



VNIVERSIDAD D SALAMANCA

INSTITUTO DE NEUROCIENCIAS DE CASTILLA Y LEÓN

TESIS DOCTORAL

PROCESOS COGNITIVOS Y AFECTIVOS EN ADULTOS MAYORES MEDIDOS POR MEDIO DEL HABLA

Autor:

Israel Martínez Nicolás

Director:

Dr. Juan José García Meilán

2022

El Dr. Juan José García Meilán, Profesor Titular del Departamento de Psicología Básica, Psicobiología y Metodología de las Ciencias del Comportamiento de la Universidad de Salamanca (España)

CERTIFICA:

Que la presente Tesis Doctoral titulada “PROCESOS COGNITIVOS Y AFECTIVOS EN ADULTOS MAYORES MEDIDOS POR MEDIO DEL HABLA”, presentada por Israel Martínez Nicolás, ha sido realizada bajo su dirección y reúne los requisitos académicos y científicos, de calidad y originalidad, para que se proceda a su defensa pública ante el tribunal designado.

Y para que así conste y obre los efectos oportunos, firmo el presente certificado en Salamanca, a 23 de junio de 2022.

Fdo: Juan José García Meilán

Agradecimientos

Probablemente nadie se imagine que lo más complicado de escribir de una tesis sean los agradecimientos. Y sin embargo es la parte que más he evitado. Demasiados problemas. Me olvidaré de alguien. Qué orden seguir. Será buena idea mencionar por un lado a aquellos a los que debo un agradecimiento académico y aquellos a los que es personal, pero ¿es que son separables? Por ello voy a empezar disculpándome, y recordando, precisamente, a aquellos que seguramente merezcan estar aquí, pero soy demasiado despistado para nombrar.

Gracias a mi director, el profesor Juan José García Meilán, que me ha guiado en todas mis etapas académicas en la universidad, a quien debo todo lo que sé, y que ha ido mucho más allá de lo esperado en su papel. Gracias al profesor Juan Carro Ramos, que, aunque no haya podido finalizar su codirección, me brindó la oportunidad que necesitaba para poder estar hoy escribiendo esto. Gracias a mis otros compañeros de investigación, especialmente a Francisco y a Olga, por su paciencia y su amabilidad. Porque, aunque esto sea mi tesis, la ciencia es colaboración.

Gracias a mis compañeras de doctorado porque nuestra ayuda mutua ha sido un salvavidas en numerosas ocasiones. Gracias también a mis compañeros de departamento y de facultad que han hecho que me sienta acogido. Gracias al resto de mis compañeros de trabajo, a veces invisibles: el personal de limpieza, de conserjería, de cafetería, de secretaría, etc. Porque la Universidad también es suya.

Gracias a todos los que me han acompañado con su amistad en este proceso. A mis compañeros con los que me formé como psicólogo Dani, Jesús y Laura. A Saúl y Guillermo, por tantos ratos de distracción. Al grupo de la D. S. por la euforia compartida: Inés, Cuau, Bea, José y Mónica. A Elena, Elsa y Alba que han sido una vía de escape. De

entre ellos quiero destacar a dos personas que han contribuido de manera especial. A Jorge, porque hemos crecido juntos y en todo este tiempo nunca he dejado de aprender con él y de él. Y a Thide, que podría estar en cada uno de los párrafos, a quien acudo siempre, con quien he pasado lo bueno y lo malo, con quien espero seguir haciéndolo.

Gracias a Julia porque me hace mejor, la recta final habría sido mucho más dura sin ella. Y gracias finalmente a mi familia, especialmente a mis padres, Carlos y María Jesús, que de entre todos los mencionados son y serán los únicos incondicionales.

La realización de esta Tesis Doctoral ha sido posible gracias a un contrato predoctoral de la Universidad de Salamanca del Programa III: Ayudas para financiar contratos predoctorales de la Universidad de Salamanca, cofinanciadas por el Banco Santander, Convocatoria 2019.

**Parte de los resultados de esta tesis doctoral ya han sido publicados en revistas con
revisión por pares:**

Martínez-Nicolás, I., Llorente, T. E., Martínez-Sánchez, F., & Meilán, J. J. G. (2021).

Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: a systematic review article. *Frontiers in Psychology*, 12, 620251. <https://doi.org/10.3389/fpsyg.2021.620251>

Martínez-Nicolás, I., Llorente, T. E., Ivanova, O., Martínez-Sánchez, F., Meilán, J. J. G.

(2022) Many Changes in Speech through Aging Are Actually a Consequence of Cognitive Changes. *International Journal of Environmental Research and Public Health*. 19(4):2137. <https://doi.org/10.3390/ijerph19042137>

Meilán, J. J., Martínez-Sánchez, F., Martínez-Nicolás, I., Llorente, T. E., & Carro, J.

(2020). Changes in the rhythm of speech difference between people with nondegenerative mild cognitive impairment and with preclinical dementia. *Behavioural Neurology*, 2020. <https://doi.org/10.1155/2020/4683573>

Resumen

El análisis automático del habla es una técnica que permite extraer información lingüística objetivamente de la señal de sonido emitida al hablar. En la producción de habla se interrelacionan numerosos sistemas para seleccionar y planificar el mensaje, dotarlo de una estructura adecuada, y enviar las señales neuromusculares a los órganos implicados en la producción del sonido. Estos procesos determinan las características de la onda sonora que se emite, de modo que el análisis del habla ha sido utilizado para detectar diversas alteraciones que afectan a estos sistemas lingüísticos como son la demencia tipo Alzheimer y el deterioro cognitivo leve.

Esta tesis indaga sobre cómo se alteran los parámetros del habla en adultos mayores afectados por diferentes deterioros cognitivos y/o afectivos. Podemos distinguir en este trabajo dos fases: En la primera se llevan a cabo dos estudios con el objetivo de hallar factores cognitivos previos a los cambios en el habla durante el proceso de envejecimiento. En la segunda, se realizan cuatro estudios que tratan de obtener combinaciones de parámetros del habla susceptibles de ser utilizados como algoritmos en la detección de varios trastornos mediante la manipulación del proceso utilizado para elicitación de habla.

En el primer estudio se concluye que diversos parámetros del habla relacionados con la duración, el ritmo, las frecuencias y el análisis espectral, sufren cambios que se relacionan con el estado cognitivo general y que de hecho podrían ser sensibles a varias etapas de un deterioro. En el segundo, se examina si esos parámetros se explican mediante procesos cognitivos específicos, encontrando una relación con el acceso lingüístico a la memoria semántica, al léxico y la función ejecutiva.

A continuación, los resultados demuestran que dentro de las personas con deterioro cognitivo leve podría haber perfiles de habla correspondientes a aquellos cuya causa subyacente es la Enfermedad de Alzheimer, y que podrían ser identificados a través de medidas del ritmo del habla. Se han utilizado dos tareas, una de lectura y otra de fluidez verbal, que permitan por medio del análisis del habla detectar con un nivel de éxito aceptable a personas con deterioro cognitivo leve y/o Alzheimer. Finalmente, se trató de extender el método del análisis del habla a la detección de depresión en mayores como un primer paso hacia el diagnóstico diferencial de depresión y demencia.

Abstract

Automatic speech analysis is a technique that allows linguistic information to be extracted objectively from the sound signal emitted during the act of speaking. Numerous systems are interrelated in speech production to select and plan the message, provide it with an appropriate structure, and send the neuromuscular signals to the organs involved in sound production. These processes determine the characteristics of the sound wave that is emitted, thus, speech analysis has been used to detect various disorders involving these systems, such as Alzheimer's dementia and mild cognitive impairment.

This thesis investigates the alteration of speech parameters in older adults affected by cognitive and/or affective impairments. We can distinguish two phases in this work. In the first one, two studies are carried out with the aim of finding cognitive factors of speech changes during the aging process. In the second, four studies are carried out in an attempt to obtain combinations of speech parameters that can be used as algorithms in the detection of various disorders by manipulating the process used to elicit speech.

The first study concludes that various speech parameters related to duration, rhythm, frequencies and spectral analysis undergo changes that are related to general cognitive state and may in fact be sensitive to various stages of impairment. In the second, we examine whether these parameters are explained by specific cognitive processes, finding a relationship with linguistic access to semantic memory, lexicon and executive function.

