivos, más o menos apropiados para aprendices hipotéticos. Los prompts instaban a los participantes a pensar por qué un material era mejor o peor para un aprendiz hipotético dado (e.g., “¿En cuál de estos es más fácil para el estudiante X hacer lo que se le pide? ¿Por qué?”). Es decir, los prompts actuaban como mensajes genéricos del tipo “¿estás seguro de estar comprendiendo bien?”. Eso puede ayudar a los aprendices a ser más escrupulosos y, por tanto, a detectar con más probabilidad posibles malentendidos.

Segundo, existen también los *error messages* o mensajes error. Éstos les hacen saber a los aprendices que hay algo erróneo en sus representaciones mentales, especificando o no qué aspecto concreto es el defectuoso. Si con los *open prompts* invitábamos al aprendiz a examinar sus representaciones en busca de posibles problemas, ahora le confesamos que, de hecho, hay un problema. Por ejemplo, Alevén y Koedinger (2002) introdujeron un tutor virtual en un CBLE que enseñaba geometría. El tutor solicitaba a los aprendices argumentar los pasos que seguían al resolver los ejercicios planteados por el CBLE y ofrecía feedback sobre dichas argumentaciones en forma de *error messages* (e.g., “Deberías ser más preciso”, “Mejor que elimines eso de tu argumentación”). O sea, Alevén y Koedinger idearon la forma de saber qué tenían sus aprendices en la cabeza para indicarles cuándo algo de eso estaba mal.

Tercero, las ayudas a la detección pueden ser también un feedback correctivo. Éste lo que hace es identificar qué aspectos son erróneos en las representaciones de los aprendices argumentando por qué son tal cosa. Ahora no sólo se indica qué aspecto concreto está mal sino las razones para sostener tal apreciación. Por ejemplo, Moreno (2004) pidió a sus participantes que diseñaran una planta, la cual debía sobrevivir a distintas condiciones ambientales. Los participantes veían una presentación multimedia que describía los diferentes ambientes y después elegían, de entre varias opciones, las raíces, tallos y hojas que creían apropiados para la planta (dadas unas condiciones ambientales). Tras hacerlo recibían feedback correctivo sobre sus elecciones (e.g., “No, un tallo corto es peligroso aquí, puesto que las hojas no podrán obtener luz solar”).

Por último, otro tipo de ayuda a la detección revela a los aprendices no sólo qué aspecto es erróneo en sus representaciones y por qué cabe considerarlo así sino que describe en detalle la representación mental que el aprendiz alberga en su mente. Si antes se decía que estaba mal y por qué, ahora lo que hacemos es dedicar un momento a reflejar con profundidad qué es lo que el aprendiz tiene en su mente. De este modo, una vez argumentado por qué *todo eso* que el aprendiz tiene en su mente está mal, no debe quedar resquicio de duda acerca de su incorrección. Este dispositivo se ha dado en llamar *impasse-trigger* por Sánchez, García-Rodicio y Acuña (2008), ya que pretende suscitar en el aprendiz la conciencia de que su comprensión no es adecuada (*impasse*), lo que debe hacerle reaccionar. Estos investigadores solicitaron a sus participantes aprender geología a partir de un CBLE. El CBLE incluía *impasse-triggers* bien como feedback sobre las respuestas de los participantes a preguntas incrustadas en el CBLE (e.g., “Dices que el magma de los volcanes compensa la pérdida de corteza en la subducción porque al aflorar se solidifica creando nueva corteza pero no es correcto: la cantidad de nueva corteza nunca podría compensar la pérdida de la subducción porque la actividad volcánica es irregular y la cantidad de magma solidificado escasa”) o como mensajes de advertencia en momentos críticos en los que los aprendices típicamente distorsionan los conceptos (e.g., “Seguramente has pensado que los choques de los Andes y el Himalaya son muy parecidos; al fin y al cabo, en ambos chocan dos placas y de en ambos nacen montañas. Pero en realidad presentan diferencias cruciales...”).

En suma, todos estos dispositivos pueden considerarse ayudas al proceso de detección. Pueden considerarse así porque todos ofrecen apoyo a los aprendices a la hora de identificar problemas en sus representaciones mentales. Eso sí, difieren en el grado de apoyo que brindan al hacerlo: los *open prompts* son los que menos apoyan, los *impasse-triggers* los que más.

3.4.2. Ayudas para reparar. Es posible identificar ayudas al proceso de reparación. Se trata de dispositivos que ofrecen los bloques que servirán para construir la explicación que repare los malentendidos. De nuevo, cabe distinguir grados de apoyo en estas ayudas. De nuevo, también, serán presentadas en orden de acuerdo con su nivel de apoyo: de menos a más.

Primero, las ayudas a la reparación pueden ser mensajes breves que revelan la *respuesta correcta* a preguntas cerradas. Atkinson, Renkl y Merrill (2003) solicitaron a sus participantes aprender cálculo de probabilidades a partir de un CBLE que contenía ejercicios resueltos. Una vez estudiado un ejercicio, los participantes debían especificar qué principio subyacía a cada paso en la resolución del ejercicio. Para ello seleccionaban posibles principios de un menú. Cuando fallaban, el CBLE mostraba la respuesta correcta (e.g., “Principio aditivo”).

Segundo, también pueden ofrecerse pistas o *hints* como bloques para construir representaciones mentales completas y correctas. El tutor desarrollado por Alevén y Koedinger (2002), al que ya nos hemos referido, incorporó también este tipo de ayuda. Para garantizar que los participantes ofrecían argumentaciones adecuadas a los pasos que iban siguiendo en la resolución de los ejercicios, el tutor les brindaba *hints* (e.g., “Cuando dos líneas paralelas son interseccionadas por una transversal, los ángulos interiores son iguales...”).

”). Así, si las representaciones mentales de los participantes presentaban algún problema (e.g., laguna, distorsión), las pistas podían resolverlo (completando o reemplazando).

Tercero, las hints pueden ser algo más generosas si las ofrecemos, por ejemplo, en forma de *fill-in-the-blank* o casillas para rellenar. Si antes se ofrecían pistas aisladas, ahora se facilita un paquete casi completo. Berthold, Eysink y Renkl (2008) hicieron a sus participantes aprender cálculo de probabilidades utilizando un CBLE. Se instó a los participantes a explicar el razonamiento subyacente a los pasos en la resolución de ejercicios resueltos. Para asegurarse de que los participantes daban buenas argumentaciones, los investigadores facilitaron fill-in-the-blanks (e.g., “Que suceda X es ___ veces tan probable como que suceda Y”).

Por último, podemos proporcionar a los aprendices elaboraciones complejas basadas en los contenidos del material. Moreno (2004) facilitó a sus participantes no sólo el feedback correctivo antes mencionado sino también *feedback explicativo*. Dicho feedback incorporaba elaboraciones sobre lo que el material había expuesto (e.g., “En este ambiente las plantas deben tener tallos largos de modo que sus hojas puedan obtener la luz...”). Ello servía para que los participantes pudieran despejar cualquier problema en sus representaciones mentales. Gerjets, Scheiter y Catrambone (2006) proporcionaron a sus participantes *explicaciones instructivas muy elaboradas* a fin de que lograran entender sin problema los principios subyacentes a los pasos en la resolución de ejercicios resueltos de probabilidad (e.g., “Los problemas de este tipo presentan dos características importantes. Primero, ...”). También Sánchez et al. (2008) desarrollaron ayudas de esta clase. Ellos las llamaron explicaciones adaptadas o *tailored explanations*, ya que abordaban problemas específicos en la representación mental de un aprendiz dado (e.g., “Es la actividad en las dorsales la que compensa la pérdida. En las dorsales...”).

En síntesis, todos estos dispositivos pueden considerarse ayudas al proceso de reparación. Pueden considerarse así porque todas asisten a los aprendices a la hora de reparar sus incompletas o defectuosas representaciones. Un asunto de interés es que difieren en el grado de apoyo que brindan: las respuestas correctas ofrecen poco, las elaboraciones bastantes más.

3.4.3. Efectividad de las ayudas para detectar y reparar. Las ayudas para detectar y reparar funcionan: quienes las reciben aprenden más que quienes no. Aunque hay también casos en los que no se da tal efectividad, los resultados globalmente considerados arrojan el resultado anunciado. Comprobar que las ayudas para detectar y reparar son efectivas puede hacerse de varias maneras. Puede compararse el aprendizaje que alcanzan quienes disfrutan de ambas en comparación con quienes no reciben ninguna de ellas. Puede también compararse el rendimiento de los que reciben ambas con el de los que reciben sólo una de ellas. Por último, es posible también explorar el efecto de una de ellas en solitario comparando el rendimiento de quienes reciben una con el de los que reciben ninguna. Vamos a ver qué ha ocurrido en cada una de los casos.

3.4.3.1. Ayudas para detectar y reparar versus nada. Los aprendices que reciben ayudas tanto para detectar problemas en sus representaciones mentales como para repararlos

aprenden más que los que no reciben ninguna de ellas. Sobre esto hay varias evidencias. Alevén y Koedinger (2002) encontraron que los participantes a los que se les brindaba apoyo en forma de error messages y hints rendían mejor que quienes habían utilizado un CBLE idéntico pero sin ayudas. El mismo patrón de resultados lo encontramos en los experimentos de Atkinson et al. (2003), esta vez empleando error messages y respuestas correctas. VanLehn et al. (2007), quienes también usaron ayudas tanto para detectar como para reparar (en múltiples formas: open prompts y error messages, por un lado; hints, fill-in-the-blanks y elaboraciones, por otro), hallaron el mismo resultado en dos de sus experimentos. De estos resultados se desprende que es posible ayudar tanto a detectar como a reparar y que si no lo hacemos, los aprendices no pueden hacer ambas cosas por sí mismos.

Hay, sin embargo, algunas evidencias que reflejan un panorama diferente. En el experimento de Conati y VanLehn (2000), en el que se hizo uso de ayudas muy parecidas a las de Alevén y Koedinger, no hubo diferencias entre los participantes que recibían ayuda y los que no la recibían. Y lo mismo ocurrió en cinco de los siete experimentos implementados por VanLehn y compañía (2007).

3.4.3.2. Ayudas para detectar y reparar versus una de ellas. Recibir las dos ayudas es mejor que recibir sólo una de ellas, a tenor de las evidencias que siguen. Por un lado, hay estudios que revelan que recibir ayudas para detectar y reparar es mejor que recibir sólo las de detección. En el experimento de Moreno (2004) los participantes que recibieron tanto feedback correctivo como el explicativo aprendieron más y mejor que quienes recibieron sólo el correctivo. Y lo mismo nos encontramos en el primer experimento de Moreno y Mayer (2005). Cabe decir que aún advertidos de los errores cometidos, los aprendices necesitan ayuda adicional para repararlos.

Ahora bien, hay también alguna evidencia de lo contrario. McKendree (1990) no pudo encontrar la misma ventaja. La investigadora pidió a sus participantes que aprendiera geometría a partir de un CBLE. Una de las condiciones incluía feedback correctivo y explicativo y la otra incluía sólo el correctivo. No hubo diferencias entre ellas.

Por otro lado, algunos estudios muestran que la combinación de ayudas para detectar y reparar es mejor que las ayudas para reparar en solitario. Moreno y Mayer (2005) manipularon en un experimento la primera parte del feedback. Así, los participantes podían recibir las elaboraciones sobre qué características de una planta la harían sobrevivir bajo ciertos ambientes de dos maneras: bien como respuesta a un feedback correctivo o bien a secas. Los que recibieron correctivo y explicativo rindieron mejor. Algo semejante se puso en evidencia en los experimentos de Sánchez et al. (2008). Quienes recibieron tanto *impassetriggers* como *tailored explanations* rindieron mejor que quienes recibieron sólo las segundas. Estos resultados sugieren que para aprovechar al máximo una ayuda de reparación, ésta debe ir precedida de una de detección. Al saber que uno debe llevar a cabo una reparación (tras recibir la ayuda de detección), la elaboración contenida en la ayuda de reparación se procesa más profundamente.

De nuevo, hay alguna prueba no coherente con esta conclusión. Sleeman, Kelly, Martinak, Ward y Moore (1989) pidieron a sus participantes que aprendiesen a resolver

problemas de álgebra. A algunos les fueron suministradas ayudas para detectar y reparar. En este caso, se identificaban las estrategias erróneas de resolución que los participantes aplicaban (feedback correctivo) para después instruirles en cómo debían resolverlo y por qué razón debían hacerlo de ese modo (feedback explicativo). En tres experimentos no hubo diferencias en el rendimiento entre los que recibieron ambas ayudas y los que recibieron sólo las segundas.

3.4.3.3. Ayudas para detectar o reparar versus nada. Contamos con evidencias de que recibir uno de los tipos de ayuda es mejor que nada. Por un lado, hay evidencias de que las ayudas para detectar en solitario son eficaces. Esto es lo que encontraron Scworm y Renkl (2006). los participantes a los que se proporcionó open prompts superaron a sus compañeros en las tareas que se les plantearon tras la fase de aprendizaje. Resultados como éste sugieren que sin algún tipo de ayuda los aprendices no son capaces de supervisar su comprensión para detectar malentendidos.

Claro que Hausmann y Chi (2002) no pueden decir lo mismo a partir de sus datos. En su experimento los participantes aprendieron biología a partir de un CBLE. La mitad recibió open prompts mientras que el resto no recibió ayudas. No hubo diferencias entre las condiciones en las medidas de aprendizaje.

Por otro lado, se ha observado que es mejor para los aprendices recibir ayudas para reparar que recibir nada. Nückles et al. (2005) solicitaron a sus participantes que aprendiesen informática con la ayuda de un tutor con el que mantenían un diálogo vía ordenador. Dependiendo de la condición experimental, los participantes recibían tailored explanations o explicaciones generales sobre informática. Los de la primera condición superaron a los de la segunda al responder una serie de preguntas que se les plantearon tras la fase de tutoría. Esto parece indicar que si los aprendices deben corregir sus representaciones mentales necesitan ayudas específicamente diseñadas para tal fin.

No obstante, hay también evidencia de que esta clase de ayudas puede resultar ineficaz. Gerjets et al. (2006) encontraron que sus explicaciones instructivas muy elaboradas tuvieron un impacto nulo en el aprendizaje.

3.4.3.4. ¿Cómo explicar la falta de efectividad (cuando se produce)? Afortunadamente, hay una forma de explicar la ineficacia de las ayudas en algunos estudios. Tal como sugieren Acuña, García-Rodicio y Sánchez (en preparación), las ayudas pueden no resultar efectivas si los aprendices cuentan ya con cierto nivel de conocimiento previo. Sólo “cuando el salto es grande”, las ayudas funcionan. Con esta expresión nos referimos a que sólo si el aprendiz debe incorporar un número considerable de ideas y relaciones a su modelo mental las ayudas funcionan como se espera.

Hay evidencias que respaldan este modo de razonar. VanLehn et al. (2007) llegan exactamente a la misma conclusión haciendo un meta-análisis de sus siete experimentos. Conati y VanLehn (2000) encontraron que, si sólo consideraban los datos de los participantes que sabían menos sobre física antes de comenzar el experimento, las ayudas eran efectivas. Y Acuña et al. también aportan datos en esa dirección. En su experimento, sólo quienes llegaban

a la sesión experimental con modelos mentales pobres sobre geología se beneficiaban de las ayudas. Quienes ya poseían un modelo con ciertas ideas y relaciones incorporadas rendían igual con y sin ayudas.

3.4.3.5. Conclusión. Habida cuenta de que los procesos de detección y reparación son críticos pero no se ejecutan con facilidad, debemos brindar ayudas para ambos. Hay un buen arsenal de ayudas disponible. En este subapartado hemos revisado estudios que ponían a prueba su eficacia. En general, las ayudas funcionan. Eso sí, son particularmente efectivas para quienes saben poco sobre el tópico a aprender y deben incorporar un buen número de ideas y relaciones.

4. Cómo presentar las ayudas para detectar y reparar.

Parece justificado, en virtud de todo lo expuesto hasta ahora, que debemos ofrecer ayudas para detectar y reparar a los aprendices. También que esas ayudas pueden resultar efectivas. Ahora la cuestión es, ¿cómo debemos presentarlas? Con esto nos referimos a que, dada la versatilidad de los CBLEs, cabe construir las ayudas de maneras diversas. Hay que decidir si las ayudas serán verbales o pictóricas y si las presentaremos en modalidad auditiva o visual. Nosotros estamos interesados en las verbales (que son a las que nos hemos referido en la sección anterior) y lo que nos ocupa es decidir qué modalidad es mejor para su presentación, la auditiva o la visual. Tal como se comprobará, no hay pautas claras sobre el uso de la modalidad en el diseño de ayudas verbales. No hay razones teóricas, porque el principio de modalidad no es aplicable a las ayudas verbales. No hay razones empíricas, porque la investigación sobre el efecto de la modalidad de las ayudas verbales arroja resultados dispares. En esta sección presentaremos el marco de investigación que nosotros proponemos para colmatar esta brecha. El marco plantea pautas sobre cómo emplear la modalidad y es coherente con los datos (en apariencia) discrepantes en la literatura.

El plan para esta sección es, pues, el que sigue. Comenzaremos describiendo qué posibilidades de diseño de ayudas nos brindan los CBLEs. Después discutiremos por qué no está claro cuál de dos de esas posibilidades es mejor: por qué no sabemos si es mejor usar lo auditivo o lo visual para las ayudas verbales. Se hará evidente que es preciso confeccionar un marco de investigación del que se deriven pautas de diseño. Y eso es lo que haremos acto seguido: presentaremos con cierta profundidad la aproximación funcional y las predicciones que se deducen de ella. Terminaremos presentando una panorámica de los experimentos que serán descritos en el capítulo próximo.

4.1. Diseñando ayudas para un CBLE: posibilidades múltiples.

Los CBLEs deben incorporar ayudas que asistan a los aprendices en la detección y reparación de problemas de comprensión. Esto es lo que se deduce de lo que hemos presentado anteriormente. Sabemos que los CBLE presentan información verbal y pictórica y a través de las modalidades auditiva y visual. De modo que las ayudas podrán ser tanto verbales como pictóricas y tanto auditivas como visuales. Esto da lugar a cuatro configuraciones posibles. ¿Cuáles exactamente?

Una opción es emplear ayudas pictóricas, es decir, no basadas en palabras. Si, además, hacemos uso de la modalidad visual tenemos *ayudas pictóricas visuales*. Estas ayudas pueden adoptar la forma de flechas o colores. Imaginemos que una animación (con su narración correspondiente) nos instruye sobre el mecanismo que hace volar a los aviones. Imaginemos que el diseñador supone que un aspecto de la animación suele pasar desapercibido. Una manera de decirnos “no olvides considerar el aspecto X” es incluir flechas que nos hagan percatarnos del valor de considerar X y tratar de prestarle suficiente atención.

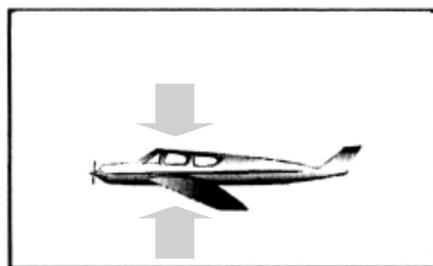


Figura 3. Ayudas pictóricas para detectar, ejemplo.

Este tipo de ayudas se han empleado en más de una ocasión, según nos consta (DeKoning, Tabbers, Rikers & Paas, 2007; Mautone & Mayer, 2001, 2007). Las ayudas pictóricas visuales pueden adoptar también la forma de colores. Imaginemos a un aprendiz que aprende física a partir de un CBLE. El CBLE le presenta ejercicios resueltos y le invita a explicar el razonamiento que hay detrás de los pasos en la resolución. El CBLE interpreta las explicaciones del aprendiz y proporciona error messages coloreando las partes de la respuesta del aprendiz que son correctas/incorrectas. Eso es lo que hace el CBLE desarrollado por Conati y VanLehn (2000) empleando el color verde para los elementos correctos y el rojo para los incorrectos.

Si, con ayudas no verbales, optamos por la modalidad auditiva tenemos las *ayudas pictóricas auditivas*. Imaginemos que el mismo CBLE utilizado por Conati y VanLehn (2000) proporcionase un aplauso o un abucheo para las explicaciones correctas o incorrectas, respectivamente. En tal caso, se trataría de una ayuda que facilita al aprendiz la detección de problemas en sus representaciones mentales. Eso sí, no sería verbal, porque no emplea palabras. Y tampoco visual porque no se ve, se oye.

Las ayudas pueden ser también verbales y pueden presentarse auditivamente para generar *ayudas verbales auditivas*. Pongamos que un aprendiz aprende botánica a partir de un CBLE. El CBLE describe distintos ambientes (condiciones de luz, agua, etc.) mediante animaciones con narración incorporada. Terminada la descripción, el CBLE solicita al aprendiz que seleccione las raíces, tallos y hojas más apropiados para que una planta prospere en el ambiente en cuestión. Si las elecciones del aprendiz no son las correctas el CBLE puede pronunciar algo como esto: “El tallo en este ambiente debería ser largo para colocar las hojas cerca del sol”. Eso permitiría al aprendiz corregir las limitaciones en sus representaciones mentales. Ayudas como estas son las que ha desarrollado Moreno (2004; Moreno & Mayer,

2005), que fueron objeto de discusión hace un momento.



Figura 4. Ayudas verbales para reparar, ejemplo.

Esas mismas ayudas podrían presentarse también en modalidad visual dando lugar a las *ayudas verbales visuales*. De hecho, la propia Moreno (e.g., Moreno & Mayer, 2002) manipuló la modalidad de presentación de estas ayudas. De modo que también las presentó en modalidad visual.

Un diseñador ha de decidir si prefiere un tipo u otro de información y si prefiere una otra modalidad. Nosotros estamos interesados únicamente en las ayudas verbales. Y, como ya venimos diciendo desde el comienzo, nos interesa saber si para estas ayudas es mejor la modalidad auditiva o la visual. Enseguida exploramos esta cuestión.

4.2. No hay pautas sobre cómo usar la modalidad para las ayudas verbales.

Como acabamos de ver, las ayudas verbales pueden presentarse tanto en la modalidad visual como en la auditiva. Una pregunta que los diseñadores de CBLE deben responder es cuál de las dos modalidades es mejor, esto es, cuál de ellas hace que las ayudas sean más eficaces promoviendo aprendizaje. No obstante, no hay aún una respuesta clara a este interrogante. Podemos acudir a argumentos teóricos y empíricos en busca de respuesta. Eso es precisamente lo que vamos a hacer enseguida. Anunciamos, eso sí, que ninguno de ellos servirá para despejar el interrogante. Específicamente, veremos que no sabemos cómo debemos presentar las ayudas porque (a) el principio de modalidad no es aplicable y porque (b) la investigación sobre el efecto de la modalidad de las ayudas presenta resultados dispares.

4.2.1. El principio de modalidad no es aplicable. El lector ducho sobre el tema habrá pensado en el principio de modalidad como un buen candidato para responder al interrogante. Pero no. Este principio sirve sólo para diseñar información verbal que acompaña a información pictórica y resulta que las ayudas no llevan asociada información pictórica (al menos no necesariamente). Vamos a ver primero qué dice este principio y qué apoyo ha recibido para después ver algunas críticas y por qué no nos sirve.

4.2.1.1. ¿Qué dice el principio de modalidad? Como es sabido, los CBLEs incluyen

tanto información verbal como pictórica. Así, un CBLE puede consistir en una animación sobre el funcionamiento de una bomba de hinchar o sobre el freno de tambor de un coche con su correspondiente información verbal describiendo lo que la animación muestra. Una buena pregunta es cómo debe presentarse esa información verbal, auditiva o visualmente. Richard Mayer y John Sweller han propuesto el *principio de modalidad* para resolver este problema. El principio de modalidad establece que allí donde la información verbal acompaña a información pictórica debe emplearse la modalidad auditiva. ¿Por qué? Muy sencillo. De esa manera se evita una sobrecarga de la memoria de trabajo y, de paso, se expande su capacidad efectiva. Lo explicamos detenidamente.

Mayer y Sweller parten de un supuesto. De acuerdo con los investigadores la memoria de trabajo posee dos sub-procesadores relativamente independientes (Mayer, 2001; Sweller et al., 1998). Ambos investigadores se apoyan en el trabajo previo de Baddeley (e.g., 1997) para sostener eso. Uno de los sub-procesadores es la *agenda visoespacial*, la cual mantiene y manipula información visual-pictórica. El otro es el *bucle fonológico*, el cual mantiene y manipula información auditivo-verbal. De este modo, las imágenes (animaciones, ilustraciones) son procesadas por la agenda visoespacial. Las palabras orales (narraciones) son procesadas por el bucle fonológico. Las palabras escritas (textos), dado que son presentadas visualmente, son procesadas por la agenda visoespacial (al menos inicialmente; eventualmente podrían transformarse a un código fonológico).

Asumida la existencia de dos sub-procesadores independientes, Mayer razona así (Mayer, 2001; Mayer & Moreno, 1998). Cuando el aprendiz debe procesar imágenes y palabras, la agenda visoespacial sufrirá una sobrecarga si los dos tipos de información se presentan visualmente. En estos casos, imágenes y palabras van en un solo sub-procesador, que consecuentemente quedará saturado. Si las palabras se presentan auditivamente y (como no puede ser de otro modo) las imágenes se presentan visualmente, la carga se reparte entre dos sub-procesadores. En estos casos, las imágenes van en la agenda y las palabras en el bucle: ninguno queda sobrecargado. Así, cabe esperar que un aprendiz aprenda más de una presentación con ilustraciones/animaciones y narraciones que de una presentación con ilustraciones/animaciones y texto.

Sobre la base del citado supuesto, Sweller razona así (Mousavi, Low & Sweller, 1995; Sweller et al., 1998). Si cada sub-procesador está asociado a una modalidad (lo visual en la agenda y lo auditivo en el bucle) y los dos son relativamente independientes, la capacidad efectiva de la memoria de trabajo está determinada por la modalidad. Usar una presentación con dos modalidades permite explotar todo el potencial de la memoria de trabajo: dos sub-procesadores que suman sus capacidades de procesamiento respectivas. Usar una presentación con una sola modalidad utiliza sólo un sub-procesador, siendo su capacidad menor que la de la conjunción de dos. Por eso, un aprendiz que recibe ilustraciones/animaciones con narraciones (y, por tanto, usa toda la capacidad de su memoria de trabajo) aprenderá más que el que recibe ilustraciones/animaciones con texto (ya que usa sólo parte de su capacidad de memoria de trabajo).

Sea por evitar la sobrecarga de un sub-procesador o por expandir la capacidad de la memoria de trabajo usando sus dos sub-procesadores, ambos investigadores predicen que las

palabras que acompañan a imágenes deben presentarse auditivamente. De esa manera se logra el mayor aprendizaje posible.

4.2.1.2. *¿Con qué apoyo cuenta el principio de modalidad?* El principio de modalidad predice que la combinación de imágenes y narraciones engendra un mayor aprendizaje que la combinación de imágenes y textos. Lo predice sobre la base de los argumentos ya presentados. Ahora, ¿se ha demostrado tal cosa? Lo cierto es que sí, y sobradamente (cf. Ginns (2005) para un meta-análisis).

Nosotros hemos contabilizado 29 experimentos en los que se pone a prueba el principio. De todos ellos, sólo un par arrojan resultados desconfirmatorios, el resto respalda el principio. En la tabla 20 figuran todos ellos. Algunas columnas necesitan explicación. A la izquierda figura la referencia del experimento y la materia o tópico sobre el que versa el material empleado en él. Los tipos de imágenes que hemos distinguido varían en la dimensión estático-dinámico. Así, las ilustraciones son imágenes estáticas (e.g., tablas, diagramas, fotografías) mientras que las animaciones son imágenes dinámicas (e.g., simulaciones, vídeos). No hemos considerado otras dimensiones como realista-esquemática o detallada-simplificada. La relación verbal-pictórico se refiere a la combinación palabras-imágenes en el material: puede ser secuencial (si las imágenes van primero y las palabras después o viceversa) o simultánea (si todo se presenta de una vez). El diseño recoge las variables independientes del experimento. La columna siguiente es de enorme interés. Una de las variables del experimento puede funcionar como factor modulador si matiza el alcance del principio de modalidad. Por ejemplo, el principio de modalidad puede quedar confirmado por determinado experimento pero sólo bajo la condición de que los participantes tengan una capacidad de memoria de trabajo particularmente limitada (indicado por una prueba). La capacidad de la memoria de trabajo actuaría así como factor modulador. Es importante señalar que (a) la variable moduladora puede ser externa al diseño (e.g., elijo usar materiales más complejos de lo habitual en el experimento) y que (b) una variable del diseño puede no ser moduladora sino servir a otros fines. Hemos diferenciado dos clases de variable dependiente, las on y las off-line. Es interesante hacer esta distinción porque el principio de modalidad se ha confirmado usando las dos clases. En resultados hacemos referencia a si el principio quedó o no respaldado (al margen de que se confirme con restricciones). Vamos a revisar algunos experimentos en los que se confirma el principio, algunos en los que no se confirma y vamos a ver qué restricciones (factores moduladores) es preciso hacer al principio.

Tabla 20. Experimentos ($N = 29$) sobre el impacto de la modalidad de la información verbal (auditiva/visual) acompañando (secuencial/simultáneamente) información pictórica (animación/ilustración) sobre distintas variables (on y off-line).

Experimento	Tópico	Tipo imagen	Relación verbal-pictórico	Diseño	Factor modulador	Variable dependiente	Resultados
Brünken, Steinbacher, Plass & Leutner, 2002 (1)	Sistema circulatorio humano	Ilustración	Simultánea	Modalidad (auditivo VS visual)	No	On-line (rendimiento en tarea secundaria, <i>dual-task</i>)	Positivos
Brünken et al.,	Geografía	Ilustración	Simultánea	Modalidad	No	On-line	Positivos

2002 (2)	de Florencia			(auditivo VS visual)		(rendimiento en tarea secundaria, <i>dual-task</i>)	
Brünken, Plaas & Leutner, 2004 (1)	Sistema circulatorio humano	Ilustración	Simultánea	Modalidad (auditivo VS visual)	No	On-line (rendimiento en tarea secundaria, <i>dual-task</i>)	Positivos
Brünken et al., 2004 (2)	Geografía de Florencia	Ilustración	Simultánea	Modalidad (auditivo VS visual)	No	On-line (rendimiento en tarea secundaria, <i>dual-task</i>)	Positivos
Craig, Gholson & Driscoll, 2002 (2)	Formación tormentas	Animación	Simultánea	Modalidad (auditivo VS visual VS ambas)	No	Off-line (recuerdo y transferencia)	Positivos
Harksamp, Mayer & Suhre, 2007 (1)	Conducta animal	Ilustración	Simultánea	Modalidad (auditivo VS visual)	No	Off-line (recuerdo y transferencia)	Positivos
Harksamp et al., 2007 (2)	Conducta animal	Ilustración	Simultánea	Modalidad (auditivo VS visual) x Tiempo (largo VS corto)	Sí, tiempo	Off-line (recuerdo y transferencia)	Positivos
Jeung, Chandler & Sweller, 1997 (1)	Geometría	Ilustración	Simultánea	Modalidad (auditivo VS visual) x Pistas pictóricas (sí VS no)	Sí, pistas pictóricas (para imágenes complejas)	Off-line (problemas)	Positivos
Jeung et al., 1997 (2)	Geometría	Ilustración	Simultánea	Modalidad (auditivo VS visual) x Pistas pictóricas (sí VS no)	Sí, pistas pictóricas (para imágenes complejas)	Off-line (problemas)	Positivos
Jeung et al., 1997 (3)	Geometría	Ilustración	Simultánea	Modalidad (auditivo VS visual) x Pistas pictóricas (sí VS no)	Sí, pistas pictóricas (para imágenes simples)	Off-line (problemas)	Positivos
Kalyuga, Chandler & Sweller, 1999 (1)	Soldadura de materiales	Animación	Simultánea	Modalidad (auditiva VS visual VS ambas)	No	Off-line (preguntas y problemas)	Positivos
Kalyuga et al., 2000 (1)	Taladrado	Ilustración	Simultánea	Modalidad (auditiva VS visual) x Imagen (sí VS no) x Fase (novato VS experto)	Sí, fase	On y Off-line (carga mental, problemas)	Positivos
Leahy, Chandler & Sweller, 2003 (1)	Temperatura y hora día	Ilustración	Simultánea	Modalidad (auditiva VS visual) x Complejidad	Sí, complejidad material	Off-line (preguntas)	Positivos

				(alta VS baja)			
Mayer & Moreno, 1998 (1)	Formación tormentas	Animación	Simultánea	Modalidad (auditiva VS visual)	No	Off-line (recuerdo y transferencia)	Positivos
Mayer & Moreno, 1998 (2)	Frenos de tambor coche	Animación	Simultánea	Modalidad (auditiva VS visual)	No	Off-line (recuerdo y transferencia)	Positivos
Mayer, Gayle & Mayer, 2003 (1)	Circuitos eléctricos	Animación	Simultánea	Modalidad (auditiva VS visual)	No	Off-line (transferencia)	Positivos
Moreno & Mayer, 1999 (1)	Formación tormentas	Animación	Simultánea	Modalidad (auditiva VS visual cerca VS visual lejos)	No	Off-line (recuerdo y transferencia)	Positivos
Moreno & Mayer, 1999 (2)	Formación tormentas	Animación	Secuencial	Modalidad (auditiva VS visual) x Relación (secuencial VS simultánea)	No	Off-line (recuerdo y transferencia)	Positivos
Mousavi et al., 1995 (1)	Geometría	Ilustración	Simultánea	Modalidad (auditiva VS visual VS ambas)	No	Off-line (problemas)	Positivos
Mousavi, Low & Sweller, 1995 (2)	Geometría	Ilustración	Simultánea	Modalidad (auditiva VS visual)	No	Off-line (problemas)	Positivos
Mousavi et al., 1995 (3)	Geometría	Ilustración	Secuencial	Modalidad (auditiva VS visual)	No	Off-line (problemas)	Positivos
Mousavi et al., 1995 (4)	Geometría	Ilustración	Secuencial	Modalidad (auditiva VS visual)	No	Off-line (problemas)	Positivos
Seufert, Schütze & Brünken, 2008 (2)	Biología	Ilustración	Simultánea	Modalidad (auditiva VS visual) x Memoria trabajo (alta VS baja)	Sí, capacidad memoria trabajo	Off-line (preguntas recuerdo y transferencia)	Positivos
Tabbers, 2006	Formación tormentas	Animación	Simultánea	Modalidad (auditiva VS visual)	No	Off-line (recuerdo y transferencia)	Negativos
Tabbers, Martens & van Merriënboer, 2004	Psicología de la instrucción	Ilustración	Simultánea	Modalidad (auditiva VS visual) x Pistas pictóricas (sí VS no)	(Ritmo de presentación)	On y off-line (carga mental, recuerdo y transferencia)	Negativos
Tindall-Ford, Chandler & Sweller, 1997 (1)	Instalación eléctrica	Ilustración	Simultánea	Modalidad (auditiva VS visual cerca VS visual lejos)	No	Off-line (problemas)	Positivos
Tindall-Ford et al., 1997 (2)	Instalación eléctrica	Ilustración	Simultánea	Modalidad (auditiva VS visual)	No	Off-line (problemas)	Positivos

Tindall-Ford et al., 1997 (3)	Instalación eléctrica	Ilustración	Simultánea	Modalidad (auditiva VS visual) x Complejidad (alta VS baja)	Sí, complejidad material	On y off-line (carga mental, problemas)	Positivos
Wouters, Paas & van Merriënboer, 2008	Cálculo probabilidades	Animación	Simultánea	Modalidad (auditiva VS visual) x Prompts (sí VS no)	Sí, prompts	Off-line (transferencia)	Positivos

En lo que sigue vamos a repasar alguno de los experimentos de la tabla. Comenzaremos con uno que utilizó animaciones con su correspondiente información verbal presentada simultáneamente. Mayer & Moreno (1998, Exp. 2) solicitaron a sus participantes que aprendiesen cómo funcionan los frenos de tambor de un coche a partir de un CBLE. El CBLE, como hemos anunciado, contenía tanto animaciones como información verbal acompañándolas simultáneamente. La información verbal se presentó bien auditivamente (narración) o visualmente (texto). Tras consultar el CBLE los participantes trataron de recordar todo lo posible sobre lo que habían estudiado y resolvieron tareas de transferencia. Los participantes en la condición narración superaron a los de la condición texto en ambas tareas.

Cuando la imagen es una ilustración y la información verbal se presenta simultáneamente también la narración es superior al texto. Mousavi et al. (1995, Exp. 2) pidieron a un grupo de participantes que aprendiese cómo resolver problemas de geometría usando para ello un material instructivo. El material presentaba ejercicios resueltos de geometría incluyendo ilustraciones (figuras geométricas) e información verbal asociada, todo ello presentado de una vez. La mitad de los participantes recibía la información verbal en narración y la otra mitad en texto. Una vez consultado el material los participantes tenían que resolver problemas parecidos y diferentes a los consultados en la fase de estudio. El tiempo que necesitaron para resolverlos fue la variable dependiente. Los resultados mostraron que quienes recibieron narraciones emplearon menos tiempo que los que recibieron texto.

La ventaja de la narración sobre el texto se da también cuando la presentación de la animación y la información verbal no es simultánea. Moreno y Mayer (1999, Exp. 2) hicieron que sus participantes aprendiesen cómo se forman las tormentas a partir de un CBLE. El CBLE incluía tanto animaciones como información verbal que describía lo que tenía lugar en ellas. La mitad de los participantes veía primero un fragmento de animación, después la parte correspondiente de información verbal, luego otro fragmento de animación, información verbal asociada, y así sucesivamente. La otra mitad recibía animación e información verbal todo de una vez. Otra variable fue la modalidad, por supuesto: unos recibían lo verbal en audio y otros visualmente. Al terminar de consultar el CBLE ya sabemos lo que tocaba: recordar y resolver tareas de transferencia. ¿Resultado? Una nueva confirmación del principio de modalidad: los de la condición narración superaron a los de la condición texto en ambas medidas. Y lo interesante es que no hubo interacción con el factor simultáneo/secuencial, esto es, lo auditivo gana en las dos circunstancias.

El mismo resultado se ha encontrado con ilustraciones e información verbal presentadas secuencialmente. Mousavi et al. (1995, Exp. 4) replicaron el experimento que hemos descrito hace un instante pero introduciendo un factor adicional: presentación simultánea/secuencial de lo pictórico y lo verbal. Los participantes en la condición con modalidad auditiva emplearon menos tiempo que aquellos en la condición con modalidad visual independientemente de la variable simultáneo/secuencial.

Es interesante que no sólo la predicción que se deriva del principio sino los razonamientos que lo sustentan han sido confirmados. El principio asume que la modalidad visual es peor porque debe compartir recursos de memoria con los que deben dedicarse a la imagen, dejando inutilizados los del bucle fonológico. Tanto una competición por los recursos como una expansión de la capacidad efectiva de la memoria de trabajo deben reflejarse en medidas on-line. Es decir, quienes reciben la información verbal en modalidad visual deben disponer de menos capacidad de procesamiento durante la consulta del material. Para averiguar si sucede tal que así es posible hacer uso de dos técnicas. Una son las escalas de esfuerzo o carga mental. Una escala de este tipo es cumplimentada por el propio participante en el experimento. En varios momentos durante la consulta del material se le pregunta al participante “¿qué nivel de actividad mental has mantenido para entender el material hasta este punto?”. El participante debe responder juzgando el nivel de actividad que estima de cero a siete (o de cero a nueve, según la escala). Pese a que alguien pueda pensar que la valoración subjetiva es imprecisa, lo cierto es que los autores de la escala afirman que exhibe niveles razonables de fiabilidad y validez (e.g., Sweller et al., 1998). Tindall-Ford et al. (1997, Exp. 3) usaron esta escala. Los investigadores solicitaron a sus participantes que aprendiesen a montar instalaciones eléctricas a partir de un material con ilustraciones (tablas) e información verbal presentados simultáneamente. Quienes recibieron la información verbal en forma de narraciones manifestaron experimentar una carga mental significativamente menor que la de quienes recibieron textos. Por consiguiente, la información en modalidad visual consume más recursos de memoria que la auditiva, lo que es consistente con el principio.

La otra técnica es la *dual-task*. En este caso se trata de plantear una tarea doble a los participantes. Por supuesto, hacer dos cosas a la vez es difícil. La cuestión es que lo es más cuanto más exigente resulta la tarea primaria. Dada una tarea secundaria sencilla, si el rendimiento en ésta es peor cuando la tarea primaria es X que cuando es Y, cabe concluir que X demanda más recursos que Y. Brünken et al. (2002, Exp. 1) plantearon una *dual-task* a sus participantes. La tarea primaria era aprender de un CBLE que incorporaba ilustraciones con su correspondiente información verbal sobre el sistema circulatorio humano. La mitad del tiempo lo verbal se presentaba en audio y la otra mitad visualmente (diseño intra-sujetos). La tarea secundaria era pulsar la barra espaciadora cada vez que una “A” que aparecía en la parte superior de la pantalla cambiaba de color. La variable dependiente era el tiempo que tardaban los participantes en pulsar. Los investigadores se aseguraban de que los participantes, en efecto, aprendían sobre el sistema circulatorio planteando unas tareas de aprendizaje (si no los participantes podrían simplemente haberse concentrado en la “A” ignorando todo lo demás). En la condición narración los tiempos de respuesta eran significativamente menores que los de la condición texto. Lo mismo sucedió en el segundo experimento en el que usaron un CBLE sobre la geografía de Florencia. Como la tarea secundaria debe realizarse empleando recursos de la agenda visoespacial y, a tenor de lo ocurrido con la condición texto, cabe

deducir que, en efecto, la modalidad visual consume recursos de la agenda robando espacio a las imágenes. De nuevo, principio confirmado.

4.2.1.3. Circunstancias moduladoras. Como se aprecia, tanto con animaciones como con ilustraciones, tanto en las presentaciones secuenciales como en las simultáneas y tanto con medidas on-line como con medidas off-line el principio de modalidad se ha visto confirmado. Ahora bien, del listado de 29 experimentos unos cuantos sirven para acotar la generalizabilidad del principio. El principio no funciona bajo cualquier circunstancia y algunos experimentos han identificado cuáles son las circunstancias que anulan el efecto de modalidad. Una de las circunstancias es la complejidad del material a aprender. Sólo si es complejo el principio de modalidad funciona. Leahy et al. (2003, Exp. 1) presentó a sus participantes una ilustración (diagrama) acompañada de información verbal presentada bien visual o bien auditivamente sobre un tópico de física. Otra variable independiente fue el nivel de complejidad de los materiales: baja (que incluía pocas relaciones entre los elementos) o alta (muchas relaciones). Una vez consultado el material los participantes respondían a preguntas sobre lo aprendido. La modalidad auditiva fue mejor que la visual sólo cuando el material era complejo. Otra es la capacidad de la memoria de trabajo. Seufert et al. (2008, Exp. 2) encontraron que el efecto aparecía sólo cuando la capacidad de la memoria de trabajo (medida con una prueba parecida a la de Daneman y Carpenter (1980)) era baja. Y hay más, tal como se observa en la tabla que hemos confeccionado. Quien esté interesado puede consultarla. Para nuestros propósitos es suficiente con las que hemos referido.

4.2.1.4. Problemas. Con todo, el principio de modalidad tiene algunos problemas. Para empezar, no siempre se ha visto confirmado (y no nos referimos a las circunstancias que acotan su generalizabilidad). En un experimento publicado en las actas de un congreso Tabbers (2006) replicó punto por punto el experimento 1 de Mayer y Moreno (1998). En él los participantes aprendían sobre la formación de montañas a partir de animaciones e información verbal presentadas simultáneamente. Como siempre, la modalidad de lo verbal fue manipulada: auditiva versus visual. Como siempre también los participantes recordaban y resolvían tareas de transferencia al terminar la consulta de los materiales. Pues bien, Tabbers no encontró diferencias significativas entre las condiciones en ninguna de las medidas. Tabbers et al. (2004) no sólo no encontraron confirmación al principio de modalidad sino que se toparon con el resultado contrario: los de la modalidad visual rindieron significativamente mejor que los de la auditiva. Siendo dos contra 27 podemos hacer la vista gorda.

En segundo lugar, se basa en una mala interpretación del modelo de Baddeley. Tanto Mayer como Sweller confunden tipo de información (verbal/pictórica) con modalidad (auditiva/visual). Baddeley sostiene que el bucle fonológico procesa información no auditiva sino verbal, articulable: “a pesar de su presentación visual, los sujetos lo pronunciarán [lo verbal] subvocálicamente” (2003, p. 191). Mayer y Sweller asumen que lo que entra a través del oído se dirige al bucle y que lo que entra a través de la vista se dirige a la agenda, sea la información que sea (verbal o pictórica). No es verdad. Como decimos, Baddeley afirma que la información verbal, incluso si es visual, se dirige al bucle. Y es que hay pruebas de que ante una presentación visual de palabras se produce una activación fonológica (e.g., van Orden, Johnston & Hale, 1988). No sólo eso, el sonido de coche, de un terremoto o de una aspiradora no es procesado por el bucle porque no es articulable. Luego hay una disociación entre tipos

de información y modalidades en el modelo de Baddeley que Mayer y Sweller parecen soslayar felizmente. No obstante, no es menos cierto que la información verbal presentada visualmente *también* consume recursos de la agenda visoespacial. ¿Cuáles son las pruebas? Una es que cuando hay que “visualizar” lo que dice un texto, quienes leen el texto resuelven peor tareas relacionadas con él que quienes lo escuchan (Levin & Divine-Hawkins, 1974). El visualizar la historia relatada en el texto es algo que implica el uso de la agenda visoespacial, claro está. El hecho de que la presentación visual del mismo sea peor que la auditiva sugiere que lo escrito comparte recursos con el procesamiento de imágenes. Otra prueba es que el movimiento de los ojos a través de las palabras impresas es gestionado por la agenda visoespacial (Postle, Izdikowski, Sala, Logie & Baddeley, 2006). Así, aunque sólo sea por la coordinación de los movimientos oculares, la información verbal presentada vía visual pone en juego la agenda visoespacial. Por lo tanto, aunque Mayer y Sweller parecen ignorar (lo cual es incorrecto) que lo verbal presentado visualmente es finalmente procesado por el bucle, también es cierto que necesita consumir recursos de la agenda. Así que tenemos que, de algún modo, darles la razón a Mayer y Sweller pero invitándoles, eso sí, a recalcar eso del *at least initially* (Mayer suele puntualizar que las palabras impresas se procesan, al menos inicialmente, por la agenda visoespacial).

En tercer lugar, sólo las medidas on-line y las presentaciones secuenciales son una verdadera prueba para el principio. Cuando nos son presentadas animaciones (que tienen un montón de *frames* que se suceden sin parar haciendo a las cosas cambiar de posición, forma y color y haciéndolas aparecer y desaparecer) y al mismo tiempo tenemos que procesar información verbal presentada visualmente, nos enfrentamos a un problema perceptivo, no de memoria (sobre esto insiste Tabbers (2002) también). No es que se consuman los recursos de la agenda dejando pocos para la animación, ¡es que uno no puede mirar a dos puntos al mismo tiempo! Así que los experimentos que usan animaciones simultáneas a lo verbal no son prueba válida para el principio de modalidad (que, recordemos, nos habla sobre un problema de memoria de trabajo, no de percepción). Ahora, los experimentos en los que la presentación de lo pictórico y lo verbal se sucede secuencialmente sí son pruebas válidas para el principio: en ellos no se confunde el problema perceptivo con el de memoria. Algunas medidas on-line funcionan también en este sentido. Una dual-task puede revelar que, en efecto, el problema es de memoria y no (sólo) perceptivo. Si al procesar imágenes e información verbal disponemos de pocos recursos en la agenda visoespacial es porque lo verbal está absorbiendo capacidad de procesamiento de ésta. No obstante, hemos de decir que los experimentos de Brünken et al. (2002) tienen un problema. La tarea secundaria que usan también plantea un problema perceptivo (no de memoria) a la condición que presenta lo verbal vía visual. Una vez más, quienes procesan la información verbal en la pantalla no pueden dirigir su mirada arriba para ver si la “A” cambia de color. ¡No pueden mirar!, no es que su agenda visoespacial esté saturada (bueno, puede estarlo pero esa no es forma de averiguarlo). La tarea secundaria que serviría es una en la que la agenda visoespacial se pone en juego sin comprometer la capacidad perceptiva de los sujetos. Por ejemplo, si los sujetos tuviesen que rotar figuras mentalmente o cualquier otra tarea que implique razonamiento pictórico, la agenda estaría en juego y la percepción se dedicaría a la tarea primaria. En definitiva, aunque algunas demostraciones del principio de modalidad no son todo lo estrictas que deberían, seguimos contando con pruebas (las de estudios con presentaciones secuenciales) que lo avalan. Y lo que sigue también es un apoyo para esta conclusión.

Cuarto, lo que a veces se considera presentación simultánea es en realidad secuencial. En algunos experimentos se presentan ilustraciones (que son imágenes estáticas, *sin* decenas de frames que se suceden) y, al tiempo, la información verbal. Se considera que la presentación es simultánea pero al ser una imagen fija nunca sabremos el tipo de procesamiento que realiza el aprendiz. Puede dedicar un rato a leer/escuchar la información verbal (en función de si se presenta visual o auditivamente) para después atender a algún aspecto de la imagen; después puede repetir el movimiento y así sucesivamente. Si el aprendiz alterna la consulta de imágenes y palabras la presentación *se vuelve*, de hecho, secuencial. Esto significa que muchos de los experimentos en los que hicimos constar “presentación simultánea” son en realidad experimentos con presentaciones secuenciales. Mira por dónde esto supone un apoyo adicional para el principio. Hace un instante decíamos que sólo lo secuencial es prueba estricta del principio: ahora vemos que hay más secuenciales de lo que parece.

El balance final es positivo. El principio de modalidad ha sido confirmado un número más que suficiente de veces. Está documentado que algunos factores modulan su efecto pero esto no es sino la excepción que confirma la regla. Hemos señalado algunos problemas que, en todo caso, el principio ha sabido encajar. De modo que suscribimos el principio: allí donde se presentan palabras e imágenes de cuya conjunción hay que aprender, usa la modalidad auditiva para lo verbal.

4.2.1.5. ¿Por qué no nos sirve el principio de modalidad para diseñar ayudas (verbales)? Si hay que procesar conjuntamente imágenes y palabras el principio de modalidad especifica qué modalidad emplear para las éstas. Pero ¿qué pasa si no hay imágenes que integrar con las palabras? ¿Qué pasa si se presentan palabras sin más? ¿Qué ocurre si se presentan palabras e imágenes pero éstas no deben integrarse con aquéllas? Pues ocurre que el principio de modalidad no nos ofrece pautas.

Resulta que las ayudas verbales (en la gran mayoría de los casos) no llevan asociadas imágenes. Nótese que decimos asociadas, es decir, imágenes que hay que integrar con la información contenida en las ayudas. Sirva como demostración que ni en las ayudas de Atkinson et al. (2003), ni en las de Gerjets et al. (2006), ni en las de VanLehn et al. (2007), ni en las de Nückles et al. (2005), y son sólo unos ejemplos, hubo imágenes que procesar conjuntamente con las ayudas. Si las palabras de la ayuda no deben integrarse con ninguna imagen, el principio no puede determinar qué modalidad saldrá ganadora. Luego el principio no nos sirve para contestar a nuestra pregunta de cómo diseñar ayudas.

4.2.2. Los resultados divergentes. Hasta la fecha son pocos los estudios que han explorado el impacto de la modalidad de las ayudas (verbales) sobre el aprendizaje. Que sepamos, sólo cuatro estudios han abordado el tema. Un problema es que son pocos. El otro, más serio, es que no hay un patrón claro de resultados.

Moreno y sus asociados (Moreno & Mayer, 2002; Moreno, Mayer, Spires & Lester, 2001) llevaron a cabo varios experimentos contrastando ayudas (verbales) auditivas y visuales. En los experimentos los participantes debían aprender botánica a partir de un CBLE.

El CBLE presentaba animaciones, con narración incorporada, que describían las condiciones de diferentes ambientes (e.g., “este ambiente tiene escasa lluvia y poca luz solar”). Terminada la descripción, se pedía a los participantes que eligiesen las raíces, tallos y hojas que considerasen oportunos para que una planta sobreviviese en el ambiente descrito. Se presentaban ocho alternativas para elección. El CBLE proporcionaba feedback explicativo sobre las elecciones de los participantes (e.g., “En este ambiente el tallo debería ser suficientemente largo como para llevar las hojas hasta la luz solar”). El feedback servía para reparar los problemas en las respuestas de los participantes. En cuatro experimentos (Moreno & Mayer, 2002, Exp. 1 & 2; Moreno et al., 2001, Exp. 4 & 5) los investigadores manipularon la modalidad de presentación del feedback explicativo. Tras utilizar el CBLE los participantes resolvían tareas de recuerdo y transferencia. En los cuatro experimentos el feedback explicativo auditivo fue mejor que el mismo feedback presentado visualmente; y esto ocurrió tanto en las medidas de recuerdo como en las de transferencia. O sea, ayuda auditiva mejor que visual.

Atkinson (2002) implementó un estudio muy similar. En esta ocasión los participantes aprendieron cálculo de probabilidades a partir de un CBLE. El CBLE contenía ejercicios resueltos que iba presentando sucesivamente. Los ejercicios, a su vez, contenían información verbal (el enunciado) y matemática (los pasos para resolverlo). Los participantes recibieron ayudas para reparar posibles malentendidos. Estas ayudas eran elaboraciones que revelaban el razonamiento subyacente a los pasos de la resolución. Atkinson manipuló la modalidad en que las elaboraciones fueron presentadas. Después de utilizar el CBLE los participantes resolvieron ejercicios similares y distintos a los consultados en la fase de estudio. En el Experimento 1 quienes recibieron elaboraciones auditivas rindieron mejor en los ejercicios similares con respecto a quienes las recibieron en modalidad visual. No hubo diferencias en los ejercicios distintos. En el Experimento 2 no hubo diferencias ni en una ni en otra variable. En fin, ayuda auditiva igual (o mejor) que visual.

Graesser, K. Moreno, Marineau, Adcock, Olney y Person (2003) también tienen en su haber un experimento parecido a los anteriores. Sus participantes hubieron de aprender informática a partir de un CBLE. El CBLE presentaba diagramas (incluyendo pequeñas narraciones introductorias simultáneas) que describían configuraciones informáticas (e.g., “un ordenador conectado con otro a través de un módem y un servidor”). El CBLE planteaba alguna cuestión compleja al participante, para la cual éste debía encontrar una respuesta satisfactoria. El participante tecleaba su respuesta y el CBLE ofrecía ayudas diversas. Turno a turno la respuesta se iba mejorando hasta completarse. Las ayudas que el CBLE ofrecía eran de distintos tipos. Por un lado, había ayudas destinadas a detectar problemas en las contribuciones del participante: open prompts forzando a los participantes a revisar sus representaciones (e.g., “dime más”, “¿algo que añadir?”) y error messages identificando elementos erróneos en ellas (e.g., “casi”, “te falta un poco”). Por otro lado, había ayudas orientadas a la reparación de los problemas detectados: en forma de hints y fill-in-the blanks (e.g., “¿y qué hay del disco duro?”, “las memorias principales de una CPU son la RAM y ___”) o de complejas elaboraciones (e.g., “El CD-Rom es otro recurso de almacenamiento... ”). en un experimento Graesser et al. manipularon la modalidad de presentación de las ayudas. Después de utilizar el CBLE los participantes resolvieron varias tareas: preguntas de recuerdo y de transferencia. Tanto en unas como en otras, la modalidad resultó ser un factor no

significativo. O sea, ayuda auditiva igual que visual.

4.2.3. Conclusión. La conclusión es que no hay conclusión. Tenemos datos a favor de la modalidad auditiva y datos a favor de la equivalencia entre modalidades. Las evidencias empíricas no ofrecen pues un patrón claro y, en virtud de ello, no cabe extraer pautas de diseño.

4.3. Un marco de investigación: la aproximación funcional.

Hemos comprobado hace un instante que aún no sabemos cómo debe utilizarse la modalidad en el diseño de ayudas por dos razones. Una, el principio de modalidad no es aplicable. Dos, los resultados de la poca investigación realizada hasta la fecha son discrepantes. Lo que necesitamos entonces es un marco de investigación que permita interpretar esas discrepancias empíricas y que proponga argumentos teóricos desde los que extraer pautas de diseño. Como respuesta a tal necesidad planteamos la *aproximación funcional*. La aproximación funcional está basada en tres supuestos.

4.3.1. Primer supuesto: hay distintas ayudas dependiendo de la función lingüística que cumplen. De acuerdo con la función que cumplen, distinguimos dos clases de ayudas: las *regulatorias* y las *explicativas*. Estas dos clases de ayudas difieren además en el modo como el aprendiz debe procesarlas y en el proceso de aprendizaje que facilitan. Detallamos todo esto a continuación.

4.3.1.1. Las ayudas regulatorias. Las ayudas regulatorias desempeñan una función manipulativa o regulatoria (Givón, 1984; Werner, 1988a). Esto significa que solicitan al aprendiz llevar a cabo una acción, en concreto, que revisen su comprensión en busca de malentendidos (e.g., simplificaciones, distorsiones). Tal solicitud obliga al aprendiz a decidir si satisfacerla o no. Las ayudas regulatorias son ayudas al proceso de detección, puesto que asisten al aprendiz en la identificación de problemas de comprensión.

4.3.1.2. Las ayudas explicativas. Las ayudas explicativas cumplen una función informativa o declarativa (Givón, 1984; Werner, 1988a). Esto es, transmiten información al aprendiz. Concretamente, proporcionan una elaboración sobre los contenidos del CBLE que han sido malentendidos (e.g., distorsionados) por el aprendiz. El aprendiz, por tanto, debe procesar e integrar tal elaboración. Las ayudas explicativas son ayudas al proceso de reparación, dado que asisten al aprendiz en la reconstrucción de sus representaciones mentales defectuosas.

4.3.1.3. Las ayudas en funcionamiento. Imaginemos que un aprendiz estudia geología a partir de un CBLE que incluye animaciones y narraciones simultáneas. En un momento dado, el CBLE le plantea una cuestión al aprendiz sobre un aspecto que (a tenor de lo sucedido en anteriores experimentos) suele ser malentendido.

- “Hemos visto que en la subducción se destruye corteza terrestre, ¿por qué no se consume hasta desaparecer? Necesariamente ha de haber un lugar en el que se cree corteza para compensar las cosas, ¿cuál?”

- (a) el magma de los volcanes compensa la pérdida
- (b) en las fosas de subducción se crea corteza
- (c) en las dorsales oceánicas se crea corteza”

Supongamos que, puesto que éste es un aspecto que suele malentenderse, el aprendiz selecciona la respuesta (a), que no es correcta. Entonces el CBLE responde así:

- (I) “Es cierto que el magma que aflora de los volcanes se solidifica creando nueva corteza pero, no, ésta no es la respuesta correcta. El magma de los volcanes es insuficiente ya que la actividad de éstos es muy irregular.”

Y luego el CBLE agrega:

- (II) “La nueva corteza se crea en las dorsales. Las dorsales eran grandes fisuras a lo largo de la corteza por las que emergía magma. El magma se iba solidificando y constituyendo así nueva corteza. Por tanto, hay un proceso de reciclado continuo de la corteza: se destruye en la subducción y se crea en las dorsales.”