Next, the results show that within people with mild cognitive impairment there could be speech profiles corresponding to those whose underlying cause is Alzheimer's disease and that could be identified through measures of rhythm. Two tasks, reading and verbal fluency, are proposed. Speech analysis on these tasks can be used with an

acceptable level of success to detect people with mild cognitive impairment and/or Alzheimer's disease. Finally, an attempt was made to extend the method to the detection of depression in older adults as a first step towards the differential diagnosis of depression and dementia.

Tabla de contenido

1. Introducción.....	9
2. Antecedentes del estudio	17
2.1 Producción y análisis del habla.....	22
2.2 El Análisis Automático del Habla como herramienta de cribado en salud mental.....	42
2.2.1 Análisis del habla en las enfermedades del sistema nervioso.....	44
2.2.2 Análisis del habla en los trastornos mentales, del comportamiento y del neurodesarrollo.....	48
2.2.3 Críticas y perspectiva de futuro	53
2.3 El lenguaje en la Demencia tipo Alzheimer	56
2.4 El habla en el deterioro cognitivo leve y la enfermedad de Alzheimer.....	63
3. Objetivos experimentales	101
4. Método.....	105
5. Resultados.....	115
Estudio I. Algunos cambios en el habla de los mayores se deben en realidad a cambios en su estado cognitivo	117
Estudio II. Procesos cognitivos medidos por los parámetros del habla de las personas con demencia.....	139
Estudio III. Los cambios en el ritmo del habla diferencian personas con deterioro cognitivo leve no degenerativo y personas con demencia preclínica.....	153
Estudio IV. Cribado de Deterioro Cognitivo Leve mediante el análisis de la voz en una tarea de fluidez semántica	165
Estudio V. Una herramienta de cribado del deterioro cognitivo leve y la enfermedad de Alzheimer a través del análisis del habla	175
Estudio VI. Detectando síntomas depresivos en mayores a través del habla	183
6. Discusión	195
7. Conclusiones.....	211
8. Referencias	217