La ayuda (I) es regulatoria en tanto que indirectamente le pide al aprendiz que revise su comprensión ya que algo no va bien. El aprendiz, por su parte, habrá de sopesar si merece la pena o no hacer lo que se le pide. La ayuda (II) es explicativa. Lo es porque proporciona una elaboración sobre los contenidos presentados en el CBLE que el aprendiz debe procesar detenidamente para reparar su modelo mental defectuoso.

4.3.2. Segundo supuesto: cada modalidad tiene características específicas, las cuales son potencialmente ventajosas. Cada modalidad del lenguaje presenta una característica específica. Es decir, cada cual tiene una propiedad que es inherente a esa modalidad (no depende de factores coyunturales) y exclusiva (la otra modalidad no puede disponer de la propiedad de su contraria).

4.3.2.1. El control visual. La modalidad visual proporciona control al aprendiz. Esto significa que el aprendiz puede (a) releer, (b) regular el ritmo de lectura y (c) enfocar en cada momento el segmento (e.g., palabra, oración) de su interés. Hay evidencias de que el control puede beneficiar la adquisición de nuevas ideas y relaciones. Por ejemplo, cuando se da la oportunidad de releer los aprendices rinden mejor en tareas de aprendizaje que quienes leen un sola vez (Rawson & Kintsch, 2005). Cuando se impide volver a consultar un segmento anterior la comprensión se resiente (Just et al., 1982). Cuando el tiempo de consulta de las ideas se restringe es más difícil retenerlas (Kintsch & Keenan, 1973). En resumen, la modalidad visual proporciona control y el control facilita la adquisición de nuevas ideas y relaciones.

4.3.2.2. La expresividad auditiva. La modalidad auditiva presenta expresividad (i.e., prosodia). A través de recursos como el énfasis, el fraseo o el ritmo la expresividad encarna la intención comunicativa del hablante y su actitud hacia lo que dice (e.g., si lo considera urgente o importante). Está documentado que nos apoyamos en la prosodia para inferir tales cosas (e.g., Baum & Nowicki, 1998; Brennan & Williams, 1995; Haskard et al., 2008).

Existen evidencias de que el tener acceso a la actitud de otros puede afectar a la disposición que mostramos a realizar cierta tarea (Deci, Eghrari, Patrick & Leone, 1994; Reeve et al., 2002; Wild, Enze, Nix & Deci, 1997). Por ejemplo, Reeve et al. (2002) observaron que cuando un instructor de chino demostraba considerar importante aprenderlo (e.g., “es muy útil aprender chino”, “es algo muy valorado”) los aprendices manifestaban tener una mayor disposición a aprenderlo. Wild et al. (1997) encontraron que cuando un profesor de magia demostraba estar intrínsecamente motivado para enseñar (e.g., “me interesa mucho enseñaros”) sus aprendices manifestaban mayor disposición a aprender. En resumen, la modalidad auditiva exhibe expresividad, la cual encarna la actitud del hablante, y las actitudes de otros hacia una tarea favorecen nuestra disposición a realizarla.

4.3.3. Tercer supuesto: existe una interacción función x modalidad. Proponemos que existe una interacción entre función y modalidad. Las ayudas explicativas deberían beneficiarse del control mientras que las ayudas regulatorias deberían beneficiarse de la expresividad. Presentamos los argumentos que subyacen a estas afirmaciones.

Por una parte, quien recibe una ayuda explicativa debe procesar e integrar ciertos contenidos. Por otra parte, sabemos que el control favorece la adquisición de nuevas ideas y relaciones. Por lo tanto, el control debería facilitar el procesamiento de las ayudas explicativas. El control haría posible a los aprendices releer, regular el ritmo de lectura y enfocar en cada momento el segmento de interés, algo útil cuando se trata de reparar sus modelos mentales defectuosos. Quien recibe una ayuda regulatoria debe decidir si realizar o no lo que se le pide. El control afecta al modo como incorporamos nuevas ideas y relaciones pero no a nuestra disposición a hacer algo. En consecuencia, el control no debería facilitar el procesamiento de las ayudas regulatorias.

Por un lado, quien recibe una ayuda regulatoria debe sopesar si satisfacer o no lo que se le demanda. Por otro lado, sabemos que el acceso a la actitud de otros afecta positivamente a nuestra disposición a hacer algo. Por consiguiente, la expresividad (que da acceso a la actitud de otro) debería facilitar el procesamiento de las ayudas regulatorias. La expresividad daría acceso a los aprendices a la actitud que quien les demanda hacer algo alberga sobre ese algo; ello sería útil cuando se trata de decidir si realizarlo o no. Quien recibe una ayuda explicativa debe procesar e integrar ciertos contenidos. La expresividad afecta a la disposición que tenemos a hacer algo pero no al modo como incorporamos nuevas ideas y relaciones. En consecuencia, la expresividad no debería facilitar el procesamiento de las ayudas explicativas.

4.3.3.1. Predicciones. De este tercer supuesto se desprenden dos predicciones. Primero, las ayudas explicativas visuales son mejores que las ayudas explicativas auditivas. Segundo, las ayudas regulatorias auditivas son mejores que las ayudas regulatorias visuales. Estas dos predicciones son las que se pondrán a prueba en los experimentos. Confirmarlas supondría respaldar la aproximación funcional y obtener las pautas de diseño de ayudas que andamos buscando.

4.4. ¿Encajan los resultados anteriores en la aproximación funcional?

La aproximación funcional hace predicciones precisas sobre el comportamiento de

cada modalidad. Específicamente, predice que las ayudas explicativas visuales son mejores que idénticas ayudas presentadas auditivamente y que las ayudas regulatorias son mejores que idénticas ayudas presentadas visualmente. Ahora la pregunta es qué hay de los resultados anteriores sobre el impacto de la modalidad de las ayudas sobre el aprendizaje (Atkinson, 2002; Graesser et al., 2003; Moreno & Mayer, 2002; Moreno et al., 2001), ¿encajan esos resultados en la aproximación funcional?

En los experimentos de Moreno y sus asociados (Moreno & Mayer, 2002; Moreno et al., 2001) los participantes aprendían botánica y recibían ayudas para reconstruir sus modelos mentales defectuosos, es decir, recibían ayudas explicativas. Las ayudas fueron presentadas bien en la modalidad visual o bien en la auditiva. A la luz del marco de investigación que nosotros proponemos, las ayudas visuales deberían haber sido mejores que las auditivas. Sin embargo, no fue así. De hecho, quienes recibieron ayudas auditivas superaron a quienes las recibieron en modalidad visual. De modo que estos resultados no encajan en la aproximación funcional. Una manera de explicar esta inconsistencia es asumir que las ayudas explicativas de Moreno y asociados contenían elaboraciones muy simples. Simples porque incluían pocas ideas y relaciones. Esto es lo que cabe concluir tras un análisis de las ayudas explicativas de Moreno et al. (véase el Appendix (A)).

Esta explicación es consistente con la investigación en comprensión de textos y obliga a puntualizar una de las predicciones del marco propuesto. En la investigación pasada se ha encontrado que cuando los textos a aprender son simples, la modalidad en la que son presentados es irrelevante: visual es igual que auditiva (Gernsbacher et al., 1990; Kintsch & Kozminsky, 1977; Kintsch et al., 1975; Smiley, Oakley, Worthen, Campione & Brown, 1977). Ahora bien, cuando el texto a aprender es complejo, la modalidad se vuelve decisiva: lo visual es mejor que lo auditivo (Green, 1981). Evidencias como éstas reclaman un ajuste de la primera predicción que planteamos: las ayudas explicativas *complejas* son mejores si se presentan en modalidad visual que si se presentan en modalidad auditiva. Con el ajuste pertinente, el marco permite explicar los resultados de Moreno y asociados.

Y ¿qué hay de los resultados de Atkinson (2002)? Puede esgrimirse el mismo argumento. En sus experimentos los participantes aprendían a resolver problemas de probabilidad a partir de un CBLE. El CBLE incluía ayudas para reconstruir las representaciones incompletas y/o incorrectas de los participantes, es decir, ayudas explicativas. Por supuesto, las ayudas se presentaron bien en una modalidad, bien en la otra. La aproximación funcional habría predicho un beneficio para los que las recibieron en modalidad visual. ¿Qué sucedió? Quienes recibieron las ayudas explicativas visuales aprendieron tanto como los que las recibieron auditivamente. Luego de nuevo hay una inconsistencia. Para resolverla basta examinar de cerca las ayudas explicativas empleadas por Atkinson (cf. Appendix (A)). De nuevo se aprecia que contienen elaboraciones muy simples, esto es, que incluyen pocas ideas y relaciones.

Veremos ahora si el resultado de Graesser et al. (2003) encaja en la aproximación funcional. Sus participantes aprendieron informática a partir de un CBLE con ayudas tanto para identificar problemas en sus representaciones mentales (regulatorias) como para repararlos (explicativas). Las ayudas fueron presentadas en modalidad visual, en unos casos, y

en modalidad auditiva, en otros. La aproximación funcional habría predicho un beneficio de las regulatorias auditivas sobre las visuales y un beneficio de las explicativas visuales sobre las auditivas. Lo que revelaron los resultados es que no hubo diferencias entre las condiciones. Esto *no* es inconsistente con el marco de investigación que proponemos. No lo es porque Graesser et al. no distinguieron tipos de ayuda y la modalidad fue manipulada para las ayudas en bloque: todas visuales o todas auditivas. Esto impidió explorar la posible interacción función x modalidad que la aproximación funcional plantea. Así que no podemos determinar si los resultados de Graesser y asociados encajan o no en nuestro marco.

En suma, los datos de la investigación pasada encajan en la aproximación funcional. Moreno et al. y Atkinson no pudieron encontrar ventaja de las ayudas explicativas visuales sobre las auditivas. Eso lo explicamos alegando que la complejidad de sus ayudas explicativas fue mínima y la interacción modalidad visual-ayuda explicativa, predecimos, se da sólo cuando la ayuda es compleja. Tal interpretación es consistente con los resultados en la investigación sobre comprensión de textos y supone un ajuste del marco que proponíamos. Los resultados en Graesser et al. no pueden considerarse ni consistentes ni inconsistentes con la aproximación funcional. Los investigadores no mantuvieron bajo control el tipo de ayuda utilizada, lo que hizo imposible explorar la interacción función x modalidad.

4.5. Panorámica de los experimentos que siguen.

Llevamos a cabo una serie de experimentos en los que los participantes aprendieron geología (i.e., la tectónica de placas) a partir de un CBLE. El CBLE incluía animaciones acompañadas de narraciones. Adicionalmente, el CBLE incorporaba ayudas tanto regulatorias como explicativas que incidían sobre los aspectos que los aprendices malentienden. Cada tipo de ayuda se presentó bien en modalidad visual bien en modalidad auditiva para generar un diseño 2 x 2: regulatoria/explicativa x auditiva/visual. Una vez consultado el CBLE los participantes debían resolver tareas de recuerdo y de transferencia, cuyas puntuaciones sirvieron como variables dependientes.

Esperábamos que quienes recibían las ayudas regulatorias auditivas rindiesen mejor que quienes recibían las mismas ayudas en modalidad visual. Esperábamos que quienes recibían ayudas explicativas visuales rindiesen mejor que quienes recibían idénticas ayudas presentadas en modalidad auditiva. Esperábamos el mismo resultado en las tareas de recuerdo, que exigían a los participantes retener la información clave presentada en el CBLE, y en las de transferencia, que exigían aplicar los conocimientos adquiridos para resolver problemas.

Un asunto importante es que preguntarse qué modalidad es mejor para presentar ciertas ayudas tiene sentido si y sólo si esas ayudas son eficaces. Por eso en los experimentos incorporamos condiciones de control (sin ayudas). Una vez hechas las comparaciones preliminares, las que nos garantizan que las ayudas funcionan, es cuando contrastamos el efecto de su modalidad de presentación.

Nota: Lo que sigue está redactado en inglés. Se trata de un requisito para obtener el *Doctorado Europeo*. Concretamente, El requisito segundo (Artículo 14, RD 56/2005) establece que “parte de la tesis doctoral, al menos el resumen y las conclusiones debe redactarse en una de las lenguas oficiales de la UE distinta a alguna de las oficiales en España”.

Chapter 4. Experiments²⁸.

0. Abstract.

Because learning from computer-based learning environments can be difficult for learners, they usually include some kind of verbal aid. When constructing these aids, designers must decide which modality is best to use for their presentation, the auditory or the visual modality. The problem is that there are not clear guidelines in this respect for two reasons: the modality principle is not applicable and the little empirical research on the question provides diverging results. A research framework was proposed and tested empirically. Based on several arguments, the framework provides guidelines about how to use modality in the design of aids. Specifically, it was predicted that aids to the monitoring process (regulatory aids) should be presented auditorily rather than visually whereas aids to the revision process (explanatory aids) should be presented visually rather than auditorily. In four experiments the participants learned geology from a computer-based learning environment including either auditory or visual aids to the monitoring and revision processes and solved learning tasks afterwards. In the Experiments 1 and 2 the results showed that the participants receiving visual aids for the revision process outperformed those receiving identical aids presented auditorily. However, the participants receiving auditory aids to the monitoring process showed no advantage over their counterparts, even when a pedagogical agent was included. In the Experiments 3 and 4 the complexity of the learning materials was increased to make the aids to the monitoring process more pertinent and, in so doing, make the impact of their modality more substantial. In the Experiment 4 it was found that, for learners who have low levels of self-regulation and literacy, the auditory aids to monitoring provided by a pedagogical agent were better than the visual aids in enhancing performance on the learning tasks. The results have implications for the design of computer-based learning environments. Theoretical implications will also be discussed.

Keywords: computer-based learning environments; verbal aids; visual modality; auditory modality; monitoring and revision processes.

28 Based on two articles:

(i) Sánchez, E. & García-Rodicio, H. (2008). The use of modality in the design of verbal aids in computer-based learning environments. *Interacting with Computers*, 20, 545-561.

(ii) García-Rodicio, H & Sánchez, E. (submitted for publication). How should designers present aids to the monitoring process? Evidence for the superiority of the auditory over the visual modality. *International Journal of Human Computer Studies*.

1. Introduction.

Computer-based learning environments (CBLEs henceforth) pose learners with some challenges, such as having to monitor their understanding and repair the problems they detect when doing so. That is the reason why CBLEs usually include some kind of aid. Because CBLEs are very flexible, these aids can be designed in multiple ways. One of the questions designers have to solve is which modality is best for the presentation of verbal aids, the visual or the auditory modality. As will be shown, there are not clear guidelines in this respect. Therefore, the main goal of this study was to present and test a research framework, which provides guidelines about how to present verbal aids regarding their modality.

1.1. Aids to monitor and revise misunderstandings in CBLEs.

Provided that learning from CBLEs has difficulties for learners, the use of any kind of support or aid is warranted. CBLEs demand learners to deal with complex topics such as biology (Azevedo & Cromley, 2004) or geology (Graesser, Wiley, Goldman, O'Reilly, Jeon, & McDaniel, 2007), to integrate verbal and pictorial information (Ainsworth, Bibby, & Wood, 2002), to develop learning paths through a hyperspace (Chen, Fan, & Macredie, 2006), and to overcome the limitations of the working memory (Sweller, van Merriënboer, & Paas, 1998).

All these circumstances can make learners to construct poor mental representations. In order to not do so, learners have (a) to monitor their understanding to detect flaws in their mental representations and (b) to generate explanations to revise and repair them. These are the *monitoring* and the *revision* processes, respectively, which can be placed in the broad category of self-regulation skills (Chi, 2000; deLeeuw & Chi, 2002; Hacker, 1998; Otero, 2002; Thiede & Dunlosky, 1999; Winne, 1995). As can be noted, these processes are critical in order for learners to reach a deep understanding from CBLEs.

However, once again executing these processes is difficult for many learners (e.g., Azevedo, Guthrie, & Seibert, 2004; Chi, Bassok, Lewis, Reiman, & Glaser, 1989; Pirolli & Recker, 1994). On the one hand, learners usually fail in executing the monitoring process: they have trouble when having to estimate their level of understanding and to notice inconsistencies between two propositions (Commander & Stanwyck, 1997; Maki, Foley, Kajer, Thompson & Willer, 1990; Markman, 1979; Otero & Campanario, 1990; Otero & Kintsch, 1992; Scommer & Surber, 1986). On the other hand, learners usually fail in executing the revision process: they have trouble when having to generate explanations or to use available information to repair the problems of understanding (Cataldo & Oakhill, 2000; Coté, Goldman, & Saul, 1998; Kintsch & Kintsch, 1995; Otero & Campanario, 1990; Renkl, 1997, 2002). For all these reasons, aids are indispensable.

Aids are, then, devices helping learners to overcome those difficulties. Thus, aids can - among other things- help learners to monitor their problems of understanding, to revise them, or both. For instance, Moreno and Mayer (2005) provided the participants with corrective feedback helping them to detect flaws in their mental representations (e.g., “No, a short stem

here is dangerous for the plant.”) while learning botany. In an earlier experiment, Moreno (2004) provided the participants with explanatory feedback, which help them to revise and repair the gaps and flaws in their mental representations (e.g., “For this environment plants should have long stems...”). Many experiments have shown that the aids to the monitoring and revision processes are indeed effective (e.g., Alevén & Koedinger, 2002; Atkinson, Renkl, & Merrill, 2003; Berthold, Eysink, & Renkl, 2008; Conati & VanLehn, 2000; Graesser, Lu, Jackson, Mitchell, Ventura, Olney, & Louwerse, 2004; Moreno, 2004; Moreno & Mayer, 2005; Sánchez, García-Rodicio, & Acuña, 2008; Schworm & Renkl, 2006; VanLehn, Graesser, Jackson, Jordan, Olney, & Rosé, 2007).

1.2. Designing aids in CBLEs: the use of modality when constructing verbal aids.

Because CBLEs are flexible, aids can be designed in multiple ways. First, aids can be pictorial or verbal. Thus, aids can be arrows directing learners' attention to relevant aspects of an animation (e.g., Mautone & Mayer, 2001) or feedback messages assisting learners in monitoring their understanding, such as those mentioned before (e.g., Moreno & Mayer, 2005). Second, verbal aids can be presented either in the auditory or the visual modality. One key question is which of these modalities is best for the presentation of verbal aids. This is the main concern of this study.

With respect to the impact of the modality of verbal aids on learning, the research available is little and has some problems. The first problem is that the *modality principle* (Mayer, 2001; Sweller et al., 1998) is not applicable for verbal aids. The modality principle suggests that verbal contents accompanying pictures should be presented auditorily rather than visually. In so doing, (a) learners are prevented from suffering working memory overload, which is due to the processing of verbal and pictorial information by the same (visual) system, and (b) the effective size of their working memory is expanded, as its two processors are used (Moreno & Mayer, 1999; Mousavi, Low, & Sweller, 1995). The point is that for this to happen, words and pictures have to be presented together (either sequentially or simultaneously) but verbal aids do not (necessarily) have related pictures. Moreover, contents accompanying pictures accomplish one function (i.e., describe the events shown by the pictures) but verbal aids can accomplish different functions (e.g., assist learners in monitoring their understanding). For these two reasons the modality principle is not applicable for the design of verbal aids.

The second problem is that the research available on the impact of the modality of verbal aids provides diverging results. To our knowledge only three sets of experiments have been carried out so far. Moreno and her colleagues (Moreno & Mayer, 2002; Moreno, Mayer, Spires, & Lester, 2001) incorporated either auditory or visual explanatory feedback (i.e., aids to revision) into a CBLE teaching botany. The participants in the auditory feedback condition performed better than those receiving visual feedback. Atkinson (2002) provided explanatory aids helping learners to revise their understanding of the rationale behind the solution steps of worked-out examples. These aids were presented in either the visual or the auditory modality. In most of the comparisons, there were no differences between the modalities. Graesser, K. Moreno, Marineau, Adcock, Olney, and Person (2003) incorporated several tutoring aids into a CBLE teaching computer literacy. These were aids to both monitoring and revision and

were presented either visually or auditorily. The results showed that there were no differences between the conditions. Overall, the research is little and there are mixed results regarding the impact of the modality of verbal aids on learning.

In the light of these problems, a conclusion can be drawn: there are not clear guidelines about how to use modality in the design of verbal aids. Hence, it seems reasonable to construct a research framework able to explain those diverging results and to formulate testable predictions. This is what is presented next.

1.3. The functional approach: a research framework.

The functional approach is the research framework we propose. The approach is based on three assumptions.

1.3.1. First assumption. The first assumption is that each modality has specific advantages. The visual modality allows readers to *control* the text. This means that they can regulate the pace of reading, re-read, or focus on the segment of the text they want. These circumstances can be beneficial in learning contents (Just, Carpenter, & Woolley, 1982; Kintsch & Keenan, 1973; Rawson & Kintsch, 2005). For instance, Rawson and Kintsch (2005) showed that when learners reread a text learn more than when only one reading is allowed. Kintsch and Keenan (1973) found that when time is constrained (that is, when readers cannot regulate the pace of reading) readers find it difficult to recall the ideas in texts. In sum, the visual modality provides control whereas control is useful in acquiring concepts from an instructional material.

The auditory modality, in turn, has *expressiveness*. The expressiveness can reflect or embody the communicative intention of the speaker (i.e., illocutionary force such as stating or requesting) and his/her attitude towards the message (e.g., whether he/she considers important what he/she is saying). There is evidence that people rely on expressiveness to notice speakers' intentions and attitudes (Baum & Nowicki, 1998; Brennan & Williams, 1995; Haskard, S. Williams, DiMatteo, Heritage, & Rosenthal, 2008). For instance, Brennan and Williams (1995) asked their participants to rate, on the basis of the expressiveness some individuals showed, the confidence these individuals had in answering to general-knowledge questions (e.g., "Who was the first president of the USA?"); the researchers found a positive and significant correlation between the confidence the respondents reported and that inferred by the participants. Moreover, there is evidence that knowing the attitude of others affects the attitude one has towards some task: if their attitude is positive we are more willing to perform it (Deci, Eghrari, Patrick, & Leone, 1994; Reeve, Jang, Hardre, & Omura, 2002; Wild, Enze, Nix, & Deci, 1997). Wild et al. (1997) found that when instructors provided evidence of being intrinsically motivated their learners in such a way they were more willing to perform a task, with respect to the learners provided with evidence of an extrinsically motivated instructor. In sum, the auditory modality has expressiveness, which embodies others' attitudes, whereas others' attitudes make us willing to perform a given task.

1.3.2. Second assumption. There are different types of verbal aid. They differ on (a) the learning process they facilitate (i.e., their function), (b) the illocutionary act they perform,

and (c) the way learners process them. Thus, the *regulatory aids* are those (a) helping learners to monitor their understanding by looking for gaps and flaws in their mental representations. These aids consist of (b) requests asking (either directly or indirectly) learners to do something, i.e., revise their understanding. Hence, when receiving regulatory aids learners have to (c) decide whether or not it is worthy to do what the aid is asking them to do (revise their understanding) and then modify their behavior accordingly. The *explanatory aids*, in turn, are those (a) helping learners to revise and repair their misunderstandings. These aids are (b) sets of assertions involving elaborations on the to-be-learned contents. When receiving explanatory aids learners have to (c) process carefully the elaborations.

Let us see how regulatory and explanatory aids work. Imagine some learners are asked to learn geology from a CBLE. During the use of the CBLE, a question is posed: “The crust is destroyed in the subduction process, why does it remain the same size?” The learner answers that “volcanoes are the solution” and the CBLE reacts this way: (1) “...No, the magma from the volcanoes cannot compensate for the loss, as it is very little...” And then: (2) “The new crust is actually created in the ridges, as the magma surfacing through them solidifies forming new plate material. This means there is a recycling loop...” The former reaction (1) of the CBLE can be considered a regulatory aid, as it helps the learner to detect flaws in his/her mental representation and (indirectly) asks him/her to revise it. The latter (2) can be considered explanatory since it provides an elaboration fixing the flaws in her mental representation and expects her to process it deeply.

1.3.3. Third assumption. There is an interaction between modality and type of verbal aid. On the one hand, learners processing an elaboration on the to-be-learned contents might find it useful to control the text. This control would allow them to process the elaboration as deeply as it is required (e.g., reading at a slow pace) to repair their understanding. Recall that control was beneficial for learners acquiring concepts from instructional materials. Hence, there may be a relation between the visual modality and the explanatory aids: the visual explanatory aids should be better than the auditory explanatory aids.

On the other hand, learners receiving demands to revise their understanding might find it useful to know speaker's intention (i.e., request) and attitude (i.e., he/she considers important what he/she is saying) on the basis of his/her expressiveness. This expressiveness would allow learners to realize they are asked to revise their mental representations and the importance of doing so and, as a consequence, doing it promptly. After all, others' attitudes affect how motivated we are towards a given task. Hence, there may be a relation between the auditory modality and the regulatory aids: the auditory regulatory aids should be better than the visual regulatory aids.

These predicted relations were called the *visual explanation superiority hypothesis* and the *auditory regulation superiority hypothesis*, respectively. These predictions were tested empirically by means of four experiments, as will be shown.

1.4. Prior mixed results fit into the functional approach.

The functional approach we have just presented predicts visual explanatory aids to outperform auditory explanatory aids and auditory regulatory aids to outperform visual regulatory aids. Does the scenario in prior research (Atkinson, 2002; Graesser et al., 2003; Moreno & Mayer, 2002; Moreno et al., 2001) fit into the functional approach?

In the experiments of Moreno and her associates (Moreno & Mayer, 2002; Moreno et al., 2001) the participants learned botany and received aids to revise their understanding, in other words, explanatory aids. Specifically, the participants were provided with visual or auditory aids in the form of explanatory feedback. It follows from our research framework that visual explanatory aids should be better than the auditory ones but this was not the case: actually, Moreno et al. found that the auditory modality was better than the visual one in all the learning measures. So this does not fit into the predictions of the functional approach. A careful examination of the explanatory feedback in Moreno's experiments may sort out the problem: the explanations were simple (as they comprised few ideas and relations between them; see Appendix (A) for an analysis). One may explain the visual non-advantage in Moreno et al.'s studies as the result of using simple explanations, which made the control available in the visual modality not so helpful.

This interpretation is in line with with prior research on text comprehension. The auditory and the visual modalities make no difference when readers are provided with simple (explanatory) texts to comprehend (Gernsbacher, Varner, & Faust, 1990; Kintsch & Kozminsky, 1977; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Smiley, Oakley, Worthen, Campione, & Brown, 1977). But they do when the (explanatory) texts are complex, as found in the experiment of Green (1981): in these cases, the visual modality is better than the auditory one. Accordingly, if visual explanatory aids are to be better than identical aids presented auditorily, they have to be complex. This refines the visual explanation superiority hypothesis we suggested above: it is only applicable for complex explanations.

What about the results in Atkinson's study (2002)? The same interpretation might be true for the results in that study. Participants learned to solve word-problems on proportion receiving explanatory aids helping them to understand the rationale behind the solution steps. In most of the comparisons, there were no differences between the visual and the auditory explanatory aids. The non-advantage of the visual explanatory aids could be explained in the same fashion as before: explanations can be considered simple (as they comprised few ideas and relations among them; see Appendix (A)). Once again, prior results fit into the research framework we have suggested.

Graesser et al. (2003) had participants learn computer literacy with the help of aids to monitor and revise their understanding. That is, they provided learners with both regulatory and explanatory aids. They found that the modality of the aids was not significant. This does not fit the functional approach. One possible explanation is that, as Graesser et al. did not distinguished regulatory from explanatory aids (modality was manipulated for the whole set of aids), the possible interaction between type of aid and modality remained unexplored.

Overall, the functional approach is able to explain divergent findings in previous studies. Once the functional approach is provided with empirical support, these mixed evidence will be discussed again.

1.5. Overview of the study.

A set of experiments in which participants learned from a CBLE were carried out. Both regulatory and explanatory aids were incorporated into a CBLE teaching geology (specifically, the plate tectonics theory). All the CBLEs included identical animations with concurrent narrations but there were different versions depending on the modality of the aids. A 2 x 2 factorial design was used with the modality of the regulatory aids (visual/auditory) and the modality of the explanatory aids (visual/auditory) as the between-subjects factors. Thus, there were four versions of the aids in the CBLEs. After using the CBLE, the participants solved retention and transfer tests, which were used as the dependent variables.

In the light of the framework suggested above, some predictions were made. We expected the participants receiving visual explanatory aids to outperform those receiving auditory explanatory aids. We expected the participants receiving auditory regulatory aids to outperform those receiving visual regulatory aids. This would be true for the performance in the retention test, since the aids would allow the participants to better recall basic concepts in plate tectonics, and also for the performance in the transfer test, since the aids would allow the participants to revise and repair their mental representations in such a way it would be possible for them to apply the knowledge acquired to novel situations.

An important question is that exploring the impact of the modality of explanatory and regulatory aids is reasonable only if the aids prove to be effective. For that reason, control conditions (without aids) were included in the experiments.

2. Experiment 1: The effect of the visual and the auditory modalities of different verbal aids on learning.

In this experiment the impact of the modality of verbal aids on learning was examined. The participants learned plate tectonics from a CBLE and received regulatory and explanatory aids presented either visually or auditorily. Then the participants solved retention and transfer tests. The goal was to test the predictions of the functional approach. Specifically, those participants receiving aids presented in a fashion compatible with the framework presented above were expected to outperform their counterparts.

It is worthy of consideration that exploring the differences between the modalities in which the regulatory and explanatory aids are presented is relevant if and only if these aids are effective. For this reason it is essential to gather support for the effectiveness of the regulatory aids before testing the impact of their modality. Accordingly, a control condition was introduced into the experiment involving no aids (control). As will be shown later, preliminary analyses were conducted comparing learning from two different conditions, namely, one including aids (using the data from the condition including auditory regulatory

aids and visual explanatory aids) and one without aids (using the data from the control condition). It should be noted that there is evidence that this type of aids are indeed effective (Alevén & Koedinger, 2002; Atkinson et al., 2003; Berthold et al., 2008; Conati & VanLehn, 2000; Graesser et al., 2004; Moreno, 2004; Moreno & Mayer, 2005; Sánchez et al., 2008; Schworm & Renkl, 2006; VanLehn et al., 2007) but we considered appropriate to conduct an additional test exploring this for one reason. Although the results from different experiments provide evidence in favor of the effectiveness of the regulatory and explanatory aids, this does not necessarily mean that these aids were effective in this very experiment. More accurately, these aids are not necessarily effective for the participants and materials used here. A more stringent test is that including a control condition.

2.1. Participants and design.

Sixty-five undergraduate students enrolled in educational psychology courses at the University of Salamanca (Spain) participated in this study. They were randomly assigned to one of five conditions. Thirteen participants served in the auditory regulatory aids and visual explanatory aids condition (AR-VE), 13 in the visual regulatory aids and auditory explanatory aids condition (VR-AE), 13 in the visual regulatory aids and visual explanatory aids condition (VR-VE), 13 in the auditory regulatory and auditory explanatory aids condition (AR-AE), and 13 in the control condition including no aids (control). The mean age of the sample was 21. Eighty one percent of the sample consisted of female students and 19 percent of male students. Although there was a majority of women, genre does not seem to be a critical variable, as observed in similar studies (Buisine & Martin, 2007; Moreno & Flowerday, 2006). All the participants had experience in using computers.

The experiment had a 2 x 2 factorial design with the modality of the regulatory aids (visual/ auditory) and the modality of the explanatory aids (visual/ auditory) as the between-subjects factors and also included a control condition (with no aids). The performances in the retention and transfer tests were the dependent variables. General reading comprehension skill and prior domain knowledge were used as control variables.

2.2. Materials.

For each participant materials consisted of a prior knowledge test, a general reading comprehension skill test, the CBLE, a retention test, and a transfer test. They all are described next.

2.2.1. Prior knowledge test. The prior knowledge test consisted of a paper-and-pencil test and comprised six open-ended questions. These questions tested on basic notions about geology and issues directly addressed in the CBLE that was presented later. The questions in this test were the following. “What is a plate? Where are plates located in the Earth's internal structure?”, “Is it possible for continents to be permanently moving? Explain why.”, “Is it possible to find fossil on the peak summit of a mountain? Explain why.”, “Explain how volcanoes are formed.”, “How are mountains formed?” One additional question was a naming task in which students had to name the parts of an illustration. The Earth's internal structure was showed, distinguishing its three layers.

2.2.2. General reading comprehension skill test. The general reading comprehension skill test consisted of eight multiple-choice questions about a narration, which were presented on the computer screen. It was a Gernsbacher and Varner test (1988) translated into Spanish by Díez and Fernández (1997). The narration was about a trader who travelled to Cabo Verde looking for new trade routes. The trader helped people in Cabo Verde to kill the rats living there, as it was an uncomfortable situation for Cabo Verdians. Participants solved questions about the narration such as “How did the King of Cabo Verde feel when the cats killed all the rats?” or “Why were there soldiers holding lances during the dinner?”

2.2.3. Computer-based learning environment. The CBLE included both (a) animations with concurrent narrations and (b) aids²⁹. The animations with concurrent narrations described several events concerning plate tectonics, namely, the three layers in the internal structure of the Earth and their relations, the convection currents as the origin of plates' movements and collisions, the collision between a continental and an oceanic plate and its consequences on the Earth's surface, the Andes range as an example of this type of collision, the collision between two continental plates and its consequences on the Earth's surface, the Himalaya range as an example of this type of collision. The narrations and animations lasted 386 seconds (six minutes approx.). They were presented in a sequential (instead of hyperspatial) fashion. The mental model we wanted the participants to build from the animations and narrations was the following. “The Earth's core is a big sphere made of iron and is very warm. The core makes the magma (i.e., a doughy substance made of melt rocks) in the mantle heat up and, thus, approach the Earth's surface or crust. When the magma gets cold again, it moves away from the crust. As a result, the convection currents are formed: the magma is continuously moving up and down. The convection currents push the plates, as they are floating on the magma, making them move away from each other. When moving, the plates can collide with other plates. There are different kinds of collisions depending on the plates engaged in the crashes. The plates can collide and one can sink (in the trenches) inside the Earth originating mountains with volcanoes (e.g., the Andes range); or they can collide and move vertically originating mountains without volcanoes (e.g., the Himalaya range). When a plate sinks in the trenches it is destroyed and becomes magma again.”

The CBLE also included aids helping the participants to detect and repair possible problems of understanding. On the basis of prior studies (e.g., Sánchez et al., 2008), the most common misunderstandings were identified. Aids were provided to assist the comprehension of those aspects typically misunderstood. There were several aspects typically simplified or mixed-up by the participants in prior experiments (in which different instructional materials were used in such a way that an inappropriate design of them could not be pointed out as the cause of the misunderstandings). One of these aspects was the following. The specific features of the Andes and the Himalaya plates' collisions were mixed-up. Many participants underestimated the particular characteristics of each type of collision, building a mixed model in which features of both types of collision were included. Accordingly, an aid episode was created involving two parts. The regulatory part helped the participants to monitor their learning process to detect the misunderstanding (confusion between Andes and Himalaya). The explanatory part assisted the participants in revising and repairing their flawed mental

²⁹ The CBLEs were developed using the ToolBook II 6.1 (Asymetrix Corporation, 2001).

representations. When processing the regulatory part, the participants had to understand how important was to check their mental representations and to start doing it. When receiving the explanatory part, the participants have to process carefully an elaboration on the to-be-learned contents. The transcription of the aid is the following:

- [The CBLE poses a question to the participant] “After watching these plate collisions a question might have raised: What exactly is the relationship between the Andes and the Himalaya plate collisions? Choose one of these answers: (a) the collisions are very similar to each other, (b) the collisions are very different from each other” [The participant chooses (a)]
- [Regulatory aid] “It is true that in both collisions there are two plates crashing and in both collisions mountains are formed. However, that is not the correct answer: the collisions present differences that are critical to understand plate tectonics.”
- [Explanatory aid] “In the Himalaya collision two continental plates are involved. That is, the plates have identical weight and thickness. That explains why the collision is a head-on crash. Conversely, in the Andes collision the plates involved are different: one is continental and the another is oceanic. This means there is one plate, the oceanic, heavier than the other one. That explains why it sinks below the continental plate and goes back to the mantle. Furthermore, the plates in the Himalaya collision push and compress each other forming a mountain without volcanoes. Conversely, in the Andes collision the oceanic plates sinks putting pressure on the continental plate and forming cracks on it, through which magma from the mantle surfaces forming mountains with volcanoes.”

(On the basis of prior experiments we considered appropriate to assume that all the participants would select the incorrect answer (a). Therefore, as a default strategy, the CBLE selected the answer (a) while saying “as many learners do, you might have chosen (a)”. Thus, every participant -except those in the control condition- received the regulatory and the explanatory aid).

Depending on the experimental condition the participants received each of these aids presented either in one modality or the other. The participants in the control condition received no aids.

There are three aspects of the aids that require a close examination. First, in order to develop the auditory sections, an experienced teacher was recruited (Emilio Sánchez). He recorded both the regulatory and the explanatory aids and the resulting audio recordings were attached to the CBLE. In both types of aid, the teacher who recorded the audio recordings attempted to employ auditory expressiveness. Two judges selected those recordings which they considered more natural and convincing. Criteria for considering a recording convincing were that it must clarify the speaker's communicative intention and provide additional clues about his/her attitude toward the message. In the regulatory part, a convincing recording was one clarifying that the communicative intention was to warn about the possible misunderstanding (regarding differences between the Andes and the Himalaya) and showing an attitude of urgency and importance (i.e., the speaker stressed the urgency of revising the misunderstanding and the seriousness of the problems related to thinking about differences between the Andes and the Himalaya as something trivial). In the explanatory part, a

convincing recording was one talking about the differences between the Andes and the Himalaya in an intelligible pace and accent, clarifying that the intention was to make some statements (not warn, request, or others), and showing a neutral attitude.

A second question regarding the design of the aid episode is that there were no pictures related to it. This meaning that learners did not have to mentally integrate verbal information in the aid episode with any picture. Moreover, the aids were not merely narrating the events depicted by pictures but they were accomplishing different functions (helping learners to identify misunderstandings, helping learners to revise their misunderstandings). These are the reasons why the modality principle was not applicable for verbal aids.

Finally, it should be noted that the explanatory aid consisted of a complex explanation. That is, it comprised many ideas and interconnections (see the Appendix (A)), which demands learners to grasp a considerable number of propositions and establish the corresponding links among them.

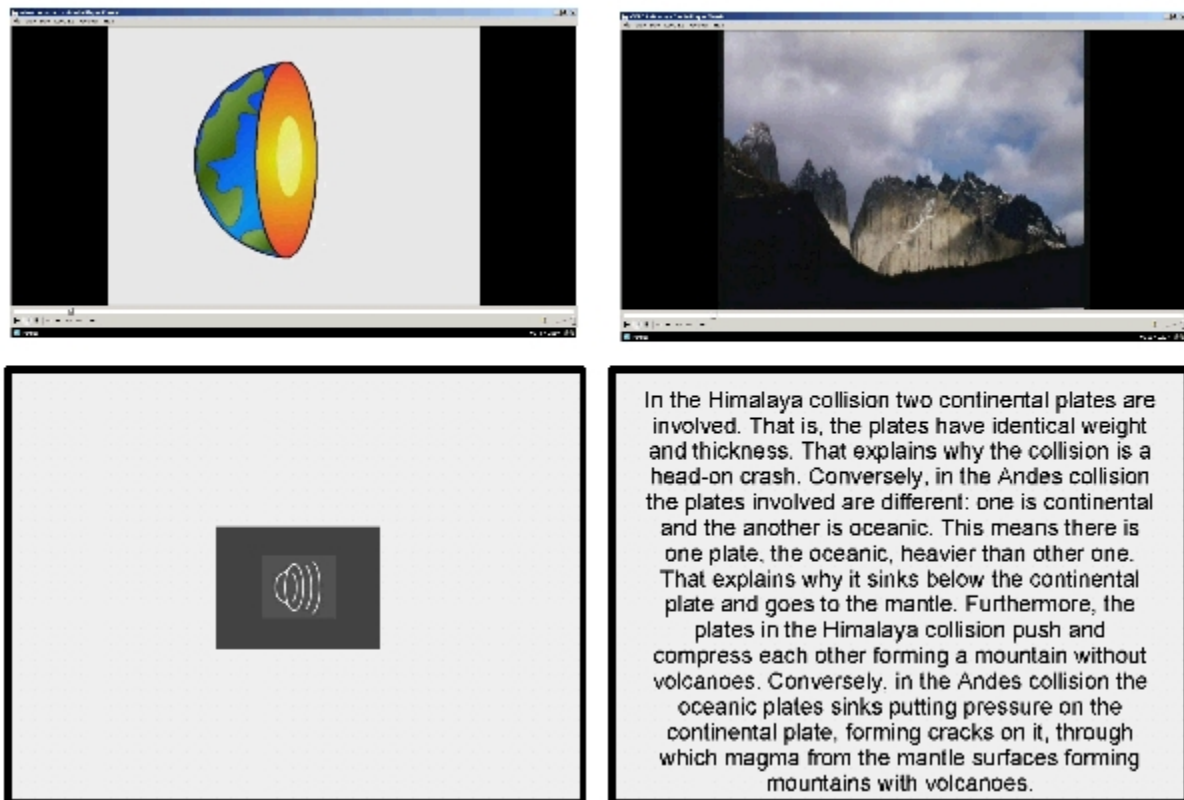


Figure 1. Screenshots from the computer-based learning environment. The pictures above correspond to the animations with concurrent narrations on plate tectonics. The pictures below correspond to an aid presented either in auditory modality (left) or visual modality (right).

2.2.4. Retention test. The retention test consisted of a paper-and-pencil test which contained eight open-ended questions. These questions required the participants to recall key

information that was presented in the CBLE. The questions were the following. “Which are the main plate movements?”, “What is a trench?”, “Which are the similarities between the Andes and the Himalaya plate collisions?”, “Is any type of plate material destroyed in the subduction?”, “What are the differences between the Andes and the Himalaya plate collisions?”, “What happened to the continental plate in the Andes collision?”, “What happened to the ocean floor in the Himalaya collision?”, “Why does the oceanic plate sink in the Andes collision?”

2.2.5. Transfer test. The transfer test consisted also of a paper-and-pencil test and comprised eight open-ended questions. The questions presented some hypothetical conditions to participants, who had to predict some results on the basis of those conditions and the knowledge acquired from the CBLE. The questions in this test were the following. “How could a mountain like the Himalaya be formed in Holland?”, “If the Andes' volcanoes stopped erupting, what would you think about it?”, “If new plate material is created permanently in the ridges, why does the Earth's surface remain the same size?”, “Is it possible for the Himalaya mountains to have volcanoes in the future?”, “How should the coast in the nearness of the Andes chain look like?”, “Imagine you find a fossil in the peak summit of a mountain, is it a mountain like the Andes or the Himalaya?”, “Is it possible for the Himalaya chain to have a ridge in the future?”, “Imagine the mountains in the Andes stop growing, what would happen?” There was also a task including an illustration in which participants had to place a volcano and explain why it was there.

2.3. Procedure.

The participants were tested simultaneously in groups of 10 to 15 participants per session. Each participant was seated in front of his/her individual computer and headphones. Learning materials were presented on Acer personal computers (Intel Pentium III processor), which included 17 inches flat monitors. First of all, the participants received some basic instructions. These instructions were as follows: “Thank you for participating in this experiment. We are interested in how we can help people to learn from computer-based materials. You will be asked to use a computer-based learning environment on plate tectonics. After using the materials you will have to solve two sets of questions. Please, solve all the tests in silence. Do not forget requesting for your tests when you have finished viewing the presentation. Before using the computer materials, we want you to fill in a prior knowledge on geology. Please, try to remember all the things you know about it.” After receiving these instructions, the prior knowledge test was delivered and the participants solved it in silence. It took no more than 12 minutes. When everyone finished filling in the prior knowledge test, the participants started using the CBLE. An experimenter ran the computerized materials on the participants' computer. The experimenter randomly ran the version of the CBLE so that participants were randomly assigned to each condition. The CBLE lasted about 480 seconds (eight minutes approx.) including the animations with narrations (386 seconds, six minutes approx.) and the aids. The participants viewed the animations with concurrent narrations describing the events listed before about the plate tectonics theory. In addition, they received the regulatory and explanatory aids. These aids were intended to assist the participants in detecting and repairing misunderstandings, as was explained above.

Whenever an aid was presented visually, the participants could use it as follows. Because we were interested in exploring the effect of the control of the visual modality, it was allowed for participants viewing aids on the screen to take as long as they wanted in doing so (times were recorded). Moreover, the whole text was presented in an all-at-once fashion on the screen so that the participants could freely manage their time, exploring the segment of the text they wanted.

Whenever an aid was presented auditorily, participants could use it as follows. The audio recordings were played once. The auditory regulatory aid lasted 33 seconds whereas the auditory explanatory aid lasted 66 seconds. Audio recordings were created by an experienced teacher and they were judged as natural and convincing by two judges, as was explained before. The recordings were intended to show maximum expressiveness.

The participants in the AR-VE condition listened to the regulatory aid through their headphones. The audio recording was played once. Then, the participants viewed the explanatory aid on the computer screen. They could freely manage their time. The participants in the VR-AE condition viewed the regulatory aid on the computer screen having control on their reading. Then, they listened to the explanatory aid once. The participants in the VR-VE condition viewed both the regulatory and the explanatory aids on the screen having control on their reading. The participants in the AR-AE condition listened to both the regulatory and the explanatory aids. The corresponding audio recordings were played once. Finally, the participants in the control condition received no aids after watching the animations with concurrent narrations.

After using the CBLE (including the animations with narrations and the aids) the participants were given the retention test. Solving the test took no more than 10 minutes. Then they were given the transfer test and they had no more than 15 minutes to solve it. Finally, they were tested on general reading comprehension skill. They read a story about an explorer on the computer screen and then answered a set of multiple-choice questions on a paper. After this final test was collected, participants were seen off. Each session lasted about 60 minutes.

2.4. Scoring.

A rater scored all the questionnaires unaware of the condition of each participant. Approximately 30% of the tests were also scored by a second rater. Interrater agreement was above .85 in the prior knowledge, the retention, and the transfer tests. Disagreements were resolved by consensus.

A template with possible answers was developed for the prior knowledge, retention, and transfer tests. It included accurate, correct but incomplete, and incorrect answers. They yielded 2, 1, or 0 points, respectively.

2.4.1. Prior knowledge test. Accurate answers for the question about the plates in the prior knowledge test were those stating that they are blocks of the superficial layer of the Earth. Accurate answers for the plates' location question included saying that they are above the mantle, as they are part of the superficial layer. Accurate answers for the question about

continents moving included stating that continents move because they are also plates or part of a plate. Accurate answers for the question about fossils were those stating that mountains comprise blocks that were under the sea in the past. Accurate answers for the question about volcanoes formation included mentioning plates collision and cracks in one of the plates involved as the causes. Accurate answers for the question about mountain formation were those mentioning plates collision. Each element correctly named in the question about the Earth's layers yielded 1 point. Total scores ranged from 0 to 15 in the prior knowledge test.

2.4.2. General reading comprehension skill test. Each right answer in the general reading comprehension skill test yielded one point. Wrong answers yielded zero points. Total scores ranged from 0 to 8.

2.4.3. Retention test. Accurate answers for the question about main plate movements included mentioning divergent and convergent movements (distinguishing the Andes and the Himalaya types). Accurate answers for the question about the trench were those stating that it is a trough where the oceanic plate material is destroyed. Accurate answers for the question about similarities between the Andes and the Himalaya included stating that in both cases two plates collide and new mountains are formed. Accurate answers for the question about plate destruction were those saying that the oceanic plate is destroyed in the Andes collision. Accurate answers for the question about differences and similarities between the Andes and the Himalaya collisions included mentioning the plates involved, the processes, and the effects on the Earth's surface of each collision type. Accurate answers for the question about the continental plate in the Andes were those stating that it suffers from oceanic pressure and, hence, cracks are formed on it. Accurate answers for the question about the ocean floor in the Himalaya included saying that it was destroyed, dragged, and finally integrated into the mountain formed. Accurate answers for the question about the sinking of the oceanic plate in the Andes were those stating that the oceanic plate sinks as a result of its weight. The range of total scores was 0-16 in the retention tests.

2.4.4. Transfer test. Accurate answers for the question about locating a volcano in a picture included locating it and explaining the reason behind that location. Accurate answers for the question about a mountain in Holland included mentioning that there should be two continental plates colliding and finally folding upwards. Accurate answers for the question about volcanoes without eruptions included saying that the collision is over so there is no longer subduction, no pressure, and, then, no cracks through which magma emerges. Accurate answers for the question about volcanoes in the Himalaya were those arguing that it will not be volcanoes in the Himalaya because no plate sinks and, hence, there is no pressure on the continental plates forming cracks so that is not possible for the magma to surface. Accurate answers for the question about the coast in the Andes included mentioning that it should have craggy cliffs. Accurate answers for the questions about fossils were those choosing the Himalaya as the place where it is possible to find fossils and explaining why. Accurate answers for the question about a trench in the Himalaya were those mentioning that it is not possible for the Himalaya to have a trench because the two plates involved in the collision are equal and, hence, none of them sinks. The range of total scores was 0-16 in the transfer test.

2.5. Results.

The variables under analysis were prior knowledge, general reading comprehension skill, retention, and transfer tests scores. The times taken on the visual regulatory and explanatory aids were recorded and they will be showed below. Prior knowledge and general reading comprehension skill scores were analyzed using a one-way analysis of variance. This was done to ensure that all the conditions had comparable levels on these two variables. As will be shown, there were differences on the general reading comprehension skill; accordingly, this variable was entered as a covariate in the subsequent analyses. Preliminary analyses were conducted in order to ensure that the aids were indeed effective. A one-way analysis of covariance was used comparing the scores in the retention and transfer tests of the condition including aids (AR-VE) with those of the condition including no aids (control). The AR-VE condition was used for these contrasts (instead of others) because this condition was expected to be the best. After these analyses, retention and transfer tests scores were analyzed using a factorial analysis of covariance with the modality of the regulatory aid (visual/auditory) and the modality of the explanatory aid (visual/auditory) as the between subjects factors. An alpha of .05 was used throughout this study. Eta-squared (η^2) was calculated as a measure of the effect size whenever there was a significant effect (or a trend). All the scores are showed in Table 1.

Table 1. Means (*M*) and standard deviations (*SD*) of all the conditions in all the control and learning measures: Experiment 1.

	AR-VE		VR-VE		AR-AE		VR-AE		Control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Prior Knowledge	5.88	2.87	5.92	2.29	6.00	2.24	6.73	2.89	5.92	2.66
G. Reading C. S.	3.85	1.63	5.46	0.97	5.38	0.77	5.54	1.20	4.62	1.56
Retention	8.35	2.76	8.46	1.63	7.54	0.85	8.00	2.01	7.00	1.53
Transfer	7.23	2.86	6.85	1.95	4.23	2.65	4.08	1.98	4.23	2.09

Note. The maximum scores were 15 for prior knowledge, 8 for general reading comprehension skill, 16 for retention, and 16 for transfer.

2.5.1. Control variables. Two one-way analysis of variance were conducted in order to ensure that all the conditions had the same level in prior knowledge and general reading comprehension skill. With respect to prior knowledge there were no significant differences, $F(4, 60) = 0.25$, $MSE = 6.78$, $p > .50$. Post-Hoc pairwise comparisons (using Fisher's LSD) confirmed that there were no differences between the conditions on this measure (all p 's $> .10$). With respect to general reading comprehension skill there were significant differences, $F(4, 60) = 4.30$, $MSE = 1.61$, $p < .005$, $\eta^2 = .29$. Post-Hoc pairwise comparisons showed that the AR-VE condition was the worst (p 's $< .05$). This is the reason why we decided to enter this variable as a covariate in the subsequent analyses.

2.5.2. Preliminary analyses. Preliminary analyses were aimed at demonstrating that the aids were effective. With respect to the retention test, the participants in the AR-VE

condition scored higher than their counterparts. The difference was reliable, as confirmed by a one-way analysis of covariance, $F(1, 23) = 4.73$, $MSE = 4.24$, $p < .05$, $\eta^2 = .17$. This means that the participants receiving aids were more able to recall the key concepts of the plate tectonics theory, as compared with the participants serving in the control condition. With regard to the transfer test, again the participants receiving aids outperformed those receiving no aids, $F(1, 23) = 9.70$, $MSE = 6.40$, $p = .005$, $\eta^2 = .30$. This indicates that the participants in the AR-VE condition were better than their counterparts in generating solutions for the transfer tasks.

In sum, these results indicate that the regulatory and the explanatory aids were effective in helping learners to construct coherent mental representations on plate tectonics. As long as the aids were effective it seemed reasonable to explore which configuration of the modality of these aids was the best.

2.5.3. Retention test. The conditions including the visual explanatory aid (namely, the AR-VE and the VR-VE conditions) scored higher than their counterparts (namely, the AR-AE and the VR-AE conditions) in the retention test. An analysis of covariance revealed that the modality of the explanatory aid was a significant factor, $F(1, 47) = 5.56$, $MSE = 3.09$, $p < .05$, $\eta^2 = .11$. This indicates that the participants receiving the visual explanatory aid recalled basic concepts about plate tectonics better than those who received the auditory explanatory aid. This is consistent with our predictions. The scores in the retention tests were similar between the auditory regulatory aid conditions (namely, the AR-VE and the AR-AE conditions) and the visual regulatory aid conditions (namely, the VR-VE and the VR-AE conditions). Not surprisingly, an analysis of covariance showed that the modality of the regulatory aid was not significant, $F(1, 47) = 0.45$, $p = .50$. This means that the participants receiving the auditory regulatory aid did not recall basic concepts better than the participants receiving the visual regulatory aid. This is not in line with the predictions we formulated. The interaction between the two factors was not significant, $F(1, 46) = 1.88$, $p > .10$, indicating that the results reported above were not qualified by any interaction.

2.5.4. Transfer test. Regarding the transfer questions, scores were not equal for all conditions, with participants in the visual explanatory aid conditions performing better than those in the auditory explanatory aid conditions. An analysis of covariance revealed that the difference was reliable, $F(1, 47) = 22.23$, $MSE = 5.54$, $p < .001$, $\eta^2 = .32$. That is to say, the participants in the visual explanatory aid conditions were more able to apply the knowledge they had gained to solve tasks than did those in the auditory explanatory aid conditions. This confirmed our prediction. The scores in the transfer test were not different between the conditions including the auditory regulatory aid and those including the visual regulatory aid. This was confirmed by an analysis of covariance, $F(1, 47) = 0.96$, $p > .10$. Accordingly, the participants receiving the auditory regulatory aid and those receiving the visual regulatory aid were equally able to apply their knowledge on plate tectonics. This was not consistent with the predictions of the functional approach. As occurred in the retention test, the interaction between the two factors was not significant, $F(1, 47) = 0.45$, $p > .50$. This indicates that the results we have just reported were not qualified by any interaction.

2.5.5. Time-recordings. The participants in the visual regulatory aid conditions spent

less time ($M = 24.76$ seconds, $SD = 3.63$) viewing the regulatory aids than did the participants in the auditory regulatory aid conditions listening to that aid (33 seconds, $SD = 0$). The same goes for the explanatory aid: those in the visual explanatory aid conditions took less time ($M = 46.67$ seconds, $SD = 4.96$) viewing it than those in the auditory explanatory aid conditions (66 seconds, $SD = 0$). The differences were confirmed by non-parametric Mann-Whitney's U -tests (p 's $< .05$). Non-parametric tests were used because parametric tests cannot be applied if the standard deviation of one condition is zero.

2.6. Discussion.

Overall, the aids have proved to be effective and support has been gathered for the visual explanation superiority hypothesis but not for the auditory regulation superiority hypothesis. These two results need further comment.

The participants receiving the explanatory aid presented in the visual modality outperformed those receiving the same aid presented in the auditory modality. This was true both in the retention and transfer tests. In the light of the functional approach, we expected explanations to be better comprehended when presented visually, since the visual modality provides control to learners, which can be of help when processing elaborations on the contents. The results are consistent with our predictions. In order to interpret why this was the case, several explanations can be suggested. First, one might think that participants in the visual explanation conditions took more time viewing the explanatory aid than those listening to it. But this was not so. As the time-recordings showed, the participants having control over the explanatory aid did not take more but less time than those listening to that aid. The participants in the visual explanatory aid conditions dedicated a 25% less time (approximately) than their counterparts on average. Another possibility is to think that the participants receiving the visual explanatory aid read it (completely) more than once. As we did not count the number of readings it is not possible to rule out this possibility. However, it seems not very much likely that they did so, as they spend 330 milliseconds (46,670 milliseconds / 141 words) per word on average. According to some evidence, readers need about 350 milliseconds to access the meaning of a word (see for example Kintsch, 1998). A third possible explanation is that the control allowed these participants to freely manage the time available, even if it was short. This is consistent with patterns of visual search during reading (e.g., Just et al., 1982; Rayner, Juhasz, & Pollatsek, 2005): readers fixate longer on certain words, review prior segments when they need it, and so on. The point is not the total time spent but the way this time is distributed. Therefore, participants in the visual explanatory aid conditions distributed the time available in an efficient way. This is particularly relevant when learners are provided with complex texts, such as the explanatory aids used here.

The reinterpretation we made before with respect to mixed evidence in prior research has been confirmed. We argued that, if there was no advantage of the visual explanatory aids over the auditory explanatory aids in the studies of Moreno and her colleagues (Moreno & Mayer, 2002; Moreno et al., 2001) and Atkinson (2002), it was due to the use of simple (instead of complex) explanations. This was in line with prior findings in (explanatory) text comprehension (for complex texts the visual modality is better than the auditory modality

(Green, 1981); for simple texts the modality does not play a role (Gernsbacher et al., 1990; Kintsch & Kozminsky, 1977; Kintsch et al., 1975; Smiley et al., 1977), though). In our Experiment 1 the explanatory aid was a complex, instead of simple, elaboration on the to-be-learned contents (see Appendix (A)); and, in this case, the visual modality outperformed the auditory modality. This demonstrates that previous findings fit into the functional approach.

With regard to the findings of Graesser et al. (2003), the results we have found may allow us to reinterpret them. As was argued before, one possibility is that in the experiment of Graesser et al. regulatory and explanatory aids (which were actually used) were not distinguished and, thus, not controlled: modality was manipulated for the whole set of aids, without exploring the possible interaction between type of aid and modality. It is not clear what would happen if they had kept different aids under experimental control. In the light of the results in our Experiment 1, one may expect visual explanatory aids to be better than auditory explanatory aids. Another possibility is to think that the regulatory aids in the experiment of Graesser and associates were pervasive whereas the explanatory aids were very rare. So far, modality seems to be not relevant when giving regulatory aids (this is what we have found in the Experiment 1). If the aids in Graesser et al.'s experiment were mainly regulatory, then one can expect modality to be not significant, as was the case. This might explain why, although visual explanatory aids can be better than auditory explanatory aids (as in our Experiment 1), this advantage remained canceled. If this second interpretation is true, it is not clear what would happen if they had used more explanatory aids.

Our results regarding the advantage of the visual modality in explanatory aids extend the work of Green (1981) on (explanatory) text comprehension in two ways. First, they replicate her findings. Second, whereas she used expository texts, we employed explanations in the form of verbal aids incorporated into a CBLE.