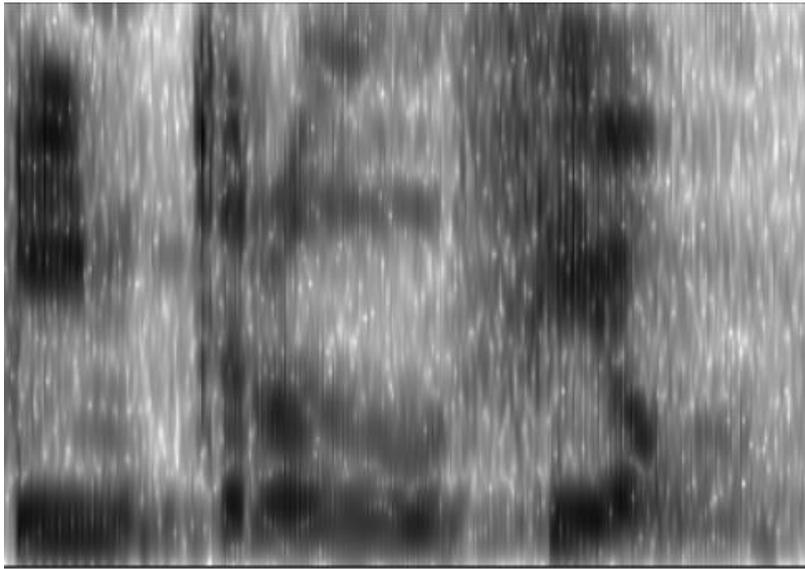
Índice de Figuras

Figura 1. “Ondas sonoras periódicas simples con la misma amplitud y con diferente frecuencia” por Joaquim Llisterri licencia bajo BY CC 4.0.....	23
Figura 2. “Ondas sonoras periódicas simples con la misma frecuencia (100 Hz) y con diferente amplitud” por Joaquim Llisterri licencia bajo BY CC 4.0.....	24
Figura 3. Representación del aparato fonador y su relación con parámetros del habla.	26
Figura 4. Espectrograma de las cinco vocales del español en su variante europea.	27
Figura 5. Representación de la distribución de la intensidad y las frecuencias de una onda simple y una compleja.	29
Figura 6. Espectrograma con bandas de 500 Hz.	30
Figura 7. Espectrograma de un hablante leyendo el principio del Quijote.....	32
Figura 8. Parámetros y su clasificación.	33
Figura 9. Proceso de validación de un parámetro del habla como posible biomarcador.	39
Figura 10. Diagrama de flujo PRISMA del proceso seguido para seleccionar los artículos incluidos en la revisión.	69
Figura 11. Evaluación de la calidad de los estudios descriptivos mediante la lista de comprobación JBI, y su calificación es como riesgo de sesgo alto, bajo o incierto para cada ítem.....	88
Figura 12. Proporción de estudios descriptivos con evaluación de riesgo de sesgo bajo, alto o incierto.	88
Figura 13. Evaluación de la calidad de los estudios de cribado mediante la lista de comprobación QUADAS-2 y su clasificación como riesgo de sesgo alto, bajo o incierto.	91
Figura 14. Proporción de los resultados de la evaluación de riesgos de sesgo según el Quadas-2.....	91
Figura 15. Número de artículos publicados descriptivos y de cribado obtenidos en la revisión.	92
Figura 16. Tareas utilizadas para elicitación de lenguaje en los estudios descriptivos y de cribado. Si un estudio utilizó varias tareas, se han incluido todas.	97
Figura 17. Diagrama de los objetivos de la investigación.....	104

Índice de tablas

Tabla 1. Estudios descriptivos del habla de las personas con DCL o EA.	71
Tabla 2. Estudios de cribado de DCL y EA.	77
Tabla 3. Resumen de los parámetros alterados en las personas con DCL y EA.	94
Tabla 4. Características de la muestra.	122
Tabla 5. Influencia del género, el estado cognitivo (MMSE) y las categorías de edad en los parámetros del habla implicados en la predicción de la EA.	124
Tabla 6. Comparaciones por pares de parámetros con un efecto principal de la edad o el estado cognitivo. * es significativo a $p < 0,05$ y ** es significativo a $p < 0,001$	125
Tabla 7. Características de la muestra.	143
Tabla 8. Resultados de las regresiones para cada parámetro.	145
Tabla 9. Comparación de los resultados en los estudios I y II.	148
Tabla 10. Características del algoritmo del DAV-EA.	155
Tabla 11. Características de la muestra.	157
Tabla 12. Comparación de los perfiles cognitivos de los grupos nodDCL y preEA.	158
Tabla 13. Comparación de los parámetros del habla de los grupos nodDCL y preEA.	161
Tabla 14. Datos descriptivos de la muestra.	167
Tabla 15. Datos descriptivos de los parámetros discriminantes.	170
Tabla 16. Características de la función discriminante.	171
Tabla 17. Datos descriptivos de la muestra.	176
Tabla 18. Centroides de grupo de las funciones.	178
Tabla 19. Datos descriptivos de los parámetros discriminantes.	179
Tabla 20. Parámetros del habla que forman las funciones discriminantes. Los parámetros indicados con la letra L corresponden a la tarea de lectura, y los indicados con la P, a la de fluidez.	180
Tabla 21. Resultados de clasificación.	181
Tabla 22. Datos descriptivos de la muestra.	186
Tabla 23. Datos descriptivos de los parámetros discriminantes.	188
Tabla 24. Características de la función discriminante.	189
Tabla 25. Resultados de clasificación.	190
Tabla 26. Datos descriptivos de la muestra con los grupos formados según la clasificación.	191

1. Introducción



1. Introducción

Las sociedades occidentales se caracterizan demográficamente por el progresivo envejecimiento de la población, lo que consecuentemente trae consigo el aumento de diversas patologías asociadas a la edad avanzada. Patologías cuya mayor expresión cognitiva es la demencia. La demencia es un trastorno neurocognitivo caracterizado por la aparición en la persona de un deterioro en dos o más dominios cognitivos en relación al estado normalizado en la cohorte de edad del individuo y/o a su rendimiento previo; El deterioro viene acompañado a menudo de cambios en el comportamiento y, además, el deterioro es de una gravedad suficiente como para interferir en las actividades de la vida diaria de la persona. El origen de los deterioros incluye diversas afecciones del sistema nervioso, infecciones, traumatismos, abuso de sustancias, deterioro neurocelular y carencias nutricionales. La causa más común de demencia es la enfermedad de Alzheimer (EA). El Alzheimer es una alteración neurodegenerativa primaria que se produce por el acúmulo patológico de ciertas proteínas en determinadas zonas del cerebro y la muerte de las células infectadas.