We found no support for the auditory regulation superiority hypothesis. This was indicated by the fact that the participants in the auditory regulatory aid conditions did not outperform those who served in the visual regulatory aid conditions. The same pattern was found in the retention and the transfer tests. We predicted the auditory regulation to be better than the visual regulation because the former provided additional information -such as the communicative intention and the speaker's attitude. This additional information would be relevant only when dealing with regulatory aids, since they assist learners in deciding whether or not it is worthy to revise their understanding. In those cases, knowing what the speaker or assistant wants and his/her attitude towards the action he/she is asking the learner to perform would be beneficial. However, this was not true in this experiment. Why was not the auditory modality in regulation as decisive as we expected? One possibility is that the auditory modality needs further clues to make the additional information it provides more trustworthy. For instance, the auditory modality would need not only expressive clues (i.e., prosody) but a visible speaker or agent. This is in line with the following arguments. Moreno and her colleagues (Moreno & Flowerday, 2006; Moreno et al., 2001) have suggested that agents incorporated into CBLEs promote a sense of social agency, which means that learners are more engaged in learning when they think they are interacting with another human. That engagement, in turn, enhances learning. This rationale is in line with the notion of *persona effect* (Dehn & van Mulken, 2000; Johnson, Rickel, & Lester, 2000). According to these

suggestions, all clues that enhance the sense of social agency (e.g., visibility of the agent) may foster learning. Thinking of the assistant or agent as a human would make the expressive additional information more trustworthy, which might reinforce the impact of the auditory regulatory aids. This rationale is described more thoroughly in the next section.

An alternative explanation for the failure in confirming the superiority of the auditory regulation over the visual one is that the regulatory aid was irrelevant. In the preliminary analyses a condition including the regulatory and the explanatory aids was compared with a control condition including no aids. The former performed better than the latter on both the retention and transfer tests. In the light of that we concluded that the aids were effective. Nevertheless, the specific impact of each aid was not tested. In order to do so a 2 x 2 factorial design should have been used (presence/absence of the regulatory aid, presence/absence of the explanatory aid). A design like that makes it possible to roughly test the contribution of each aid. As that was not the comparison we made, it is not possible to rule out the possibility that, although the combination of regulatory plus explanatory aids was effective, it was due to the effectiveness of the explanatory aid in isolation. If that was the case, then a possible explanation for the null impact of the modality of the regulatory aids is that regulatory aids are irrelevant. Nevertheless, there are evidence that the regulatory aids (*per se*) are indeed helpful. These come from studies using different materials and participants from those used here (Atkinson et al., 2003; Moreno & Mayer, 2005; Schworm & Renkl, 2003) but also from studies carried out in our lab (Acuña, García-Rodicio, & Sánchez, 2008 and see also the Experiments 3 and 4 presented here), in which both materials and participants were very similar to the ones in the Experiment 1. Because of these evidence an explanation based on the assumption of regulation's ineffectiveness seems to be not very much plausible.

3. Experiment 2: Exploring the impact of the modality of different aids with a pedagogical agent.

The second experiment had two goals. First, to explore more thoroughly the impact of the auditory modality when giving regulatory aids. Second, to replicate the results found in the first experiment regarding the superiority of the visual explanatory aids.

In this experiment, it was hypothesized that the visibility and physical presence of the pedagogical agent would make the additional information delivered through the auditory modality more trustworthy and, hence, effective. According to the *social agency theory* (Moreno & Flowerday, 2006; Moreno et al., 2001), when learners think about the pedagogical agent as a human, their engagement is enhanced and, thus, learning is promoted. The more clues (e.g., visibility, presence), the more humanity is perceived (also in line with the notion of persona effect, Dehn & van Mulken, 2000; Johnson et al., 2000). We argued that this sense of humanity would be beneficial only for regulatory aids. When dealing with explanatory aids, learners have to assimilate the contents comprised by them without taking special decisions. Conversely, regulatory aids make learners to take decisions (“Should I revise my understanding? Or, alternatively, should I do anything different from that?”). This means that learners provided with regulatory aids have to reject doing something and perform the action expressed in the aid. In so doing they have to assess whether it is worthy or not to carry out

that action. The additional information delivered through the expressiveness of the auditory modality may facilitate such a decision taking (after all, knowing the attitude of others towards certain task affects your own attitude, e.g., Wild et al., 1997; and recall that others' attitudes are embodied by expressiveness, e.g., Brennan & Williams, 1995). The point is that for such additional information to be trustworthy, a real life-like agent should provide the regulatory aids. When this condition is attained agent's intentions and attitudes become more believable and reliable. As a consequence, the impact of regulatory aids should be more powerful as well.

In the light of the above, we decided to introduce a real agent into the CBLE presentation. The agent was not an animated virtual agent but an actual human being (an experienced teacher). This person came along with the participants through their session using the CBLE and took part whenever they received aids presented in the auditory modality. Therefore, auditory aid recordings from the Experiment 1 were replaced by actual utterances from a human being, who uttered them in a trustworthy fashion. The agent, then, was not only visible but physically present in the experimental session. Both visibility and presence were considered clues of humanity. It is interesting that prior studies have also introduced a human being as a pedagogical agent in the experimental session (e.g., Azevedo, Moos, Greene, Winters, & Cromley, 2008), although they had a different purpose.

It should be noted that the impact of the agents' visibility has been already tested empirically, finding somewhat discouraging results (see Dehn & van Mulken, 2000). Both when pedagogical agents are used as deliverers of the contents learners have to learn (Craig, Gholson, & Driscoll, 2002; Mayer, Gayle, & Mayer, 2003; van Mulken et al., 1998) and when pedagogical agents are used as assistants giving aids (Atkinson, 2002; Baylor & Ryu, 2003; Graesser et al., 2003; Mitrovic & Suraweera, 2000; Moreno et al., 2001) the effect of visibility on learning is nonexistent. For instance, Baylor and Ryu (2003) had participants solve psychopedagogical consultation problems with the help of multiple agents. These agents gave different aids, such as corrective feedback ("that's a good idea") or explanatory feedback ("but... the students won't get it just from the definitions and the laws... they need to interact with the information"). There were three types of agent: visible and animated, static, and invisible. There were no differences on the learning measures between the conditions.

When considering the agents' fulfilling an assistance function (aid providers rather than contents deliverers) and its effect on learning measures, why was the agents' visibility ineffective? We argue that visibility is relevant only for regulatory aids, not for explanations. Perceiving additional clues (such as agent's attitude) is not beneficial in learning from an explanation, as the learner have no decisions to take while doing so. However, these clues would be useful when having to take decisions, as they facilitate the decision taking. In the experiments we have just cited designers gave visibility to agents providing explanatory but not regulatory aids (in Atkinson and Moreno et al.) or an agent providing the two types of aid without controlling it (in Baylor & Ryu and Graesser et al.). This would explain why those researchers found no effect of agents' visibility. Additionally, it could be argued that agents' visibility is not strong enough; hence, clues such as physical presence are also required. Our Experiment 2 differed from earlier work in two ways. We attempted to solve those problems by (a) examining the impact of agent's features either in explanatory or regulatory aids and (b)

providing our agent with additional clues (i.e., he was physically present) besides visibility.

An agent can provide not only clues making his/her utterances more trustworthy but also other paralanguage devices such as gestures. The use of gestures may make the agent more effective, provided that gestures can accomplish a communicative/pedagogical function. As was shown in the chapter 2, the problem is that the value of speakers' gestures is controversial. To start with, it is not clear if gestures are meant to be communicative. Bavelas, Chovil, Lawrie, & Wade (1992) claim that gesturing is communicative, as speakers produce more gestures when there is an addressee than when there is not. Krauss (1998) claims that gesturing aids in lexical retrieval because when speakers are prevented from gesturing their speaking rate decreases. McNeill (1985) claims that gestures constitute thought, since they are parallel to speech, revealing that they aid in formulating the message. Second, evidence that listeners or learners rely on the gestures of a speaker/agent is mixed. For instance, Buisine and Martin (2007) found that deictic gestures were useful in promoting better recall on their participants, who learned how to use a remote control and a photocopier. But Craig et al. (2002) found the same gestures useless for participants learning about lightning formation. For all these reasons we expected gestures produced by the agent, if any, to be not helpful. As has been argued, we expected agents' visibility and presence to make regulatory aids more effective; that was all.

We had the same predictions as in the Experiment 1. This means that the auditory modality and humanity of the verbal aids was expected to be helpful only for regulations but not for explanations. The control of the visual modality would be helpful only when dealing with explanatory aids. Therefore, we expected the participants in the visual explanatory aid conditions to outperform those in the auditory explanatory aid conditions; we further expected the participants in the auditory regulatory aid conditions to outperform those in the visual regulatory aid conditions. This would be true both on the retention and transfer tests scores.

As the materials and the participants used in this experiment were very similar to those used in the Experiment 1, no control condition was included. This means that we assumed the aids to be effective, as was the case in the Experiment 1.

3.1. Participants and design.

Sixty-one undergraduate students enrolled in educational psychology courses at the University of Salamanca (Spain) participated in this experiment. They were randomly assigned to one of the four conditions. Sixteen participants served in the auditory regulatory aids and visual explanatory aids condition (AR-VE), 14 in the visual regulatory aids and auditory explanatory aids condition (VR-AE), 16 in the visual regulatory aids and visual explanatory aids condition (VR-VE), and 15 in the auditory regulatory aids and auditory explanatory aids condition (AR-AE). The mean age of the sample was 22. Ninety-one percent of the sample consisted of female students and nine percent of male students. As was argued in the Experiment 1, this proportion of women may not represent a problem. All the participants had experience in using computers.

As in the Experiment 1, we used a 2 x 2 factorial design with the modality of the

regulatory aids (visual/ auditory) and the modality of the explanatory aids (visual/ auditory) as the between-subjects factors. In this experiment, the auditory modality consisted of a visible human agent providing the aids. Retention and transfer tests were again the dependent variables. General reading comprehension skill and prior domain knowledge were used as control variables.

3.2. Materials.

For each participant materials consisted of a prior knowledge test, a general reading comprehension skill test, the CBLEs, a retention test, and a transfer test. All the materials were identical to those used in the Experiment 1 except for the following changes.

3.2.1. Computer-based learning environment. The CBLE was the same as that used in the Experiment 1, except for the aids presented in the auditory modality. The CBLE consisted of the same animations and narrations and aids. However, in this experiment the auditory aids (both regulatory and explanatory) were provided by a human agent who was in the same room as the participants. In both types of aid, the human agent attempted to make use of the auditory expressiveness. The agent was an experienced teacher (Emilio Sánchez). As was done in Experiment 1, two judges assessed the appropriateness of the agent's utterances. A convincing regulatory aid was one clarifying that the communicative intention was to warn about the possible misunderstanding and showing an attitude of urgency and importance. In the explanatory part, a convincing aid was one clarifying the differences between the Andes and the Himalaya in an intelligible pace and accent, expressing that the communicative intention was to make some statements, and showing a neutral attitude. The judges were present in the experimental sessions in order to ensure this was the case. Furthermore, they checked the words the agent used to see if they were exactly the same as in the visual modality. Because the words were the same, the variable under consideration was the modality not the contents of the verbal aids.

3.3. Procedure.

The procedure was identical to that used in the Experiment 1 except for some changes we will describe next. The participants were again tested simultaneously in groups of 10 to 15 participants per session. Each participant was seated in front of his/her individual computer and headphones. When everybody finished the prior knowledge test, the participants started using the CBLE. An experimenter ran the computerized materials on the participant's computer. The participants were randomly assigned to each condition. This was done in a different fashion with respect to the Experiment 1. In this second experiment, all the participants in a session were assigned to the same condition (because the human agent gave the aids to all the participants in a session). A given session was randomly designated to use certain experimental condition. When using the CBLEs, the participants first viewed the animations with concurrent narrations and then received the aids.

Whenever an aid was presented in the visual modality, the participants viewed it on the screen taking as long as they wanted in doing so. The whole text was presented in an all-at-once fashion on the screen so that participants could freely manage their time, exploring the

segment of the text they wanted. A human agent presented the auditory aids. He uttered them once attempting to make use of the auditory expressiveness. Two judges were present in the experimental sessions to check several aspects. First, whether the agent was convincing. Second, whether the words uttered were the same as those included in the visual aids. Third, times were recorded in order to ensure that they were the same as those of the recordings in the Experiment 1. These judges reported that all aspects were appropriate.

The participants in the AR-VE condition listened to the regulatory aid, which was provided by the human agent once. Then, these participants viewed the explanatory aid on the computer screen having control on their reading. The participants in the VR-AE condition viewed the regulatory aid on the computer screen having control on their reading. Then, they listened to the explanatory aid, which was provided by the human agent once. The participants in the VR-VE condition viewed both the regulatory and the explanatory aids on the screen having control on their reading. The participants in the AR-AE condition listened to both the regulatory and the explanatory aids. They were provided by the human agent once.

After using the CBLE, the participants were provided the retention and transfer tests. They had no more than 10 and 15 minutes to solve them, respectively. Finally, the participants were tested on general reading comprehension skill. Each session lasted about 60 minutes.

3.4. Scoring.

The scoring procedures were identical to those used in the Experiment 1.

3.5. Results.

The variables under analysis were prior knowledge, general reading comprehension skill, retention, and transfer tests scores. The times taken on the visual regulatory and explanatory aids were recorded and they will be showed below. Prior knowledge and general reading comprehension skill scores were analyzed using a one-way analysis of variance. This was done to ensure that all the conditions had similar levels on these variables. Retention and transfer tests scores were analyzed using a factorial analysis of variance with the modality of the regulatory aid (visual/auditory) and the modality of the explanatory aid (visual/auditory) as the between subjects factors. All the scores are showed in Table 2.

Table 2. Means (*M*) and standard deviations (*SD*) of all the conditions in all the control and learning measures: Experiment 2.

	AR-VE		VR-VE		AR-AE		VR-AE	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Prior Knowledge	6.63	3.57	6.53	2.47	7.94	2.42	8.00	2.65
G. Reading C. Skill	5.63	1.36	5.31	1.66	5.47	0.99	4.71	1.07
Retention	6.50	1.55	6.44	1.75	6.34	1.11	5.57	1.65
Transfer	12.06	2.79	10.94	2.86	8.80	2.08	8.07	2.06

Note. The maximum scores were 15 for prior knowledge, 8 for general reading comprehension skill, 16 for retention, and 16 for transfer.

3.5.1. Control variables. The performances of all the conditions in prior knowledge and general reading comprehension skill were very similar. Two one-way analysis of variance were carried out. As revealed by these analyses, there were no significant differences in prior knowledge, $F(3, 57) = 1.23$, $MSE = 8.00$, $p > .05$. The same was true for general reading comprehension skill, $F(3, 57) = 1.34$, $MSE = 1.72$, $p > .05$. Post-Hoc comparisons (using Fisher's LSD) confirmed that there were no differences among the conditions (p 's $> .10$). These results indicate that all the conditions had comparable levels in the control variables.

3.5.2. Retention test. The conditions scored equally in the retention test. An analysis of variance revealed that the modality of the explanatory aid was not significant, $F(1, 57) = 1.72$, $MSE = 2.36$, $p > .10$ on this measure. This means that the participants receiving visual explanatory aids did not recall basic concepts about plate tectonics better than those who received auditory explanatory aids. This is not consistent with the results in the Experiment 1 and it will be discussed below. An analysis of variance showed that the modality of the regulatory aid was not significant, $F(1, 57) = 1.09$, $p > .10$. This indicates that the participants receiving the auditory regulatory aid did not recall basic concepts better than the participants receiving the visual regulatory aid. The interaction between the two factors was not significant, $F(1, 57) = 0.79$, $p > .10$, indicating that these results were not qualified by any interaction.

3.5.3. Transfer test. Concerning the transfer test, the participants in the visual explanatory aid conditions outperformed those in the auditory explanatory aid conditions. An analysis of variance confirmed this observation, $F(1, 57) = 22.91$, $MSE = 6.23$, $p < .001$, $\eta^2 = .29$. This indicates that the participants in the visual explanatory aid conditions were more able to apply the knowledge they had acquired than did those in the auditory explanatory aid conditions. An analysis of variance showed that there were no significant differences between the conditions as a function of the modality of the regulatory aid, $F(1, 57) = 2.10$, $p > .10$. That is to say, the participants receiving the auditory regulatory aid were not better than those receiving the visual regulatory aid in applying the knowledge they had acquired to solve the transfer tasks. The interaction between the two factors was not significant either, $F(1, 57) = 0.10$, $p > .50$, this meaning that the results we have just reported were not qualified by any interaction.

3.5.4. Time-recordings. The participants in the visual regulatory aid conditions spent less time ($M = 25.98$, $SD = 4.18$) viewing the regulatory aids than those in the auditory regulatory aid conditions (33 seconds, $SD = 0$). Those in the visual explanatory aid conditions took less time ($M = 47.80$ seconds, $SD = 5.24$) viewing it than those in the auditory explanatory aid conditions (66 seconds, $SD = 0$). The differences were confirmed by non-parametric Mann-Whitney's U -tests (all p 's $< .05$).

3.6. Discussion.

Overall, we have gathered support for the visual explanation superiority hypothesis but not for the auditory regulation superiority hypothesis. The former result is consistent with the

findings of the Experiment 1 and with our predictions. The latter is also consistent with the findings in Experiment 1 but not with our expectations.

The participants receiving visual explanatory aids outperformed those receiving auditory explanatory aids in the transfer test. We predicted that the explanatory aids would be comprehended better in visual modality, since this makes it possible to control the presentation. This is what has been found and it is in line with the results found before. As was done in the Experiment 1, this result will receive further consideration. The visual advantage would be explained by three alternative arguments. One possibility is that the participants receiving visual explanatory aids took more time viewing them than those in the auditory explanatory aid conditions. Again, the time-recordings showed that the former took no more but less time (25% less time, approximately). Another possibility is to think that these participants read the full explanatory aid twice. This might be possible but seems somewhat unlikely, as they spend 339 milliseconds (47,800 milliseconds / 141 words) per word on average. Finally, it could be argued that these participants freely managed the time available, focusing on the segment they wanted. This is consistent with patterns of visual movements during reading (Just et al., 1982; Rayner et al., 2005).

Despite the advantage of the visual explanatory aid in the transfer scores, there was no effect in the retention scores. This needs close examination. Transfer tests are designed to elicit the use of the knowledge acquired in a productive way. They pose novel and problematic situations and learners must reason about them using the knowledge they have gained. In sum, transfer tests elicit problem-solving processes on learners. This is different from the process that is elicited by retention tests. Retention tests ask learners to recall basic information. This makes it possible for learners to reproduce what they have learned without deeply reflecting on it. Therefore, retention tests should not be considered measures of deep learning -at least in a strict sense (see for example Kintsch, 1994 for a more thorough discussion). On the contrary, transfer tests provide learners with tasks in which they have to demonstrate whether they have acquired a deep learning. The aids we introduced in the CBLE were intended mainly to promote deep learning. This is the reason why the effects might be observable only in deep learning measures. Moreover, this is consistent with prior findings: there are experiments in which the impact of aids led to an increase in performance in deep learning tasks but not in other tasks (e.g., Graesser et al., 2004; McNamara, Kintsch, Songer, & Kintsch, 1996; McNamara & Kintsch, 1996; Vidal-Abarca, Martínez, & Gilabert, 2000).

As was done in the Experiment 1, prior results in this field can be reconciled. Although in the experiments of Moreno and her colleagues (Moreno & Mayer, 2002; Moreno et al., 2001) and those of Atkinson (2002) it was observed that visual explanatory aids are not better than identical aids presented in the auditory modality, a reinterpretation can be made. The explanatory aids used in those experiments were simple (as demonstrated in the Appendix (A)), which can explain the absence of visual advantage. Consistent with research on text comprehension, only when complex rather than simple texts are used, the visual gains advantage over the auditory modality (Green, 1981). The explanatory aid in our Experiments 1 and 2 was complex and, accordingly, it was best presented in the visual instead of the auditory modality. Regarding the results in Graesser et al.'s experiment (2003) the only possible explanations from our research framework (and the results in our Experiments 1 and

2) are those included in the discussion of the Experiment 1.

The visual advantage in this Experiment 2 also represents evidence which extends the results in Green (1981). Our results are in line with her findings and extend them because she conducted only one experiment and she used expository texts instead of verbal explanatory aids introduced into a CBLE.

In this experiment it was not possible to find support for the auditory regulation superiority hypothesis, as happened in the Experiment 1. The auditory modality was enriched by adding visibility and physical presence to the agent. It was hypothesized that this would make the additional information delivered by the auditory expressiveness more trustworthy and, hence, more effective. But this was not the case. In order to explain these results, several explanations may be suggested. One possibility is that the superiority of the auditory regulatory aids might be only true when learners are facing very demanding to-be-learned materials. When dealing with very demanding materials (e.g., CBLEs involving multiple modules), the monitoring skills become even more relevant: there are a lot of decisions to make (“Should I revise that?” “Should I integrate those two concepts?”). In those situations regulatory aids become particularly helpful. As a consequence, the modality of presentation of these aids may become more relevant as well. This leaves the auditory regulation more room for gaining advantage from the visual regulation. Another possibility is that the superiority of the auditory regulatory aids is true only when learners have low self-regulation and literacy skills. Learners with low self-regulation skills find it particularly difficult to monitor their understanding. Hence, regulatory aids should be very helpful for them. To the extent that regulatory aids are more relevant the same rationale presented before can be invoked: the auditory modality has more room for gaining advantage from the visual modality. With respect to literacy skills, the rationale behind the hypothesis is the following. Regulatory aids can be conceived as *structural cues* (explicit instructions to learners to take certain actions in order for them to build a coherent mental representation from the materials; Britton, 1994; Gernsbacher, 1997; Goldman & Rakestraw, 2000) which demands learners to have specific knowledge to recognize and use these cues. The question is that visual regulatory aids require learners to have that knowledge but auditory regulatory aids do not: the auditory expressiveness embodies information about what the cues are and how to use them. These two possibilities were tested on the Experiments 3 and 4. Because of that, they are explained in detail below.

Prior research on agents' visibility has shown that seeing the agent would have a neutral effect on learning. Here it was argued that this was the case because the agents were providing learners with explanatory aids rather than regulatory (Atkinson, 2002; Moreno et al., 2001) or giving both types of aid without being aware of it (Baylor and Ryu, 2003; Graesser et al., 2003). Results did not confirm our argument. Overall, this could indicate that agents' features are less important than we expected. Further work is needed to investigate this possibility.

4. Experiment 3: Testing the superiority of the auditory regulatory aids with demanding materials.

From the functional approach it is argued that there are different types of verbal aids depending on their function. One of those types, the regulatory aids, consists of devices asking learners to do something (i.e., revise their understanding) in order to help them in monitoring their learning process. When receiving regulatory aids, learners have to decide whether or not it is worthy to perform the action the aids expect them to do (i.e., revise their understanding). In deciding so, the expressiveness of the auditory regulatory aids can be helpful, as it embodies the speaker's intention (i.e., request) and attitude towards what he/she is saying (i.e., he/she considers it very important). Therefore, one prediction from the framework is that auditory regulatory aids are better than identical aids presented in the visual modality. However, support for this prediction could not be gathered in two previous experiments. The goal of the Experiment 3 was to test this prediction under certain conditions.

In this experiment it was argued that the complexity of the to-be-learned materials would make the auditory expressiveness of the regulatory aids more relevant. More accurately, complex materials will demand learners to make a lot of decisions being more difficult for them to regulate their learning on their own (“Did I understand that notion?” “Should I read this passage again?” “Are these two notions interrelated?” and so forth). In these circumstances providing regulatory aids would be particularly helpful. The point is that this situation (in which regulatory aids are very pertinent) would leave the auditory regulatory aids more room for gaining advantage from the visual regulatory aids. Accordingly, the participants receiving auditory regulatory aids were predicted to outperform those receiving visual regulatory aids.

In keeping with this rationale, the complexity of the to-be-learned materials used in the experiment was increased with respect to that used in prior experiments. This was done by including more modules into the to-be-learned materials (see below). This required the inclusion of new questions in the prior knowledge, retention, and transfer tests. Moreover, as the number of events described in the CBLEs increased and so did the questions in the transfer tests, it was possible to distinguish between near and far transfer in the analyses.

It is worthy of consideration that exploring the differences between the modalities in which the regulatory aids are presented is relevant if and only if these aids are effective. For this reason it is essential to gather support for the effectiveness of the regulatory aids before testing the impact of their modality. This was done in other experiments (e.g., Atkinson et al., 2003; Moreno & Mayer, 2005; Schworm & Renkl, 2003) and also (indirectly) in the Experiment 1. Nevertheless, we considered appropriate to conduct an additional test exploring this for some reasons. First, as the complexity of the materials was increased (with respect to our prior experiments), the number of aids increased accordingly. This means there were new aids not previously tested. Second, although the results from different experiments provide evidence in favor of the effectiveness of the regulatory aids, this does not necessarily mean that regulatory aids were effective in this very experiment (with the participants and materials used here). A more stringent test is that including a control condition.

4.1. *Participants and design.*

Thirty-nine undergraduate students enrolled in educational psychology courses at the University of Salamanca (Spain) participated in this experiment. They were randomly assigned to one of two conditions. Thirteen participants served in the auditory regulatory aids condition (ARA), 13 in the visual regulatory aids condition (VRA), and 14 in the control condition (with no regulatory aids). The mean age of the sample was 22. The sample was 80% female (although there was a majority of women, genre does not seem to be a critical variable, as observed in similar studies (Buisine & Martin, 2007; Moreno & Flowerday, 2006)). All the participants had experience in using computers.

The experiment had a one factor design with the modality of the regulatory aids (visual/auditory) as the between-subjects factor and also involved a control condition (without regulatory aids). Retention test and (total/near/far) transfer test scores were used as the dependent variables. Prior domain knowledge and general reading comprehension skill were used as control variables.

4.2. *Materials.*

For each participant materials consisted of a prior knowledge test, a general reading comprehension skill test, the CBLEs, a retention test, and a transfer test. All these materials are described next.

4.2.1. *Prior knowledge test.* The prior knowledge test consisted of a paper-and-pencil test comprising seven open-ended questions. The questions in this test were the following. “Explain how volcanoes are formed.” “Is it possible for continents to be permanently moving? Explain why.” “How are mountains formed?” “Is it possible for the Earth surface to be recycled somehow?” “Why do some mountains have volcanoes whereas others do not?” “What is a tectonic plate?” “Where are plates located in the Earth's internal structure? Choose one of these possible configurations.”

4.2.2. *General reading comprehension skill test.* The general reading comprehension skill test was identical to that used in the Experiments 1 and 2.

4.2.3. *Computer-based learning environment.* The CBLE included both (a) animations with concurrent narrations and (b) aids. It is worthy of consideration that the number of modules (and their corresponding interrelations) or events covered by the materials was increased, as compared to the prior experiments. The animations with concurrent narrations described several events concerning plate tectonics, namely, the three layers in the internal structure of the Earth and their relations, the convection currents, the ridges and the process through which new crust is created, the convection currents as the origin of plates' movements and collisions, the collision between a continental and an oceanic plate and its consequences on the Earth's surface, the Andes range as an example of this type of collision, the collision between two continental plates and its consequences on the Earth's surface, the Himalaya range as an example of this type of collision, the destruction of crust in the trenches. The narrations and animations lasted 583 seconds (10 minutes approx.). They were

presented in a sequential (instead of hyperspatial) fashion. The mental model we wanted the participants to build from the animations and narrations was the following. “The Earth's core is a big sphere made of iron and is very warm. The core makes the magma (i.e., a doughy substance made of melt rocks) in the mantle heat up and, thus, approach the Earth's surface or crust. When the magma gets cold again, it moves away from the crust. As a result, the convection currents are formed: the magma is continuously moving up and down. The magma in the mantle surfaces through the ridges (i.e., breaks in the crust) because of the movement the currents provoke. Once it is in the surface, it gets colder and solidifies creating new crust. The convection currents also push the plates, as they are floating on the magma, making them move away from each other. When moving, the plates can collide with other plates. There are different kinds of collisions depending on the plates engaged in the crashes. The plates can collide and one can sink (in the trenches) inside the Earth originating mountains with volcanoes (e.g., the Andes range); or they can collide and move vertically originating mountains without volcanoes (e.g., the Himalaya range). When a plate sinks in the trenches it is destroyed and becomes magma again whereas the magma surfacing through the ridges solidifies creating new crust. This means there is a continuous recycling loop between the trenches and the ridges. (see also the Appendix (B).)”

The CBLE also included aids helping the participants to detect and repair possible problems of understanding. As was explained before, common misunderstandings were identified on the basis of prior studies (e.g., Sánchez et al., 2008). Aids were provided to assist the comprehension of those aspects typically misunderstood. There were three aspects typically simplified or mixed-up by the participants in prior experiments. First, the notion of ridge was distorted. A number of participants thought of ridges as isolated and small cracks instead of large cracks all over the Earth's surface; this has important implications. Second, the specific features of the Andes and the Himalaya plates' collisions were mixed-up (as was explained above). Many participants underestimated the particular characteristics of each type of collision, building a mixed model in which features of both types of collision were included. Finally, most of the participants could not grasp the idea of the recycling loop linking the activity in the ridges and that in the subduction. They grasped either the former or the latter part of the relationship but not the link between them. Accordingly, three aids were created involving two parts each. All the aids had a regulatory part (except for the control condition), which helped the participants to monitor their learning process to detect misunderstandings, and an explanatory part, which assisted the participants in revising and repairing their flawed mental representations. When processing the regulatory part, the participants had to understand how important was to check their mental representations and to start doing it. When receiving the explanatory part, the participants have to process carefully an elaboration on the to-be-learned contents. An example of these aids is the following (the one transcribed in the Experiment 1 is another example):

- [The CBLE poses a question to the participant] “The crust is destroyed in the subduction process, why does it remain the same size? Choose one of these answers: (a) magma from the volcanoes compensates for the loss, (b) new crust is created in the trenches, (c) new crust is created in the ridges” [The participant chooses (a)]
- [Regulatory aid] “It is true that from the volcanoes magma comes to the Earth's surface and solidifies so that it may be seen as new crust but, no, this is not the correct answer.

The magma from the volcanoes cannot compensate for the loss as it is very little and eruptions take place irregularly.”

- [Explanatory aid] “The new crust is actually created in the ridges. The magma surfaces through the ridges, solidifies, and creates new plate material. In so doing, the ridges compensate for the loss in the subduction process. Therefore, there is a recycling loop in the Earth's surface: new crust is created in the ridges and destroyed in the trenches.”

(The participants who chose the correct answer were told that, anyway, a more precise response was needed and, therefore, were redirected to the explanatory aid: So, in certain sense, all the possible answers were considered not completely correct).

Depending on the experimental condition, the participants received either both explanatory and regulatory aids or only explanatory aids. The participants in the ARA and VRA conditions were provided with both regulatory and explanatory aids; those in the control condition received explanatory aids but no regulatory aids at all.

Moreover, depending on the condition, the participants received the regulatory aids either in the auditory modality (the ARA condition) or the visual modality (the VRA condition). In order to develop the auditory sections, an experienced teacher was recruited (Andrea Bustos). She recorded the auditory regulatory aids and the resulting audio archives were attached to the CBLE. She was instructed to employ the auditory expressiveness. Two judges selected those recordings they considered more natural and convincing. Criteria for considering a recording convincing were that it must clarify the speaker's communicative intention and provide additional clues about his/her attitude toward the message. More accurately, a convincing recording was one clarifying that the communicative intention was to warn about a misunderstanding and showing an attitude of urgency and importance (i.e., the speaker stressed the urgency of revising the misunderstanding and the seriousness of the problems related to thinking that way). This would make learners to process the subsequent explanatory aid more carefully.

4.2.4. Retention test. The retention test consisted of a paper-and-pencil test including five open-ended questions. The questions were the following. “Why are tectonics plates moving and crashing permanently?” “What does it mean crust is being recycled continuously?” “What is a ridge?” “What are the convection currents?” “What are the differences between the Andes and the Himalaya plate collisions?”

4.2.5. Transfer test. The transfer test consisted also of a paper-and-pencil test and comprised nine open-ended questions. The questions in this test were the following. Six questions were considered near transfer questions because they addressed aspects of the plate tectonics theory directly supported by the aids. “Imagine that ridges are not long cracks all over the crust but isolated holes, what would happen?” “Would it be possible for the Himalaya chain to have volcanoes in the future?” “There are a number of volcanoes in Italy, why does Spain have no volcanoes?” “Imagine the volcanoes in the Andes stop erupting, what would had happen?” “What would you expect if the subduction process destroys more crust than that is created in the ridges?” There was also another question including an illustration in which the participants had to place a volcano and explain why it was there. Three questions were

categorized as far transfer questions because they tested on aspects not directly supported by the aids but demanded the participants to mentally run the complete plate tectonics theory to answer the questions, which was only indirectly supported by the aids. “How could you explain that plates stop moving?” “Imagine that the convection currents start moving half their speed, how would you explain that?” “Imagine that the convection currents stop working, what would you think had caused that situation?”

4.3. Procedure.

The procedure was very similar to that used in the Experiments 1 and 2. The participants were tested simultaneously in groups of 10 to 15 participants per session. Each participant was seated in front of his/her individual computer and headphones. First of all, the participants received some basic instructions. After receiving the instructions, prior knowledge tests were delivered and the participants solved them in silence. It took no more than 12 minutes. When everyone finished filling in the prior knowledge test, the participants started using the CBLE. An experimenter ran the computerized materials on the participants' computer. The experimenter randomly ran the version of the CBLE so that participants were randomly assigned to each condition. The CBLE lasted about 900 seconds (15 minutes) including the animations with narrations (583 seconds, 10 minutes approx.) and the aid episodes. The participants viewed the animations with concurrent narrations describing the events listed before about the plate tectonics theory. In addition, they received the three aids including the regulatory and the explanatory parts. These aids were intended to assist the participants in detecting and repairing misunderstandings, as was explained above.

Whenever a regulatory aid was presented auditorily, the participants could use it as follows. The audio recordings were played once. The first auditory regulatory aid lasted 27 seconds, the second one lasted 26 seconds, whereas the third one lasted 27 seconds. The audio recordings were created by an experienced teacher and they were judged as natural and convincing by judges, as explained before. The recordings were intended to show maximum expressiveness.

Whenever a regulatory aid was presented visually, the participants could use it as follows. It was allowed for the participants viewing the aids on the screen to take as long as they wanted in doing so. The whole text was presented in an all-at-once fashion on the screen.

The explanatory aids were presented visually on the screen (also in all-at-once fashion). It was allowed for all the participants (that is, in the ARA, the VRA, and the control conditions) to view these explanations as long as they wanted.

According to the above, the participants in the ARA condition listened to the regulatory aids through their headphones. The audio recordings were played once. After each regulatory aid, these participants viewed the explanatory aids on the computer screen. They could freely manage their time when doing so. The participants in the VRA condition viewed the regulatory aids on the computer screen. Then they viewed the explanatory aids. They could freely manage their time when doing so. The participants in the control condition viewed the explanatory aids on the computer screen freely managing their time.

After using the CBLE, the participants were given the retention test. Solving this test took no more than 10 minutes. Then they were given the transfer test and they had no more than 15 minutes to solve it. Finally, they were tested on their general reading comprehension skill. After this final test was collected, the participants were seen off. Each session lasted about 75 minutes.

4.4. Scoring.

A rater scored all the tests unaware of the condition of each participant. Approximately 30% of the tests were also scored by a second rater. Interrater agreement was above .81 in all the tests. Disagreements were solved by consensus.

A template with possible answers was developed for the prior knowledge, retention, and transfer tests. It included accurate, correct but incomplete, and incorrect answers for each question. These answers yielded 2, 1, or 0 points, respectively.

4.4.1. Prior knowledge test. Accurate answers for the question about volcanoes' formation were those stating that in order for volcanoes to be formed two plates must collide, one sinking under the other and putting pressure on it to create cracks through which the magma surfaces. Accurate answers for the question concerning the movement of continents include stating that continents move horizontally because they are part of tectonic plates that, in turn, are floating on magma, which is moving permanently. Accurate answers for the question about mountains' formation were those explaining that for mountains to be formed the tectonic plates must collide and push each other folding up crust material. Accurate answers for the question on the recycling of crust were those explaining that crust can be recycled, as it is continuously created and destroyed by different processes. Accurate answers for the question asking why only some mountains have volcanoes included stating that for mountains to have volcanoes one of the plates involved in the collision must sink under the other and push it to produce cracks through which magma surfaces. Accurate answers for the question about the notion of tectonic plate were those stating that plates are the blocks of the Earth's surface, which are floating on magma and can collide each other provoking multiple geologic phenomena. The correct alternative in the multiple-choice question about the position of tectonic plates was that showing the plates as parts of the Earth's surface. Total scores ranged from 0 to 14 in the prior knowledge test.

4.4.2. General reading comprehension skill test. Each right answer in the general reading comprehension skill test yielded 1 point. Wrong answers yielded 0 points. Total scores ranged from 0 to 8.

4.4.3. Retention test. The accurate answers for the retention test are described next. Accurate answers for the question on plates' movement and crashing were those stating that plates move because they are floating on magma, which in turn is moving permanently (due to the convection currents), and that plates collide because sometimes two moving plates converge. Accurate answers for the question on the recycling loop were those mentioning both the destruction of crust in the trenches and the construction of crust in the ridges.

Accurate answers for the question about the ridges included stating that they are places where the Earth's surface divides into plates and where the new crust is created. Accurate answers for the question about the convection currents included explaining that these currents are movements in the mantle, which carry the magma (and, hence, the plates) and produced by the difference on the temperature of the magma near the Earth's core and the magma near the Earth's surface. Accurate answers for the question about the differences between the Andes and the Himalaya were those mentioning the plates involved in each case (i.e., oceanic-continental and continental-continental, respectively), the processes (i.e., subduction and mutual pressure), and the consequences on the Earth's surface (i.e., mountains with volcanoes and mountains without volcanoes). Total scores ranged from 0 to 10 points in the retention test.

4.4.4. Transfer test. The accurate answers for the transfer test were the following. Accurate answers for the question on ridges as holes rather than cracks included stating that in that situation the crust could not be recycled, as the source of new plate material would be too small. Accurate answers for the question about volcanoes in the Himalaya included reasoning that it is not possible for the Himalaya to have volcanoes because there is no oceanic plate sinking, and, hence, no pressure on the continental plate provoking cracks through which the magma can surface forming volcanoes. Accurate answers for the question about volcanoes in Italy but not in Spain included arguing that for a mountain to have volcanoes one of the plates must sink under the other, put pressure on it, and create cracks; something that is possible in Italy because of the presence of both continental and oceanic plates but not in Spain because of the absence of oceanic plates. Accurate answers for the question on eruptions in the Andes range were those stating that for them to halt the collision between the plates should stop so that there would be no pressure of the sinking plate over the plate above and, hence, no cracks through which the magma can surface. Accurate answers for the question about an increase in the process of crust destruction were those predicting the crust to disappear so that the mantle would be exposed to the environment. Accurate answers for the question asking about the causes to the end of plate crashes were those mentioning that for this to happen the convection currents should stop, this being due to the cooling of the Earth's core. Accurate answers for the question on the speed of the convection currents were those stating that for the currents to speed down the temperature of the Earth's core should decrease, as it is the origin of the movements in the mantle. Accurate answers for the question about the convection currents stop working were those explaining that for that to happen the Earth's core should get cold, being impossible to heat up the magma to make it approach to the crust. Accurate answers for the question asking the participants to locate a volcano in an image were those locating the volcanoes on the continent near the trenches and explaining that those are the places where continental and oceanic plates collide. Total scores ranged from 0 to 18 points in the transfer test. Total scores ranged from 0 to 12 in the near transfer test and from 0 to 6 in the far transfer test.

4.5. Results.

The variables under analysis were prior knowledge, general reading comprehension skill, retention, and (total/near/far) transfer tests. Prior knowledge and general reading comprehension skill were analyzed using a one-way analysis of variance with condition

(ARA/VRA/control) as the between-subjects factor in order to ensure that all the conditions had the same level in these two variables. Retention and transfer tests and total performance scores were analyzed using an independent samples *t*-test comparing the performances of the condition including regulatory aids (ARA) with those of the condition with no regulatory aids (control). This was done as a preliminary analysis in order to ensure that the regulatory aids were indeed effective. The ARA condition was used for these contrasts instead of the VRA condition; this was done this way because the ARA condition was expected to be better than the VRA condition. Retention and transfer tests scores were also analyzed using an independent samples *t*-test comparing the performances of the condition with auditory regulatory aids (ARA) with those of the condition involving visual regulatory aids (VRA). This was done to test the predicted superiority of the auditory regulatory aids over the visual regulatory aids. All the scores are showed in Table 3.

4.5.1. Control measures. An analysis of variance was conducted to check whether the conditions had the same level in these measures. Regarding prior knowledge, there were no significant differences between the conditions, $F(2, 37) = 1.230$, $MSE = 6.907$, $p > .10$. With respect to general reading comprehension skill, once again there were no significant differences, $F(2, 37) = 1.298$, $MSE = 1.838$, $p > .10$. Post-Hoc pairwise comparisons (using Fisher's LSD) confirmed that there were no differences between the conditions on these measures (all p 's $> .10$). These results indicate that the participants in the three experimental conditions exhibited similar levels of prior knowledge and general reading comprehension skill.

Table 3. Means (*M*) and standard deviations (*SD*) of all the conditions in all the control and learning measures: Experiment 3.

	ARA		VRA		Control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Prior Knowledge	4.45	2.28	4.31	3.38	3.02	2.07
G. Reading C. Skill	5.15	1.57	5.69	1.11	4.85	1.34
Retention	6.38	1.86	6.00	1.91	4.29	2.26
Transfer	5.23	2.20	6.23	2.35	3.57	2.44
Near Transfer	3.23	1.64	3.92	1.71	2.43	1.83
Far Transfer	2.00	1.08	2.31	0.95	1.14	1.03

Note. The maximum scores were 14 for prior knowledge, 8 for reading comprehension skill, 10 for retention, 18 for total transfer, 12 for near transfer, 6 for far transfer.

4.5.2. Preliminary analyses. Preliminary analyses on the effectiveness of the regulatory aids are presented next. With regard to the retention test, the participants in the regulatory aids condition (ARA) scored higher than their counterparts (control). This difference was reliable, as indicated by an independent samples *t*-test, $t(25) = 2.62$, $p < .05$, $\eta^2 = .22$. This means that the participants in the condition including regulatory aids were more able to recall the key concepts of the plate tectonics theory, as compared with those in the control condition. The scores on the (total) transfer test were in the expected direction, with

the regulatory aids condition performing better than the control condition. However, the analysis showed that there were not significant differences but a trend, $t(25) = 1.85$, $p = .08$, $\eta^2 = .12$. Regarding the near transfer test, the participants in the two conditions scored similarly. The t -test confirmed this, $t(25) = 1.20$, $p > .10$. This indicates that the participants in the regulatory aids condition were not better than their counterparts in applying their knowledge to generate solutions to the problems posed in the near transfer test. On the contrary, the participants in the regulatory aids condition outperformed those in the control condition in the far transfer test, as indicated by an independent samples t -test, $t(25) = 2.11$, $p < .05$, $\eta^2 = .14$. This indicates that the participants in the regulatory aids condition were more able to bring into play the knowledge they had acquired to solve novel tasks, such as those comprised by the far transfer test. Overall, the regulatory aids were effective in promoting a complete and correct understanding of the plate tectonics theory. As long as the regulatory aids proved to be effective, it seemed reasonable to explore the impact of their modality on learning.

4.5.3. Retention test. With respect to the retention test, the participants in the ARA condition scored almost equal than those in the VRA. Not surprisingly, an independent samples t -test revealed that the modality of the regulatory aid was not significant, $t(24) = 0.52$, $p > .50$. This indicates that the participants receiving the auditory regulatory aids recalled basic concepts about plate tectonics equal than those who received the visual regulatory aids. This was not in line with our expectations.

4.5.4. Transfer test. The scores in the (total) transfer test were similar between the ARA and the VRA conditions. An independent t -test showed that the modality of the regulatory aid was not significant, $t(24) = -1.12$, $p > .10$. Identical result was found with regard to the near transfer test, $t(24) = -1.06$, $p > .10$. The same goes for the far transfer test, $t(24) = -0.77$, $p > .10$. This means that participants receiving the auditory regulatory aid were not more able to apply the knowledge they had acquired to solve the (near/far) transfer tasks as compared with the participants receiving the visual regulatory aid. Overall, the regulatory aids were not best when presented in the auditory modality, as compared with their performance in the visual modality. This was not consistent with our predictions.

4.6. Discussion.

Within the research framework described above, the auditory regulatory aids were predicted to be better than the visual regulatory aids. However, the results in this experiment did not confirm the prediction. It was assumed that, because the auditory expressiveness embodies the intentions of the speaker and his/her attitudes of towards what he/she is saying, it is easier for learners to decide whether performing the action the aid is asking them to do and, hence, to fully profit from these aids. Moreover, we hypothesized that this pattern would be more apparent when using demanding to-be-learned materials. The question now is why this rationale was not supported by the data.

There are several possible explanations for this unexpected result. One explanation for the unexpected results in the Experiment 3 may be that the auditory modality of the regulatory aids requires not only expressive features but additional clues to become trustworthy and,

hence, effective. What kind of additional clues? One possibility to reinforce auditory expressiveness is to provide learners not with an auditory recording attached to the CBLE but an actual human agent. This human agent would be both visible and present in the experimental session showing his/her communicative intention and attitude in a more trustworthy fashion. This procedure was used in the Experiment 2 failing to gather confirmation for the auditory regulation superiority hypothesis. In that experiment a human agent was both visible and physically present in the experimental session but those clues did not make the auditory regulatory aids better than the visual ones on any of the dependent measures. Despite this, in the Experiment 4 we returned to this question, since it could be also argued that extra clues of humanity might reinforce the auditory expressiveness of regulatory aids if young participants (who have low self-regulation and literacy skills, as will be argued later) are recruited. Visibility and presence was not relevant for adult learners, but what if young participants are used?

Another explanation may be that the auditory regulation advantage is only true for learners exhibiting low levels of self-regulation or literacy (or, at least, lower levels than the adults used in the Experiments 1, 2, and 3). The rationale is the following. On the one hand, because learners with low self-regulation lack the skills necessary to regulate their learning on their own, regulatory aids are particularly pertinent for them. This leaves more room to explore the impact of the modality of those aids: the more important they are, the more crucial their modality becomes. On the other hand, because regulatory aids are structural cues (e.g., Goldman & Rakestraw, 2000), they require learners to have knowledge in order to recognize them and profit from them. As a prerequisite to acquire that knowledge, learners need experience with instructional materials (Gernsbacher, 1997). Low literacy means that learners have no such an experience and, hence, not enough knowledge to recognize and use regulatory aids. But regulatory aids presented in the auditory modality prevent learners from having to apply that knowledge since the auditory expressiveness embodies speaker's intention and attitude making learners notice what the cues mean and how important it is to follow them. If this is true, then auditory regulatory aids should be better than identical aids presented in the visual modality for low literacy learners. This is what was tested in the experiment described next.

5. Experiment 4: Testing the superiority of the auditory regulatory aids in participants with low self-regulation and literacy.

The results in the Experiment 3 showed no advantage for the auditory regulatory aids with respect to the visual ones. In other words, increasing the complexity of the to-be-learned materials did not make the auditory expressiveness (of the auditory regulatory aids) more relevant. Therefore, the prediction of the functional approach under question remains unconfirmed.

A main change was introduced in the experimental design in order to gather support for that prediction. In the Experiment 4 young participants were intentionally recruited (aged 14 on average; the participants in the Experiments 1, 2, and 3 were aged 21-22 on average). This was done because these participants were assumed to have lower self-regulation and

literacy skills than the adult participants (such as those used in the Experiments 1, 2, and 3). There are some reasons to believe such a thing. First, there are a number of studies indicating that young learners exhibit lower self-regulation skills than adults (e.g., Greene & Azevedo, 2008; Hacker, 1997; Skarakis-Doyle, 2002; Sperling, Howard, Miller, & Murphy, 2002). For instance, Greene and Azevedo (2008) asked middle-school (aged 12 on average) and high-school students (aged 15 on average) to think-aloud while learning from a CBLE on the topic of biology. The results showed that the high-school students built more sophisticated mental models, which in turn depended on their ability to monitor their understanding; this suggesting that the high-school students were more able to monitor their learning process as compared with the middle-school students. Similarly, Hacker (1997) asked seventh, ninth, and eleventh-graders (from 11 to 17 years old) to look for “anything that did not fit” in texts crafted for having errors. The results showed that the higher the level of education, the more errors were detected. Taken in conjunction, these studies indicate that there is a relationship between age and self-regulation; in other words, the older the learner is, the more self-regulation skills he/she has.

Second, the adult participants in our previous experiments were students aged 21-22 on average whereas the participants in the Experiment 4 were students aged 14 on average. This means that the experience with instructional materials is seven years larger for the former participants. As has been argued, the more experience with instructional materials, the more knowledge about structural cues is gained (e.g., Gernsbacher, 1997). Therefore, an advantage of seven years equals more knowledge on the recognition and use of structural cues.

Finally, the difference between the students in the Experiments 1-2-3 and 4 lies not only on the years they have been studying but on the educational levels they have passed. Students aged 14 are involved in obligatory education whereas university students volunteer to enroll in higher education. More accurately, those studying at the university have passed two levels above obligatory education (within the Spanish system). Because of the obligatory nature of the education in lower levels, the students are exposed to a short range of instructional materials (e.g., Beck, McKeown, Sinatra, & Loxterman, 1991), this reducing the possibility of gaining experience with structural cues. The instructional materials in high education are more in number and more sophisticated (with respect to lower levels of education), this allowing university students to acquire more knowledge on the recognition and use of structural cues. Moreover, it could be argued that for learners to enroll in higher education, high levels of self-regulation are needed; this meaning that university students are likely to have more self-regulation skills than those in the high-school level (in fact, there is a correlation between monitoring skills and academic achievement which, in turn, plays a key role in entering into higher education; e.g., Otero, Campanario, & Hopkins, 1992). For all these reasons, the participants in the Experiment 4 were assumed to have lower self-regulation and literacy than those in the Experiments 1-2-3.

The rationale behind the predictions in the Experiment 4 was the following. On the one hand, to the extent that young learners have less self-regulation skills than adults, they should find regulatory aids particularly helpful. As was argued in the Experiment 3, the more relevant the regulatory aids are in a given situation, the more room is left for the auditory

modality to show its advantage from the visual modality. On the other hand, to the extent young learners have less literacy than adults, they should be less able to profit from the regulatory aids presented in the visual form, as these demand learners to apply their knowledge about structural cues (Gernsbacher, 1997; Goldman & Rakestraw, 2000). As was argued before, literacy means having experience with instructional materials, which is a prerequisite to acquire knowledge about structural cues. Without such an experience, learners are not able to recognize and use regulatory aids as cues guiding their learning process. The point is that regulatory aids presented in the auditory modality do not demand learners having that knowledge because they embody speakers' intentions and attitudes, which automatically tells learners what the cues mean and how important it is to follow them: that is, what the cues are for (i.e., a request: "I want you to revise your understanding") and how they want learners to react (i.e., promptly: "It is critical for you to do it"). In line with this rationale, the participants receiving auditory regulatory aids were expected to outperform those receiving visual regulatory aids in this experiment.

In keeping with the rationale we have just described, the sample recruited was as youngest as possible. Ninth-graders were selected according to the following criteria: (a) they had the topic of plate tectonics as part of their official curriculum and (b) they had not studied it yet (the experiment took place right before their teachers started teaching on the topic).

Another important change with respect to the Experiment 1 was that the auditory expressiveness of the regulatory aids was reinforced. As was mentioned before, it is likely that for the expressiveness of the regulatory aids to become trustworthy, more clues are needed. These clues can be the visibility or the physical presence of the pedagogical agent. This argument is supported by the social agency theory (Moreno & Flowerday, 2006; Moreno et al., 2001) and the persona effect (Dehn & van Mulken, 2000; Johnson et al., 2000). Within this framework, the effectiveness of a pedagogical agent depends on the level of humanity he/she can display. The visibility and presence can make an agent more human and, hence, more effective. Nevertheless, there are evidence that visibility is less relevant than expected, even when the agents are not contents deliverers but aid providers (Atkinson, 2002; Baylor & Ryu, 2003; Graesser et al., 2003; Mitrovic & Suraweera, 2000; Moreno et al., 2001). Most important, the same was true in our Experiment 2. So why should we still believe in clues of humanity as something relevant? The hypothesis in the Experiment 4 was that those clues of humanity might be important for young learners. Young learners might need additional clues to find the auditory expressiveness of the regulatory aids actually trustworthy. That is the reason why in the Experiment 4 a human agent was used (instead of audio recordings attached to the CBLE).

As was done in the Experiment 1, a control condition (involving no regulatory aids) was included in the Experiment 2. This was done in order to gather evidence in favor of the effectiveness of the regulatory aids, which was important given that the participants changed with respect to prior experiments.

5.1. Participants and design.

Eighty-four ninth-graders studying at several high-schools in Las Palmas (Spain)

participated in this experiment. The participants volunteered to take part in the experiment (with parents' permission). They were randomly assigned to one of three conditions. Twenty-seven participants served in the auditory regulatory aids condition (ARA), 29 participants in the visual regulatory aids condition (VRA), and 29 in the control condition (including no regulatory aids). The mean age of the sample was 14. The sample was approximately 60% female. All the participants had experience in using computers.

As in the Experiment 3, this experiment had a one factor design with the modality of the regulatory aids (visual/auditory) as the between-subjects factor and involved also a control condition (with no regulatory aids). Retention test, and (total/near/far) transfer test scores were used as the dependent variables. Prior domain knowledge and general reading comprehension skill were used as control variables.

5.2. Materials.

Once again, for each participant materials consisted of a prior knowledge test, a general reading comprehension skill test, the CBLE, a retention test, and a transfer test. All these materials were identical to those used in the Experiment 3 except for the following changes.

5.2.1. General reading comprehension skill test. The general reading comprehension skill test consisted of a paper-and-pencil test comprising ten open-ended questions. The participants had to read a text on eskimos and their way of life and then solved some questions about it. It was extracted from a Spanish reading comprehension battery (i.e., PROLEC-SE, Ramos & Cuetos, 1999). The participants solved questions such as “Who was the first person to discover the way of life of eskimos?” or “On what do the eskimos rely to survive?”

5.2.2. Computer-based learning environment. The auditory regulatory aids (in the ARA condition) were not audio archives attached to the CBLE but were provided by a human pedagogical agent, who was both visible and present in the experimental session. The agent was an experienced teacher (Nadezhna Castellano) who was instructed to use her expressiveness. A convincing regulatory aid was that clarifying that the communicative intention was to warn about a misunderstanding and displaying an attitude of importance and urgency. Special attention was paid to make it in a natural way. The experimental session was video-recorded and two judges checked the recordings to ensure that (a) the agent was both natural and convincing when providing regulatory aids and that (b) the words uttered by the agent matched those included in the visual regulatory aids. This was important because the variable under consideration was the modality of the regulatory aids not the words they comprised.

5.3. Procedure.

The participants were tested simultaneously in groups of 15 to 20 participants per session. The experiment was developed in two sessions (for each group of participants). In

one session, the participants solved the prior knowledge and general reading comprehension skill tests and there was another session in which the participants used the CBLEs and solved the retention and transfer tests. This was different from the Experiment 3. In the Experiment 4 it was not possible to take all the tests in one session. It was so because of the high-school timetable of the participants. In the first session the instructions for the experiment were delivered. These were the same as in the Experiment 1. After receiving the instructions the participants solved the prior knowledge test in silence. This took no more than 12 minutes. Afterwards, they read the text about the eskimos and solved its corresponding questions, this being the general reading comprehension skill test. After these two tasks were solved, the participants were seen off. Each first session lasted about 30 minutes.

In the second session, each participant was seated in front of his/her individual computer and headphones. The experimenter randomly assigned a given group of participants to one of the experimental conditions. This was done this way because in the ARA condition the aids were given by a human agent present in the experimental session. The experimenter ran one of the versions of the CBLE. The participants used the CBLEs. This included watching the animations with concurrent narrations and receiving the regulatory and explanatory aids. The participants in the ARA condition listened to the regulatory aids from the human pedagogical agent. She was instructed to give the aids in a natural and convincing fashion, as explained before. After each regulatory aid, these participants viewed the explanatory aids on the computer screen. They could freely manage their time when doing so. The participants in the VRA condition viewed both the regulatory and explanatory aids on the computer screen. The participants in the control condition viewed the explanatory aids on the screen but received no regulatory aids. After using the CBLE, the participants were given the retention and transfer tests. After these tests were collected, the participants were seen off. Each second session lasted about 50 minutes.

5.4. Scoring.

The scoring was identical to that described above. Only the general reading comprehension skill test required different scoring instructions. Each right answer in the general reading comprehension skill test yielded 1 point. Wrong answers yielded 0 points. The test included criteria for assessing the answers. For instance, correct answers for the question about eskimos' survival were those stating that they rely on animals to get both food and clothes. Total scores ranged from 0 to 10.

5.5. Results.

The variables under analysis were prior knowledge, general reading comprehension skill, retention, and (total/near/far) transfer tests scores. Prior knowledge and general reading comprehension skill were analyzed using a one-way analysis of variance with condition (ARA/VRA/control) as the between-subjects factor in order to ensure that all the conditions had similar performances in these two variables. In the preliminary analyses, retention and transfer tests scores were analyzed using a Mann-Whitney's *U*-test comparing the performances of the condition including regulatory aids (ARA) with those of the control condition (including no regulatory aids). This was done in order to ensure that the regulatory

aids were effective. The ARA condition was used instead of the VRA condition in these comparisons because the former was expected to be better than the latter. Non-parametric tests (such as Mann-Whitney's *U*-test) were used because the assumption of normality was not met in the dependent measures (as indicated by the Kolmogorov-Smirnov test, p 's < .05). Retention and transfer tests scores were also analyzed using a Mann-Whitney's *U*-test comparing the performances of the ARA condition with those of the VRA condition. This was done to explore the impact of the regulatory aids on learning. All the scores are showed in Table 4.

Table 4. Means (M) and standard deviations (SD) of all the conditions in all the control and learning measures: Experiment 4.

	ARA		VRA		Control	
	M	SD	M	SD	M	SD
Prior Knowledge	3.07	2.74	2.38	1.90	2.48	1.74
G. Reading Compr. Skill	5.67	2.32	4.93	1.71	5.71	1.72
Retention	2.30	2.13	0.83	1.47	1.21	1.57
Transfer	4.81	3.34	2.14	2.61	2.43	2.51
Near Transfer	3.04	2.44	1.24	1.64	1.64	1.73
Far Transfer	1.78	1.40	0.90	1.26	0.79	1.26

Note. The maximum scores were 14 for prior knowledge, 10 for reading comprehension skill, 10 for retention, 18 for total transfer, 12 for near transfer, and 6 for far transfer.

5.5.1. Control measures. A one-way analysis of variance was calculated to ensure that the conditions had the same level in these measures. The participants performed almost equal in the prior knowledge test. Accordingly, an analysis of variance revealed that there were no significant differences between the conditions, $F(2, 81) = 0.835$, $MSE = 4.656$, $p > .10$. Regarding general reading comprehension skill, there were no significant differences either, $F(2, 81) = 1.477$, $MSE = 3.723$, $p > .10$. Post-Hoc pairwise comparisons (using Fisher's LSD) confirmed that there were no differences between the conditions on these measures (all p 's > .10). These results indicate that the participants in the three experimental conditions had comparable levels of prior knowledge and general reading comprehension skill.

5.5.2. Preliminary analyses. In the preliminary analyses the performances of the condition including regulatory aids (ARA) were compared to those of the control (with no regulatory aids) condition. The participants in the former condition scored better than their counterparts on the retention test. Not surprisingly, a Mann-Whitney's *U*-test showed that the effect of the regulatory aids was significant, $Z = -2.023$, $p < .05$, $\eta^2 = .08$. This indicates that participants receiving the regulatory aids recalled basic information about plate tectonics better than those who received no regulatory aids. With regard to the transfer test, the participants receiving regulatory aids scored better than those in the control condition, $Z = -2.700$, $p < .01$, $\eta^2 = .15$. The same pattern was found in the near transfer scores, with the conditions performing differently, $Z = -2.197$, $p < .05$, $\eta^2 = .10$. This indicates that the participants in the regulatory aids condition were better than their counterparts in applying

their knowledge to solve the near transfer questions. Identical result was found in the far transfer test. The differences between the conditions were reliable, as confirmed by a Mann-Whitney's *U*-test, $Z = -2.745$, $p < .01$, $\eta^2 = .13$. This indicates that the participants receiving regulatory aids were more able to bring into play the knowledge they had acquired to solve the far transfer questions, as compared with the participants in the control condition. Overall, evidence has been gathered supporting the effectiveness of the regulatory aids.

5.5.3. Retention test. The results with regard to the impact of the modality of the regulatory aids are presented next. The participants in the ARA condition scored higher than those in the VRA condition on the retention test. Accordingly, a Mann-Whitney's *U*-test revealed that the modality of the regulatory aid was significant on this measure, $Z = -3.015$, $p < .005$, $\eta^2 = .15$. This indicates that participants receiving the auditory regulatory aids recalled key information about plate tectonics better than those who received the visual regulatory aids. This is in line with the prediction we formulated.

5.5.4. Transfer test. The same goes for the transfer test, with the participants in the ARA condition outperforming those in the VRA condition, $Z = -3.151$, $p < .005$, $\eta^2 = .17$. The pattern was the same for the near transfer test, $Z = -2.927$, $p < .005$, $\eta^2 = .16$. The differences in the far transfer test were also significant, being in the expected direction, $Z = -2.440$, $p < .05$, $\eta^2 = .10$. Taken in conjunction, these results indicate that the participants receiving the auditory regulatory aids were more able to generate solutions to transfer tasks as compared with the participants receiving the visual regulatory aids. This was consistent with our predictions.

5.6. Discussion.

The preliminary analyses revealed that the regulatory aids were effective for these participants. In fact, they were more helpful than in the Experiment 3. This was indicated by the fact that significant differences between those receiving regulatory aids and those in the control condition were found in all the learning measures. This may be explained by stating that regulatory aids were more pertinent for young learners, as compared with the adult participants used in the Experiment 3. This was so because young learners have less self-regulation skills than adults, as argued before and reported in a number of studies (Greene & Azevedo, 2008; Hacker, 1997; Skarakis-Doyle, 2002; Sperling et al., 2002).

In this experiment support for the auditory regulation superiority hypothesis was gathered. It was hypothesized that, for learners with low self-regulation and literacy, the auditory regulatory aids provided by a human agent would be better than the visual regulatory aids. More accurately, low self-regulation learners were assumed to find regulatory aids particularly helpful and low literacy learners were assumed to have low experience with instructional materials and, thus, low knowledge about structural cues and how to use them. Because of that, they were predicted to profit from the auditory presentation of the regulatory aids, since regulatory aids presented in this modality (a) embody speakers' intentions and attitudes facilitating learners to profit from the aids and (b) prevent learners from having to apply that kind of knowledge. The argument was supported by the data, as indicated by the fact that the participants in the ARA condition scored higher than their counterparts on the

retention and transfer tests.

It should be pointed out that the same finding as the one found in the Experiment 4 has been found in another recent experiment. Bustos, Sánchez, and García (2007) asked sixth-graders to learn geography from a text. During the reading, some of the participants received auditory regulatory aids (from a human agent) helping them to monitor their understanding whereas others received similar aids presented in the visual modality (inserted into the text in written format). After reading the text, the participants solved retention and transfer questions. The results showed that the participants in the auditory regulatory aids condition outperformed those in the visual regulatory aids condition both on retention and transfer scores. These results are consistent with those presented here, indicating that they are robust instead of random.

There is at least one alternative explanation for the results in the Experiment 4, which would put our interpretation into question. It has been argued that the participants in this experiment had low literacy. For the purposes of the experiment this corresponded to a low level of knowledge about structural cues. However, one might also argue that low literacy means low decoding skills (i.e., being able to access the meaning of words in its written format with speed and accuracy). If this is true, then the auditory presentation of any kind of verbal message (regardless of being a regulatory or an explanatory aid) should produce better outcomes than the visual presentation, since the auditory modality prevents learners from having to decode written words. Accordingly, the participants in this experiment profited from the auditory presentation of the regulatory aids not because these were so (regulatory instead of explanatory aids) but because the participants would have profited from the auditory presentation of any verbal material. In order to rule out this possibility, the modality of the explanatory aids should have been manipulated. An advantage of the visual modality for the explanatory aids would rule out an interpretation based on the level of decoding skills -and in favor of the functional approach. But the modality of the explanatory aids was not manipulated. This manipulation was not made because there were (apparently) conclusive evidence that the visual explanatory aids are better than the auditory explanatory aids (Experiments 1 and 2). Although such a manipulation was not carried out, there are other ways of rejecting the interpretation based on decoding skills. First, there is evidence that readers of similar or even lower age (from third to seventh-graders) than those used here perform equally well either in reading or listening comprehension. This is indicated by a highly significant correlation between the performances under the two conditions (Smiley et al., 1977) or a comparable recall of the macrostructure of a text under the two conditions (Hildyard & Olson, 1982). If readers within that age range had problems in decoding then it should have been found that listening comprehension is better than reading comprehension. But the evidence does not indicate that. Second, the general reading comprehension skill of the participants in the Experiment 4 fell within normal levels. This was indicated by the fact that their performance on the general reading comprehension skill test was within the 50th percentile (Ramos & Cuetos, 1999). Difficulties in decoding words would have yielded poor performances on this test (as there is a high correlation between reading comprehension and decoding skill; see Perfetti, 2007), but it was not so. In sum, as it seems likely that the participants in the Experiment 4 had good reading comprehension (i.e., comparable to their listening comprehension), an interpretation for the results based on the assumption that the

participants had low decoding skills may be ruled out.

In reinterpreting Graesser et al.'s (2003) results, one of the possible explanations posed in the discussion of the Experiment 1 can now be ruled out. We argued that, as the modality was not relevant for their regulatory and explanatory aids, two possible explanations could be suggested. First, there were almost no explanatory aids whereas regulatory aids were pervasive. Given that in our Experiments 1-2-3 we found the modality of regulatory aids to be irrelevant, it was reasonable to believe that that proportion of regulatory aids in Graesser et al.'s experiment was true. The second explanation was that, although there were similar proportions of explanatory and regulatory aids, the type of aid was not crossed with modality rejecting the possibility of exploring their interaction. In the light of the results in our Experiment 4 the second explanation seems most plausible, as we have found that the modality of the regulatory aids is not necessarily irrelevant.

In the Experiment 2 we found the agent's visibility to be nonsignificant (even for regulatory aids). We concluded that maybe pedagogical agents' visibility was less important than one can expect. In the light of the results in the Experiment 4 this needs some refinement. In the Experiment 4, in which a human agent (both visible and present) was used, the regulatory aids provided by her were better than identical aids presented on the computer screen (this was confirmed in the Bustos et al.'s study (2007)). This indicates that, at least under certain circumstances, agents' visibility can have an impact on learning. More research is needed to clarify which agents' features and under which circumstances turn out to be effective.

Chapter 5.

Practical and theoretical implications.

0. Summary.

Learning from CBLEs is usually difficult for learners. This is true because of the demands learners have to deal with when doing so. These demands are, among others, that learners have to monitor their emerging understanding to detect problems of understanding and to generate explanations to repair those problems, which makes many learners get into trouble. Fortunately, researchers have developed aids to help learners to detect and repair their problems of understandings, proving them to be effective. Given that CBLEs allow designers to construct these aids in many ways (i.e., pictorial or verbal aids, auditory or visual verbal aids), one key question in the design of verbal aids is which is the best modality for their presentation. As was shown, there are not clear guidelines in this respect: the modality principle is not applicable and the available research on the impact of the modality of verbal aids is not conclusive. This circumstance called for a research framework. The functional approach, the research framework we presented, assumes that depending on the type of aid, one or the other modality is best.

The functional approach is based on several assumptions. First, each modality has specific advantages. The visual modality has control, which allows learners to re-read, regulate the pace of reading, and focus on the segment of the text they want. Control can be helpful under some circumstances, facilitating learners to acquire concepts from an instructional material (e.g., Rawson & Kintsch, 2005). The auditory modality has expressiveness, which embodies the speaker's intentions and attitudes (e.g., Brennan & Williams, 1995). Knowing the attitudes of others has an impact on own attitudes: if others are motivated towards a given task we are more willing to perform it (e.g., Wild et al., 1997). Second, there are different types of aid. Regulatory aids are those assisting learners to monitor their understanding. When processing regulatory aids learners have to decide whether or not it is worthy to do what the aids ask them to do (revise their understanding). Explanatory aids are those assisting learners to revise and repair their flawed mental representations. When processing explanatory aids learners have to assimilate elaborations on the to-be-learned contents. The third assumption is that there is an interaction between type of aid and modality of presentation. The regulatory aids are expected to be best presented in the auditory modality, whereas the explanatory aids are expected to be best presented in the visual modality. These

predictions are based on several arguments. The regulatory aids are those requesting learners to revise their understanding. Thus, when receiving these aids learners have to decide whether or not to perform that action. In so doing, knowing the communicative intention and the attitude of the speaker may be helpful (recall that others' attitudes makes us more willing to do a given task), something which is possible if the expressiveness of the auditory modality is displayed. The explanatory aids are those providing elaborations on the to-be-learned contents for the learners to revise their understanding. When receiving these aids learners may find it useful to control the text (recall that control facilitates acquiring concepts from a given material), something possible if a visual explanatory aid is presented. These predictions were called the visual explanation superiority hypothesis and the auditory regulation superiority hypothesis. These two hypotheses were tested empirically by means of a set of experiments.

The superiority of the visual explanatory aids over the auditory explanatory aids was confirmed in the Experiments 1 and 2. University students learned geology from a CBLE and received explanatory aids presented either visually or auditorily. Those receiving visual explanatory aids outperformed their counterparts both on retention and transfer scores. This was in line with prior research on text comprehension, which revealed that when learners read complex explanatory texts they learn more if the visual (rather than the auditory) modality is used for the presentation of the texts (Green, 1981). The Experiments 1 and 2, however, could not gather confirmation for the auditory regulation superiority hypothesis. That was true even when the regulatory aids were delivered by a human agent both visible and present.

In the Experiments 3 and 4 the goal was to test the remaining prediction of the functional approach: the expected superiority of the auditory regulatory aids over the visual regulatory aids. In the Experiment 3, university students learned geology from a CBLE including more modules than those materials used in the two prior experiments. This made the CBLEs more complex and, as a consequence, the inclusion of regulatory aids more pertinent. Because of that, the modality of the regulatory aids was expected to have more room for showing an auditory advantage over the visual modality. However, the results in the Experiment 3 did not confirm that argument. In the Experiment 4 ninth-grade students were recruited. It was argued that (a) because of their low self-regulation, they would find the regulatory aids particularly helpful and (b) because of their low experience in using instructional materials, they had not very much knowledge on the recognition and use of structural cues, such as the regulatory aids. For these reasons, it was expected that the auditory regulatory aids would be better than the visual ones, since the former facilitate learners to profit from the regulatory aids and do not demand learners to have knowledge about structural cues. Moreover, a visible and physically present agent was used for providing the aids. The prediction was confirmed, as indicated by the superiority of the participants receiving auditory regulatory aids on the retention and transfer tests.

It should be noted that the same result of the Experiment 4 has been obtained in another recent experiment. As was mentioned, Bustos et al. (2007) found that the participants in the auditory regulatory aids condition outperformed those in the visual regulatory aids condition both on retention and transfer scores. This suggests that the results in the Experiment 4 are robust.

1. Practical implications.

The results of our set of experiments have practical implications. In the light of the findings on the superiority of the visual explanatory aids, a new guideline for designers may be that if they want learners to fully profit from explanatory aids then they should present them visually. This allows learners to process these aids as deeply as it is required. According to the results in the Experiment 4, another guideline for designers may be that regulatory aids must be presented auditorily if low self-regulation and low literacy learners are recruited, at least if human (rather than virtual or no) agents are used. Taken in conjunction these two guidelines provide clear suggestions about how to present verbal aids in CBLEs.

Instead of trivial, these guidelines provide designers with clear answers for certain questions otherwise unsolved. Recall that the modality principle (Mayer, 2001; Sweller et al., 1998) was not suitable for verbal aids because these do not necessarily have related pictures and because verbal information accomplishes functions different from describing events shown by pictures (e.g., help learners in monitoring their understanding). Recall also that prior research did not provide conclusive findings (Atkinson, 2002; Graesser et al., 2003; Moreno & Mayer, 2002; Moreno et al., 2001). These are the reasons why those new guidelines are relevant.

2. A new look at prior research.

Regarding those mixed results in prior research on the topic, a reinterpretation can be made taking into account the findings from our set of experiments. In the experiment of Graesser et al. (2003) the participants were provided with multiple aids presented either in the visual or the auditory modality. The modality had no impact on learning. One seemingly plausible explanation may be that Graesser et al. were providing different types of aid (both regulatory and explanatory aids) without keeping this factor under control. Accordingly, the possible interaction between types of aid and modalities remained blocked yielding non significant results. It is not clear what would happen if the type of aid had been kept under experimental control. From the results presented here, one would expect the explanatory aids to be best presented visually and the regulatory aids to be best presented auditorily.

In the experiments of Moreno and her colleagues (Moreno & Mayer, 2002, Moreno et al., 2001) the auditory explanatory feedback was more helpful than the same feedback presented in the visual modality. This is not consistent with either the functional approach or the results in the Experiments 1 and 2. A possible explanation for that inconsistency is that the elaborations comprised in the explanatory feedback of Moreno et al.'s were somewhat simple. This circumstance would make the control of the visual modality irrelevant, as learners would not need to carefully process the elaborations. In fact, a close examination of the explanatory feedback that these researchers gave confirms this argument (see Appendix (A)).

In the experiments of Atkinson (2002) the modality of the explanatory aids was basically not significant. This is also inconsistent with either the research framework we proposed or our findings. One way to solve the problem is to assume the explanatory aids in the Atkinson's study to be simple -rather than complex. The analyses in the Appendix (A) also

confirmed that assumption. This would explain why the visual control was not so helpful for processing the explanatory aids.

The fact that the visual explanatory aids superiority hypothesis holds only for complex explanatory aids represents a constraint to what was stated above (and is consistent with the findings in text comprehension mentioned before). Accordingly, the guideline should be formulated in this way: use the visual modality (instead of the auditory modality) for the presentation of complex (instead of simple) explanatory aids. It is not clear what would happen if Atkinson or Moreno and her colleagues had used complex explanatory aids. From the findings we have found, one would expect visual explanatory aids to be better than auditory explanatory aids. To sum up, an interpretation for the mixed results in prior research may be that the distinction between regulatory and explanatory aids was not done, leaving this factor without experimental control, and that simple rather than complex explanatory aids were used, making the visual control less relevant than expected.

3. Theoretical implications.

3.1. The functions of language.

Another implication of the results presented here is that a discussion on the functions of language can be reinstated. In the experiments we have conducted, different patterns of results were obtained depending on the type of aid in question. This represents an empirical support for the assumption that there are different functions of language, each of which elicits certain reaction on the part of the hearer. Let us explain this in more detail.

The main function of language is allowing communication among individuals. Actually, language is both a system for representing knowledge and a system of communication among humans. Both functions have advantages. A system of representation makes it possible to control our knowledge (e.g., Karmiloff-Smith, 1992), whereas communication makes it possible to get vital information via second-hand and cooperation among interdependent individuals (e.g., Pinker & Bloom, 1990). But communication may be considered the main function of language, as there are signs that it was the critical function in the human evolution. Pinker and Jakendoff (2005) argue that if representation was the main function of language, then how one could explain why the vocal tract is tailored to the demands of speech (even compromising breathing and swallowing).

And, what communication is for? Communication, in turn, has different functions. One can use language for commenting on something (e.g., “I love watching the landscape from my window”), asking someone to do something (e.g., “Shut the window, please”), demanding some information (e.g., “Where is the window?”), apologizing (“Sorry, I broke your window”), and so on. The notion of illocutionary act (Austin, 1962; Searle, 1975) is commonly used for referring to all these uses of language. But all these functions of language (as a system of communication) can be reduced to two main categories. According to some linguists (Givón, 1984; Halliday, 1973; Werner, 1988a, 1988b) there are two types of illocutionary acts (i.e., two functions of language). There is a *manipulative* function consisting of (a) communicative acts a speaker uses to manipulate the behavior of others, which (b)

affects the intentional state of a hearer. There is an *informative* function consisting of (a) communicative acts a speaker uses to inform others about something, which (b) affects the informational state of a hearer. In keeping with this categorization, the acts of requesting and asking accomplish a manipulative function whereas stating and apologizing accomplish an informative function.

Now the question is that, according to our view, regulatory aids accomplish a manipulative function whereas explanatory aids accomplish an informative function. So there is a correspondence between the types of aid distinguished here and the (communicative) functions of language.

Overall, (a) if the distinction between regulatory and explanatory aids makes it possible to interpret a number of experimental results (otherwise uninterpreted) and (b) if we assume that regulatory aids accomplish a manipulative function whereas explanatory aids accomplish an informative function, then the distinction between the manipulative and the informative functions of language is supported. This means that the experimental results presented and discussed here may support a distinction made in the field of (psycho)linguistics (Givón, 1984; Halliday, 1973; Werner, 1988a, 1998b).

It should be noted that there are other evidence in line with this distinction. First, children rely mainly on the manipulative when starting to communicate with other humans (e.g., Bates et al., 1975/1998). This means that the pattern in the development of the manipulative and the informative functions is different in each case: manipulation is mastered when information is in an early stage. This, in turn, suggest they are different from each other.

Second, when higher primates (e.g., chimpanzees) are instructed to speak (by means of a symbol device or *lexigrams*), they rely almost exclusively on the manipulative function (e.g., Brakke & Savage-Rumbaugh, 1996). Again, this may suggest that manipulation and information are different functions. So the evidence we presented here may complement these evidence gathered in other fields.

3.2. *The differences between reading and listening.*

A final implication is that a discussion on the differences between the spoken and written modalities of language can be also reinstated. Within the *simple view of reading* (Gough, Hoover, & Peterson, 1996; Hoover & Gough, 1990), it is assumed that “beyond the point of word recognition, listening and reading appear to require essentially the same processes” (Gough et al., 1996: 2). Thus, according to Gough et al. the spoken and written modalities differ on the need for decoding, being comparable in other aspects.

From such a view it is not possible to explain why the spoken modality was best comprehended under certain circumstances in the present study. The simple view of reading would predict listening comprehension to be better than reading comprehension when decoding is not mastered; and would predict such an advantage irrespective of the type of text (i.e., regulatory/explanatory). In the Experiment 4 listening comprehension was better than reading comprehension for the regulatory aids in participants showing good levels of

decoding (as indicated by their score in the general reading comprehension skill test PROLEC-SE). So, how could listening comprehension outperform reading comprehension in good decoding readers? As the reader knows, the functional approach explains such a result by stating that the spoken regulation is better than the written regulation since learners take advantage from expressiveness to decide whether or not performing what regulation asks them to do.

From the simple view of reading it is not possible to explain why the written modality was better than the spoken modality in our experiments. The simple view of reading would predict reading comprehension to be equal than listening comprehension if decoding was completely mastered or automatized. But it would never predict reading comprehension to outperform listening comprehension. In the Experiments 1 and 2 reading comprehension was better than listening comprehension for the explanatory aids. How could those results fit into the simple view of reading? The functional approach explains such a result by stating that the written explanation is better than the spoken explanation because the former provide learners with control, which they use to assimilate the elaborations comprised in the explanatory aids.

Overall, the results we found in the present study cannot be explained from the simple view of reading (Gough et al., 1996). A revision of that view can be made on the basis of the evidence gathered here.

4. Limitations and future directions.

Some limitations of the study should be pointed out. First, as stated in the introduction, given the possibilities of CBLEs, aids can be designed in multiple ways. Thus, aids can be pictorial instead of verbal. In this study pictorial aids were not investigated. Further work might be done in order to clarify the effectiveness of this kind of assistance and the way designers should provide it (e.g., DeKoning, Tabbers, Rikers, & Paas, 2007; Mautone & Mayer, 2001, 2007). Second, although the results are expected to be generalizable to other topics, only plate tectonics has been considered. A confirmation was reported in Bustos et al. (2007) who used the topic of geography. In any case, more research would explore if the conclusions from this research hold for different topics.

References.

- Acuña, S. R., García-Rodicio, H., & Sánchez, E. (en preparación). Building mental models from computer-based learning environments: The interaction between aids to monitoring and regulation processes and prior knowledge.
- Acuña, S. R., García-Rodicio, H., & Sánchez, E. (2008). When do instructional explanations work effectively? In A. Maes & S. Ainsworth (Eds.), *Proceedings EARLI Special Interest Group Text and Graphics: Exploiting the Opportunities Learning with Textual, Graphical and Multimodal Representations* (pp. 173-174). Tilburg, The Netherlands.
- Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences, 11*(1), 25-61.
- Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*, 147-179.
- Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology, 89*, 183-193.
- Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15*, 593-613.
- Allen, J. F., & Perrault, C. R. (1980). Analyzing intention in utterances. *Artificial Intelligence, 15*, 143-178.
- Asymetrix. (2001). *Toolbook Instructor II. [Computer Program, PC]*. Bellevue, WA: Asymetrix Learning Systems.
- Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*, 416-427.
- Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*, 774-783.
- Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon.
- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523-535.
- Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research, 30*, 87-111.
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Instructional Science, 56*, 45-72.
- Bach, K. (1994). Meaning, speech acts, and communication. In R. M. Harnish (Ed.), *Basic Topics in the Philosophy of Language*: Prentice-Hall.
- Bach, K. (2003). Speech acts and pragmatics. In M. Devitt & R. Hanley (Eds.), *Blackwell Guide to the Philosophy of Language*.
- Baddeley, A. (1997). *Human Memory: Theory and practice*. Exeter, UK: Psychology Press.
- Baddeley, A. (1998). Working memory. *Life Sciences, 321*, 167-173.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, 36, 189–208.

Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning? *Metacognition and Learning*, 3, 39-58.

Baron-Cohen, S. (1989). Perceptual role taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology*, 7, 113-127.

Baron-Cohen, S., Leslie, A. M., & Fruth, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21, 37-46.

Bates, E., Camaioni, L., & Volterra, V. (1975/1998). The acquisition of performative prior to speech. In A. Kasher (Ed.), *Pragmatics. Critical concepts* (pp. 274-295). xx: Routledge - Francis Taylor Group.

Bates, E., Camaioni, L., & Volterra, V. (1976). Sensoriomotor performatives. In E. Bates (Ed.), *Language and context: the acquisition of pragmatics* (pp. 49-71). New York: Academic Press.

Baum, K. M., & Nowicki, S. J. (1998). Perception of emotion: measuring decoding accuracy of adult prosodic cues varying in intensity. *Journal of Nonverbal Behavior*, 22(2), 89-107.

Bavelas, J. B. (1994). Gestures as part of speech. Methodological implications. *Research on language and social interaction*, 27, 201-221.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, 15, 469-490.

Baylor, A., & Ryu, J. (2003). Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research*, 28, 373-395.

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, 27, 108-119.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*, 11, 17-67.

Bell, L. C., & Perfetti, C. (1994). Reading skill: Some adult comparisons. *Journal of Educational Psychology*, 86, 244-255.

Berthold, K., Eysink, T. H. S., & Renkl, A. (2008). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science*, DOI 10.1007/s11251-11008-19051-z.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in Self-Explanation and Self-Regulation Strategies: Investigating the Effects of Knowledge Acquisition Activities on Problem Solving. *Cognition and Instruction*, 13, 221-252.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, 14, 325-341.

Brakke, K. E., & Savage-Rumbaugh, S. (1995). The development of language skills in bonobo and chimpanzee. Comprehension. *Language and Communication*, 15, 121-148.

Brakke, K. E., & Savage-Rumbaugh, S. (1996). The development of language skills in pan. Production. *Language and Communication*, 16, 361-380.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: a constructive versus interpretive approach. *Cognitive Psychology*, 3, 193-209.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*, 383-398.

Britton, B. K. (1994). Understanding expository text: Building mental structure to induce insights. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 641-674). New York: Academic Press.

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of Cognitive Load in Multimedia Learning with Dual-Task Methodology: Auditory Load and Modality Effects. *Instructional Science*, *32*, 115-132.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of Cognitive Load in Multimedia Learning Using Dual-Task Methodology. *Experimental Psychology*, *49*, 109-119.

Bruner, J. (1975). From communication to language. *Cognition*, *3*, 255-287.

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Buisine, S., & Martin, J. C. (2007). The effects of speech-gesture cooperation in animated agents' behavior in multimedia presentations. *Interacting with Computers*, *19*, 484-493.

Bustos, A. (en preparación). *La competencia retórica y el aprendizaje de la lengua escrita. ¿Se puede hablar de una competencia específica?*

Bustos, A., Sánchez, E., & García, J. R. (2007, November). Las competencias lectoras: ¿Cuándo los alumnos pueden beneficiarse de los recursos retóricos del texto? [Reading skills: When do students profit from rhetorical devices of texts?]. In *Poster presented at Jornada Científica Homenaje a Jesús Alegría*. Salamanca, Spain.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. *Journal of Educational Psychology*, *96*, 31-42.

Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic Boundaries in Adjunct Attachment. *Journal of Memory and Language*, *45*, 58-81.

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, *30*, 348-369.

Cataldo, M. G., & Oakhill, J. (2000). Why Are Poor Comprehenders Inefficient Searchers? An investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of Educational Psychology*, *92*(791-799).

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 35-53). Norwood, NJ: Ablex.

Chafe, W. (1985). Linguistic Differences Produced by Differences between Speaking and Writing. In D. R. Olson, A. Hildyard & N. Torrance (Eds.), *Literacy, Language, and Learning* (pp. 105-123). Cambridge: Cambridge University Press.

Chafe, W., & Danielewicz, J. (1987). Properties of Spoken and Written Language. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending Oral and Written Language* (pp. 83-113). Berkeley: Academic Press.

Chen, S. Y., Fan, J. P., & Macredie, R. D. (2006). Navigation in hypermedia learning systems: experts vs. novices. *Computers in Human Behavior*, *22*, 251-266.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of

- generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161-238). Mahwah, NJ: Lawrence Erlbaum Associates.
- Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, P. (1989). Self-explanation: how students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145-182.
- Cimpian, A., & Markman, E. M. (2007). Preschool children's use of cues to generic meaning. *Cognition*, doi:10.1016/j.cognition.2007.1007.1008.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, H. H. (2001). Conversation. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of Social and Behavioral Sciences*. London: Elsevier.
- Clark, H. H. (2002). Conversation, structure of. In L. N. Basingstoke (Ed.), *Encyclopedia of Cognitive Science*. England: Macmillan.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition*. Washington: APA Books.
- Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62-81.
- Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259-294.
- Coltheart, M. (2005). Modeling reading: The Dual Route approach. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: a handbook* (pp. 6-23). Oxford: Blackwell publishing.
- Commander, N. E., & Stanwyck, D. J. (1997). Illusion of knowing in adult readers: effects of reading skill and passage length. *Contemporary Educational Psychology*, *22*(1), 39-52.
- Conati, C., & VanLehn, K. (2000). Further results from the evaluation of an intelligent computer tutor to coach self-explanation. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Proceedings 5th International Conference ITS'2000* (pp. 304-313). Montreal, Canada: Springer-Verlag.
- Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, *105*, 658-668.
- Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, *25*, 1-53.
- Craig, S. D., Gholson, B., & Driscoll, D. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*, 428-434.
- Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *PROLEC-R, Bateria de Evaluación de los procesos lectores (revisada)*: TEA Ediciones.
- Daneman, M., & Carpenter, P. A. (1980). Individual Differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology*, *21*, 731-746.
- De Vega, M. (2002). Del significado simbólico al significado corpóreo. *Estudios de Psicología*, *23*, 153-174.

- De Vega, M. (2005). Lenguaje, corporeidad y cerebro: Una revisión crítica. *Revista Signos*, 38, 157-176.
- De Vega, M., Carreiras, M., Gutiérrez-Calvo, M., & Alonso-Quecuty, M. (1990). *Lectura y comprensión. Una perspectiva cognitiva*. Madrid: Alianza Editorial.
- Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, 62, 119-142.
- Dehn, D. M., & van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies*, 52, 1-22.
- deLeeuw, N., & Chi, M. T. H. (2002). Self-explanation: Enriching a situation model or repairing a domain model. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 55-78). Mahwah, NJ: Lawrence Erlbaum Associates.
- Díez, E., & Fernández, A. (1997). Bateria multimedia de comprensión (versión abreviada) [Comprehension multimedia inventory (summarized version)]: University of Salamanca.
- Donald, M. (1993). Précis of Origins of the modern mind: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences*, 16, 737-791.
- Dore, J. (1979). Conversational acts and the acquisition of language. In E. Ochs & B. Schieffelin (Eds.), *Developmental pragmatics*. New York: Academic Press.
- Ekman, P. (1997). Should we call it expression or communication? *Innovations in Social Science Research*, 10, 333-344.
- Elosúa, M. R., Gutiérrez, F., García Madruga, J. A., Luque, J. L., & Gárate, M. (1996). Adaptación española del Reading Span Test de Daneman y Carpenter. *Psicothema*, 8, 383-395.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27, 209-221.
- Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, 100, 233-253.
- Ferreira, F., & Anes, M. (1994). Why study spoken language? In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 33-56). San Diego: Academic Press.
- Fox-Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory and Cognition*, 29, 320-326.
- García, J. N. (1995). La intención comunicativa en autistas no verbales: un estudio observacional. *Revista de Psicología General y Aplicada*, 48, 377-391.
- García-Rodicio, H., & Sánchez, E. (submitted for publication). How should instructional designers present aids to the monitoring process? Evidence for the superiority of the auditory over the visual modality. *International Journal of Human-Computer Studies*.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8, 8-11.
- Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*, 16, 104-121.
- Gernsbacher, M. A. (1995). The Structure Building Framework: What it is, what it might also be, and why. In B. K. Britton & A. C. Graesser (Eds.), *Models of text understanding* (pp. 289-311). Hillsdale, NJ: Erlbaum.
- Gernsbacher, M. A. (1997). Coherence cues mapping during comprehension. In J.

Costermans & M. Fayol (Eds.), *Processing interclausal relationships. Studies in the production and comprehension of text* (pp. 3-21). Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Varner, K. R. (1988). *The multi-media comprehension battery. (Tech Rep. No. 88-3)*. Eugene, OR, Institute of Cognitive and Decision Sciences.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16(3), 430-445.

Gibbs, R. W. (1983). Do people always process the literal meaning of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 524-533.

Gibbs, R. W. (1986). What makes some indirect speech acts conventional? *Journal of Memory and Language*, 25, 181-196.

Gilabert, R., Martínez, G., & Vidal-Abarca, E. (2005). Some good texts are always better: text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction*, 15, 45-68.

Giins, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15, 313-331.

Givón, T. (1984). Prolegomena to discourse pragmatics. *Journal of Pragmatics*, 8, 489-516.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1-55.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, 31, 129-151.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26, 69-83.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, 3, 419-429.

Goldman, S. R., & Bisanz, G. L. (2002). Toward a functional analysis of text genres. In J. Otero, E. León & A. C. Graesser (Eds.), *The psychology of science text comprehension*. New Jersey: Lawrence Erlbaum Associates.

Goldman, S. R., & Rakestraw, J. A. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (pp. 311-335). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gómez, J. C., Laá, M. V., & Phillips, W. (1993). Protoimperativos y atención conjunta. ¿Usan los niños autistas la mirada para pedir cosas? In *Actas VII Congreso Nacional AETAPI*. Salamanca.

Gómez, J. C., Sarriá, E., & Tamarit, J. (1993). The comparative study of early communication and theories of mind: ontogeny, phylogeny, and pathology. In S. Baron-Cohen, H. Tager-Flusberg & D. Cohen (Eds.), *Understanding other minds: perspectives from autism*. Oxford: Oxford University Press.

Gómez, J. C., Sarriá, E., Tamarit, J., Brioso, A., & León, E. (1995). *Los inicios de la comunicación: estudio comparado de niños y primates no humanos e implicaciones para el autismo*. Madrid: CIDE.

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: ununnatural act. *Bulletin of the Orton Society*, 30, 179-196.

- Gough, P. B., Hoover, W. A., & Peterson, C. (1996). Some observations on the simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties* (pp. 1-13). Hillsdale, NJ: Erlbaum.
- Graesser, A. C., & Goodman, S. H. (1985). How to construct conceptual graph structures. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 363-383). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H., Ventura, M., Olney, A., et al. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers*, *36*, 180-192.
- Graesser, A. C., Millis, K. K., & Zwaan, R. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*, 163-189.
- Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N. (2003). AutoTutor improves learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo & J. Kay (Eds.), *Proceedings of artificial intelligence in education* (pp. 47-54). Amsterdam: IOS Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review* *101*, 371-395.
- Graesser, A. C., Wiley, J., Goldman, S. R., O'Reilly, T., Jeon, M., & McDaniel, B. (2007). SEEK Web tutor: fostering a critical stance while exploring the causes of volcanic eruption. *Metacognition and Learning*, *2*, 89-105.
- Green, R. (1981). Remembering ideas from text: the effect of presentation. *British Journal of Educational Psychology*, *51*, 83-89.
- Greene, J. A., & Azevedo, R. (2008). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, doi:10.1016/j.cedpsych.2008.1005.1006.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 41-58). New York: Academic Press.
- Grosjean, F., & Hirst, C. (1996). Using prosody to predict the end of sentences in English and French: Normal and brain-damaged subjects. *Language and Cognitive Processes*, *11*, 107-134.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, *12*, 175-204.
- Haberlandt, K., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General* *114*, 357-374.
- Hacker, D. J. (1997). Comprehension monitoring of written discourse across early-to-middle adolescence. *Reading and Writing: An Interdisciplinary Journal*, *9*, 207-140.
- Hacker, D. J. (1998). Self-Regulated Comprehension During Normal Reading. In D. J. Hacker, A. C. Graesser & J. Dunlosky (Eds.), *Metacognition in Educational Theory and Practice* (pp. 165-191). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Halliday, M. A. K. (1973). *Explorations in the functions of language*. London: Arnold.
- Halliday, M. A. K. (1978). *Language as social semiotic: the social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M. A. K. (1987). Spoken and Written Modes of Meaning. In R. Horowitz & J. Samuels (Eds.), *Comprehending Oral and Written language*. Berkeley: Academic Press.
- Hancher, M. (1979). The classification of cooperative illocutionary acts. *Language in*

Society, 8, 1-14.

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48, 101-119.

Harnish, R. M. (2007). Mood: What's the illocutionary point? *Current Trends in Pragmatics*.

Harskamp, E. G., Mayer, R. E., & Suhre, C. (2007). Does the modality principle for multimedia learning apply to science classrooms? *Learning and Instruction*, 465-477.

Haskard, K. B., Williams, S. L., DiMatteo, M. R., Heritage, J., & Rosenthal, R. (2008). The provider's voice: Patient satisfaction and the content-filtered speech of nurses and physicians in primary medical care. *Journal of Nonverbal Behavior*, 32, 1-20.

Hausmann, R., & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology*, 7, 4-15.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717-742.

Hildyard, A., & Olson, D. R. (1982). On the comprehension of oral vs written discourse. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 19-33). Norwood, NJ: Ablex.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 1205-1218.

Holtgraves, T. (2004). The production and perception of implicit performatives. *Journal of Pragmatics*, 37, 2024-2043.

Holtgraves, T. (2007). Automatic intention recognition in conversation processing. *Journal of Memory and Language*, doi:10.1016/j.jml.2007.1006.1001.

Holtgraves, T., & Ashley, A. (2001). Comprehending illocutionary force. *Memory and Cognition*, 29, 83-90.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2, 127-160.

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual Differences in Reading to Summarize Expository Text: Evidence From Eye Fixation Patterns. *Journal of Educational Psychology*, 94, 44-55.

Jeung, H. J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17, 329-343.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11, 47-78.

Juszyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.

Just, M. A., Carpenter, P. A., & Wolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228-238.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, 92, 1-11.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective in on*

cognitive science. Cambridge: MIT Press.

Kendeou, P., Lynch, J. S., van den Broek, P., Espin, C. A., White, M. J., & Kremer, K. E. (2005). Developing Successful Readers: Building Early Comprehension Skills through Television Viewing and Listening. *Early Childhood Education Journal*, 33, 91-98.

Kendeou, P., & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology*, 97, 235-245.

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Souther Italian conversation. *Journal of Pragmatics*, 23, 247-279.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W. (1994b). The psychology of discourse processing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. New York: Academic Press.

Kintsch, W. (1994). Text comprehension, memory and learning. *American Psychologist*, 49(4), 294-303.

Kintsch, W., & Keenan, J. (1973). Reading rate and retention as a function of number of propositions in base structure of sentences. *Cognitive Psychology*, 5(3), 257-274.

Kintsch, W., & Kintsch, E. (1995). Strategies to promote active learning from text: Individual differences in background knowledge. *Swiss Journal of Psychology*, 54, 141-151.

Kintsch, W., & Kozminsky, E. (1977). Summarizing Stories after Reading and Listening. *Journal of Educational Psychology*, 69(5), 491-499.

Kintsch, W., Kozminsky, E., Streby, W., McKoon, G., & Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 14, 196-214.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: a theoretical analysis. *Journal of Memory and Language*, 29, 133-159.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science*, 7, 54-59.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, 61, 743-754.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, 4, 343-346.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., et al. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, 35, 1437-1444.

Lang, P. J. (1994). The varieties of emotional experience: a mediation on James-Lange theory. *Psychological Review*, 101, 211-221.

Last, D. A., O'Donnell, A. M., & Kelly, A. E. (2001). The Effects of Prior Knowledge and Goal Strength on the Use of Hypertext. *Journal of Educational Multimedia and Hypermedia*, 10.

Lazonder, A. W., & Rouet, J. F. (2008). Information problem solving instruction: Some cognitive and metacognitive issues. *Computers in Human Behavior*, 24, 753-765.

Leahy, W., Chandler, P., & Sweller, J. (2003). When Auditory Presentations Should and Should not be a Component of Multimedia Instruction. *Applied Cognitive Psychology*, 17, 401-418.

León, J. A., & Pérez, O. (2002). Taxonomías y tipos de inferencias. In J. A. León (Ed.),

- Conocimiento y discurso. Claves para inferir y comprender* (pp. 45-68). Madrid: Pirámide.
- Levin, J., & Divine-Hawkins, P. (1974). Visual imagery as a prose-learning process. *Journal of Reading behavior*, 6, 23-30.
- Lieberman, A. M. (1992). The relation of speech to reading and writing. In R. Frost & L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning*. North Holland Publishers: Elsevier.
- Lin, L. M., & Zabucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology*, 23, 345-391.
- Linderholm, T., Everson, M. G., Van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of Causal Text Revisions on More- and Less-Skilled Readers' Comprehension of Easy and Difficult Texts. *Cognition and Instruction*, 18, 525-556.
- Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14, 257-274.
- Maki, R. H., & Berry, S. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 10, 663-679.
- Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 609-616.
- Mañá, A., Vidal-Abarca, E., Domínguez, C., Gil, L., & Cerdán, R. (2008). Papel de los procesos metacognitivos en una tarea de pregunta-respuesta con textos escritos.
- Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, 46, 986-992.
- Massaro, D. W. (1994). Psychological aspects of speech perception. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 219-263). New York: Academic Press.
- Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, 93, 377-389.
- Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, 99(3), 640-652.
- Mayer, R. E. (1983). Can you repeat that? Qualitative effects of repetition and advance organizers on learning from science prose. *Journal of Educational Psychology*, 75, 40-49.
- Mayer, R. E. (2001). *Multimedia Learning*. New York: Cambridge University Press.
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation. Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an active self-explaining environment: what works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95, 806-813.
- Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions, pretraining, and signaling. *Journal of Educational Psychology*, 76, 1089-1105.
- Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715-726.
- Mayer, R. E., & Moreno, R. (1998). A split attention effect in multimedia learning. Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90, 312-320.

- McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human-Computer Interaction, 5*, 381-413.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*, 440-466.
- McNamara, D. S. (2004). SERT: Self-Explanation Reading Training. *Discourse Processes, 38*, 1-30.
- McNamara, D. S., Kintsch, E., Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.
- McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247-288.
- McNamara, D. S., O'Reilly, T., Best, R. M., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research, 34*, 147-171.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92*, 350-371.
- Mitrovic, A., & Suraweera, P. (2000). Evaluating an animated pedagogical agent. *Lecture Notes in Computer Science, 1839*, 73-82.
- Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45-64.
- Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*, 99-113.
- Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology, 31*, 186-207.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358-368.
- Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94*(3), 598-610.
- Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology, 97*(1), 117-128.
- Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*(2), 177-213.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 498-550.
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319-334.
- Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: the contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry, 27*, 657-669.
- Nation, K., & Snowling, M. J. (1998). Semantic Processing and the Development of Word-Recognition Skills: Evidence from Children with Reading Comprehension Difficulties.

Journal of Memory and Language, 39, 85-101.

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., et al. (1998). Frontotemporal lobar degeneration. *Neurology*, 51, 1546-1554.

Nückles, M., Wittwer, J., & Renkl, A. (2005). Information About a Layperson's Knowledge Supports Experts in Giving Effective and Efficient Online Advice to Laypersons. *Journal of Experimental Psychology: Applied*, 11, 219-236.

O'Brien, J. E., & Albrecht, E. J. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 777-784.

Olson, D. R. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review*, 47, 257-281.

Otero, J., Campanario, J. M., & Hopkins, K. D. (1992). The relationship between academic achievement and metacognitive comprehension monitoring ability of Spanish secondary school students. *Educational and Psychological measurement*, 52, 419-430.

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, 3, 229-235.

Otero, J. C. (2002). Noticing and fixing difficulties in understanding science texts. In J. Otero, J. A. León & A. C. Graesser (Eds.), *The Psychology of Science Text Comprehension* (pp. 281-307). Mahwah, N.J.: Lawrence Erlbaum Associates.

Otero, J. C., & Campanario, J. M. (1990). Comprehension evaluation and regulation in learning from science texts. *Journal of Research in Science Teaching*, 27, 447-460.

Paul, R., & Cohen, D. J. (1985). Comprehension of Indirect Requests in Adults with Autistic Disorders and Mental Retardation. *Journal of Speech and Hearing Research*, 28, 475-479.

Penney, C. G. (1989). Modality effects and the structure of short-term working memory. *Memory and Cognition*, 17, 398-422.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357-383.

Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language*, 24, 503-518.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 000-000.

Pierrehumbert, J., & Hirschberg, J. (1990). The Meaning of Intonation in the Interpretation of Discourse. In P. Cohen, J. Morgan & M. Pollack (Eds.), *Intentions in Communication* (pp. 271-311). Cambridge MA: MIT Press.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707-784.

Pinker, S., & Jakendoff, R. (2005). The Faculty of Language: What's Special about it? *Cognition*, 95, 201-236.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction*, 12, 235-275.

Plaut, D. C. (2005). Connectionist approaches to reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: a handbook* (pp. 24-38). Oxford: Blackwell Publishing.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12, 61-86.

Postle, B., Idzikowski, Ch., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The

selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 100-120.

Ramos, F. J., & Cuetos, F. (1999). *Prolec-SE. Evaluación de los Procesos Lectores en alumnos de 3er ciclo de Educación Primaria y de Educación Secundaria [Prolec-SE. Reading Processes of Middle and High-School Students Assessment]*. Madrid: TEA.

Ratcliff, G., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior*, *17*, 403-417.

Ravid, D., & Berman, R. A. (2004). Information Density in the Development of Spoken and Written Narratives in English and Hebrew. *Discourse Processes*, *41*, 117-149.

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology*, *97*(1), 70-80.

Rayner, K., Juhasz, B. J., & Pollatsek, A. (2005). Eye movements during reading. In C. Hulme & M. Snowling (Eds.), *Handbook of Reading Research* (pp. 79-97). Oxford: Blackwell.

Reeve, J., Jang, H., Hardre, P., & Omura, M. (2002). Providing a Rationale in an Autonomy-Supportive Way as a Strategy to Motivate Others During an Uninteresting Activity. *Motivation & Emotion*, *26*, 183-207.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*, 1-29.

Renkl, A. (2002). Learning from worked-out examples: Instructional explanations supplement self-explanations. *Learning and Instruction*, *12*, 529-556.

Rouet, J. F. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers*, *15*, 408-428.

Rubman, C. N., & Waters, H. S. (2000). A, B Seeing: The Role of Constructive Processes in Children's Comprehension Monitoring. *Journal of Educational Psychology*, *92*, 503-514.

Sánchez, E., García, J. R., & Bustos, A. (en prensa). ¿Es la necesidad de decodificar las palabras escritas la única diferencia entre la comprensión del lenguaje oral y la comprensión de textos? In

Sánchez, E., & García-Rodicio, H. (2008). The use of modality in the design of verbal aids in computer-based learning environments. *Interacting with Computers*, *20*, 545-561.

Sánchez, E., García-Rodicio, H., & Acuña, S. R. (2008). Are instructional explanations more effective in the context of an impasse? *Instructional Science*, DOI 10.1007/s11251-11008-19074-11255.

Scardamalia, M., & Bereiter, C. (1987). Two models of the processes in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics*. Cambridge: Cambridge University Press.

Schafer, A., Carter, J., Clifton, C., & Frazier, L. (1996). Focus in relative clause construal. *Language and Cognitive Processes*, *11*, 135-163.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2008). The effects of realism in learning with dynamic visualizations. *Learning and Instruction*, *xxx*, 1-14.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: verbatim, propositional, and situational representations. *Journal of Memory and Language*, *25*, 279-294.

- Schnotz, W. (2002). Towards an Integrated View of Learning From Text and Visual Displays. *Educational Psychology Review, 14*, 101-120.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning* (pp. 49-69). Cambridge: Cambridge University Press.
- Schommer, M., & Surber, J. R. (1986). Comprehension monitoring failure in skilled adult readers. *Journal of Educational Psychology, 78*, 353-357.
- Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education, 46*, 426-445.
- Searle, J. R. (1969). *Speech acts*. Cambridge: Cambridge University Press.
- Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics 3: Speech acts* (pp. 59-82). New York: Academic Press.
- Searle, J. R. (1975). A taxonomy of illocutionary acts. In K. Gunderson (Ed.), *Language, Mind, and Knowledge* (pp. 1-29). Minneapolis.
- Searle, J. R. (1979). *Expression and meaning*. Cambridge: Cambridge University Press.
- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*, 227-237.
- Seufert, T., Schütze, M., & Brünken, R. (2008). Memory characteristics and modality in multimedia learning: An aptitude treatment interaction study. *Learning and Instruction, xx*, 1-15.
- Shapiro, A. (1998). Promoting active learning: the role of system structure in learning from hypertext. *Human-Computer Interaction, 13*, 1-35.
- Skarakis-Doyle, E. (2002). Young children's detection of violations in familiar stories and emerging comprehension monitoring. *Discourse Processes, 33*, 175-197.
- Sleeman, D., Kelly, A. E., Martinak, R., Ward, R. D., & Moore, J. L. (1989). Studies of diagnosis and remediation with high school algebra students. *Cognitive Science, 551-568*.
- Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. (1977). Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Journal of Educational Psychology, 69*, 381-187.
- Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language, 32*, 25-38.
- Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language, 17*, 3-23.
- Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*, 51-79.
- Stiles, W. B. (1981). Classification of intersubjective illocutionary acts. *Language in Society, 10*, 227-249.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.
- Tabbers, H. K. (2002). *The modality of text in multimedia instructions: Refining the design guidelines*.: Unpublished doctoral dissertation, Open University of the Netherlands, Heerlen.
- Tabbers, H. K. (2006). Where did the modality effect go? In *Proceedings of the SIG 2 Meeting on Text and Graphics Comprehension* (pp. 27-30).

- Tabbers, H. K., Martens, R. L., & Van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology, 74*, 71-81.
- Tan, A., & Nicholson, T. (1997). Flashcards revisited: Training poor readers to read words faster improves their comprehension of text. *Journal of Educational Psychology, 89*, 276-288.
- Tannen, D. (1982). The oral/literate continuum in discourse. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 1-16). Norwood, NJ: Ablex.
- Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17*, 273-281.
- Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: an analysis of selection items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*, 1024-1037.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When Two Sensory Modes Are Better Than One. *Journal of Experimental Psychology: Applied, 3*, 257-287.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, L. M. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition, 73*, 89-134.
- Tversky, B., Morrison, J. B., & Bretancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies, 57*, 247-262.
- Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers gestures facilitate students learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*, 187-204.
- Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating Memory-Based and Constructionist Processes in Accounts of Reading Comprehension. *Discourse Processes, 39*, 299-316.
- van Dijk, T. (1977). Context and cognition: Knowledge frames and speech act comprehension. *Journal of Pragmatics, 1*, 211-233.
- van Dijk, T., & Kintsch, W. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 213-345.
- Van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- van Mulken, S., André, E., & Müller, J. (1998, January). *The persona effect: How substantial is it?* Proceedings of HCI on People and Computers XIII.
- van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 371-386.
- VanLehn, K., Graesser, A. C., Jackson, G., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31*, 3-62.
- Vidal-Abarca, E., Martínez, G., & Gilabert, R. (2000). Two procedures to improve instructional text: Effects on memory and learning. *Journal of Educational Psychology, 92*(1), 1-10.
- Voss, J. F., & Silfies, L. N. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction, 14*, 45-68.
- Werner, E. (1988b). A formal computational semantics and pragmatics of speech acts. In x. xx (Ed.), *Proceedings of the 12th conference on Computational linguistics* (pp. 744 -

749). Budapest: Association for Computational Linguistics.

Werner, E. (1988a). Toward a theory of communication and cooperation for multiagent planning. In *Theoretical aspects of reasoning about knowledge. Proceedings of the 2nd Conference*: Morgan Kaufman Publishers.

Wild, T. C., Enzle, M. E., Nix, G., & Deci, E. L. (1997). Perceiving others as intrinsically or extrinsically motivated: effects on expectancy formation and task engagement. *Personality and Social Psychology Bulletin*, *23*, 837-848.

Wilson, D. (1994). Relevance and Understanding. In G. Brown, K. Malmkjaer, A. Pollitt & J. Williams (Eds.), *Language and understanding* (pp. 35-58). Oxford: Oxford University Press.

Wilson, D., & Sperber, D. (2002). Relevance theory: A tutorial. In Y. Otsu (Ed.), *Proceedings of the Third Tokyo Conference on Psycholinguistics* (pp. 45-70). Tokyo: Hituzi Shobo.

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, *38*, 1559-1579.

Winne, P. H. (1995). Self-regulation is ubiquitous but its forms vary with knowledge. *Educational Psychologist*, *30*, 223-228.

Wouters, P., Paas, F., & Van Merriënboer, J. J. G. (2008). Observational learning from animated models: Effects of modality and reflection on transfer. *Contemporary Educational Psychology*, doi:10.1016/j.cedpsych.2008.1003.1001.

Yuill, N. M., Oakhill, J. V., & Parkin, A. J. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology*, *80*, 351-361.

Zwaan, R. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 35-62). San Diego: Elsevier.

Zwaan, R., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162-185.

Appendix (A).

Complexity analyses of the aids in the present experiments and in prior research (following the system of Graesser and Goodman (1985) for constructing *conceptual structures*).

I. Example of the explanatory aids in Moreno and Mayer (2002); Moreno et al. (2001).

“The stem should be long enough to put the leaves in the sun.” (Moreno et al., 2001: 184).

Analysis:

Statement node 1: The stem should be long enough
 Statement node 2: Long stems put the leaves in the sun
 Consequence arc: Statement node 1 [causes] Statement node 2

The explanatory aid involves two statement nodes and one casual link.

II. Example of the explanatory aids in Atkinson (2002).

“We need to set up another proportional relationship to determine the production time.” (Atkinson, 2002: 420).

Analysis:

Statement node 1: We need to set up another proportional relationship
 Statement node 2: The setting up determines the production time
 Consequence arc: Statement node 1 [causes] Statement node 2

The explanatory aid involves two statement nodes and one casual link.

III. Analyses of some of our aids (Exp. 1 & 2).

III. 1. The regulatory aid (87 words in Spanish):

“Usually the people who watch this presentation tend to elaborate a simplified conception of the plate collisions process; thus, probably you only saw that plates collide so that mountains are formed both in the Andes and in the Himalaya by the same principle. However, there are important differences between both plate collisions that play a big role in clarifying what plate tectonics is. You must consider the following ideas.”

Analysis:

Statement node 1: Usually people (those who watch this presentation) tend to

elaborate a simplified conception of the plate collisions process
 Statement node 2: (Probably) You saw that plates collide so that mountains are formed
 Statement node 3: There are important differences between plate collisions
 Statement node 4: Differences play a big role in clarifying what is plate tectonics
 Statement node 5: You must consider the following ideas
 Consequence arc: Statement node 1 [causes] Statement node 2

The explanatory aid involves five statement nodes and one casual link.

III. 2. The explanatory aid (141 words in Spanish):

“There are two main differences. First, in the Andes one plate sinks because of its weight. In this collision, there is one continental plate and one oceanic plate, each with its own composition. The denser plate, namely, the oceanic plate, sinks and returns to the mantle. In the Himalaya, both plates are continental and so have the same composition. When the plates collide they push each other and neither sinks. Second, the effects of these two types of collisions are quite different. In the Andes the sinking plate pushes the plate above, putting pressure upon it and forming cracks. Magma emerges through the cracks and thus volcanoes are formed. In contrast, in the Himalaya magma does not emerge through the cracks; instead, both plates push each other without producing cracks so that a mountain without volcanoes is formed.”

Analysis:

Statement node 1: There are two main differences (between the Andes and the Himalaya)
 Statement node 2: In the Andes one plate sinks
 Statement node 3: The sinking plate has certain weight
 Statement node 4: There is one continental and one oceanic plate
 Statement node 5: Each one has its own composition
 Statement node 6: The oceanic plate sinks and returns to the mantle
 Statement node 9: The oceanic plate is the denser plate
 Statement node 10: In the Himalaya there are two continental plates
 Statement node 11: Continental plates have the same composition
 Statement node 12: When plates collide they push each other and neither sinks
 Statement node 13: Effects are quite different
 Statement node 14: In the Andes the sinking plate pushes the plate above
 Statement node 15: The pushing put pressure and forms cracks
 Statement node 16: Magma emerge through the cracks
 Statement node 17: Volcanoes are formed
 Statement node 18: In the Himalaya magma does not emerge through the cracks
 Statement node 19: Both plates push each other without producing cracks
 Statement node 20: A mountain without volcanoes is formed
 Consequence arc: Statement node 4 [causes] Statement node 3
 Consequence arc: Statement node 10 [causes] Statement node 11
 Consequence arc: Statement node 13 [causes] Statement node 14

Consequence arc: Statement node 14 [causes] Statement node 15
Consequence arc: Statement node 15 [causes] Statement node 16
Consequence arc: Statement node 18 [causes] Statement node 17
Consequence arc: Statement node 17 [causes] Statement node 19

The explanatory aid involves 19 statement nodes and seven casual links.

Appendix (B).

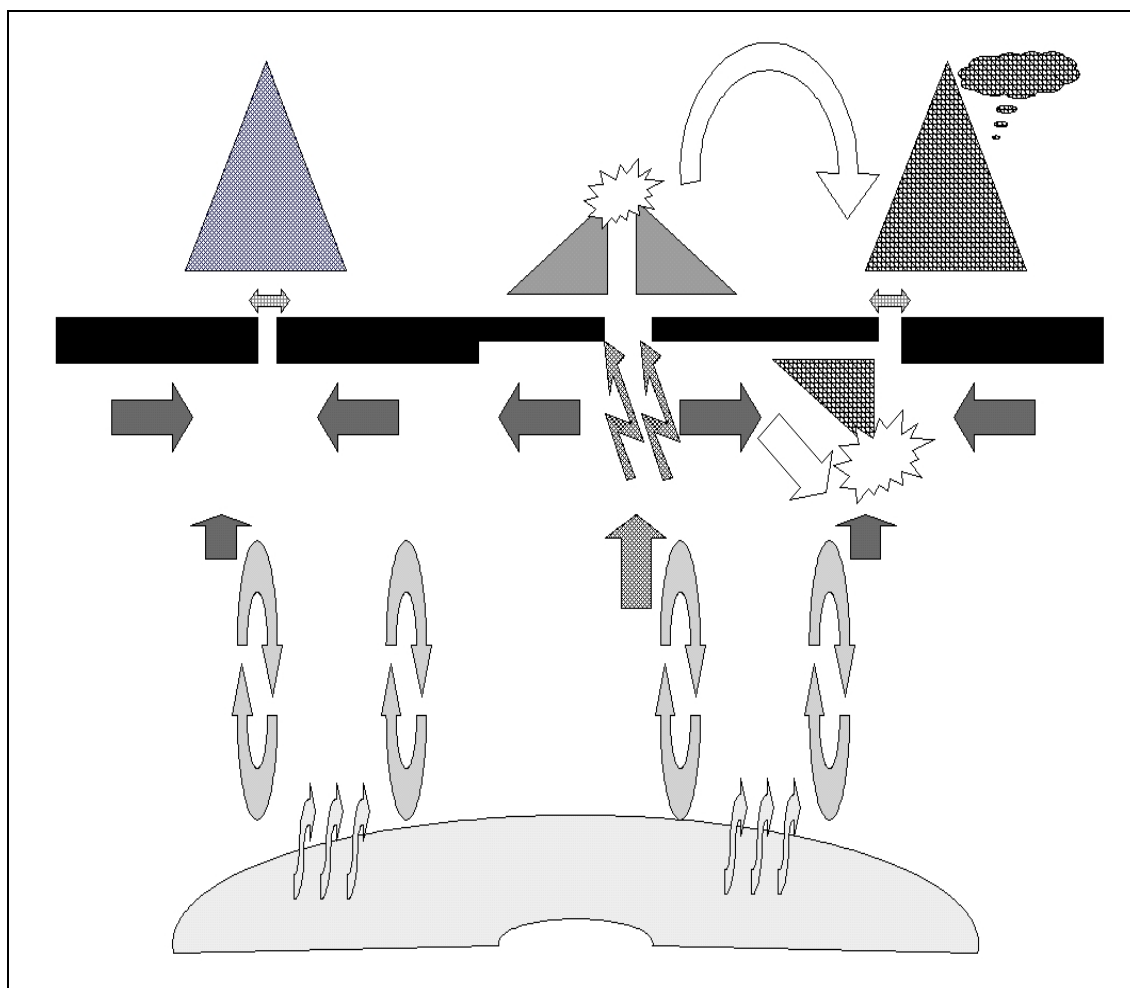


Diagram representing the topic covered in the CBLE
(From the bottom to the top of page)

The Earth's core transmitting heat. The magma currents in the mantle moving up and down because of the heat from the core. The plates floating on magma and moving back and forth because of the currents. The magma surfacing through the ridges. The magma in the surface solidifying and creating new crust (center). A mountain being formed because of the plates collision (left). A mountain with volcanoes being created because of the plates collision (right). A plate, involved in that collision, sinking because of its weight in the trench. The crust of that plate being destroyed. The recycling loop between the new crust being created in the ridges and the crust destroyed in the trenches.