Los criterios diagnósticos tanto del DSM-V (American Psychiatric Association, 2013) como de la CIE-11 (OMS, 2022) coinciden en señalar que la demencia debida a la enfermedad de Alzheimer tiene un inicio insidioso y una progresión gradual cuya queja inicial suele ser un déficit de memoria y aprendizaje, al que se suma el de otros dominios cognitivos. En 2011, la Alzheimer's Association y el National Institute on Aging (NIA-AA) emitía unos criterios similares (McKhann et al., 2011). Sin embargo, a partir de 2018, se abandona una definición basada en las consecuencias clínicas de la enfermedad, en favor de una basada en la patología subyacente medida a través de biomarcadores (Jack et al., 2018). La razón de este cambio se argumenta sobre el elevado grado de error que se produce utilizando criterios basados en la conceptualización de la enfermedad como

síndrome. La enfermedad de Alzheimer se diagnostica con seguridad sólo a través de una autopsia, lo que ha revelado que muchas de las personas diagnosticadas clínicamente no muestran posteriormente rasgos neuropatológicos de la enfermedad, y que muchas personas sin síntomas de deterioro sí que muestran tales cambios al fallecer (Nelson et al., 2011). Así, la propuesta de los nuevos criterios NIA-AA se basa en el uso de tomografía por emisión de positrones y el análisis del líquido cerebrospinal que permita detectar una acumulación patológica de péptido beta amiloide y proteína tau. Por este motivo, la demencia en general, y la demencia tipo Alzheimer en particular, es para la ciencia uno de los mayores retos de futuro.

De forma similar, el deterioro cognitivo leve (DCL) o trastorno neurocognitivo leve para el DSM-V y la CIE-11, es una entidad que hace referencia a un déficit objetivo de uno o más dominios cognitivos de la persona, pero que no es lo suficientemente grave como para afectar al funcionamiento y las actividades de la vida diaria del mayor. Sobre este trastorno se han posado todas las miradas que tratan de realizar un diagnóstico temprano de la demencia, ya que este deterioro puede tener una entidad propia o, por el contrario, ser una etapa temprana en el desarrollo de la EA. Así, las recomendaciones de la NIA-AA descartan la categoría de DCL debido a enfermedad de Alzheimer, por la nueva Enfermedad de Alzheimer con presencia de un Deterioro Cognitivo Leve, haciendo hincapié de nuevo en la causa y no en la manifestación clínica.

A pesar de ser ésta una visión acertada, lo cierto es que el diagnóstico a través de biomarcadores es costoso e invasivo, como reconoce la NIA-AA en su propio informe, y queda a la espera del desarrollo de nuevos biomarcadores que corrijan esos puntos débiles, siendo, en el fondo más unas recomendaciones del camino a seguir que unos criterios establecidos. Añaden, además, que tampoco son criterios prácticos para muchos

grupos de investigación, dejando la puerta abierta a la evaluación cognitiva como método de diagnóstico.

Tradicionalmente, el método utilizado de diagnóstico por excelencia ha sido la evaluación neuropsicológica por medio de baterías de test. Las baterías presentan la ventaja de que permiten realizar un perfil de las capacidades y, por tanto, los déficits del mayor, acordes al criterio de diagnóstico establecido por la batería. Permiten así, un trabajo individualizado y una monitorización de su evolución. Sin embargo, presentan el inconveniente de que requieren mucho tiempo en su administración y, por tanto, tienen un elevado coste, además de que a veces los resultados son poco clarificadores.

En este contexto surge la necesidad de buscar métodos fiables para detectar las demencias que tengan un bajo coste tanto económicamente como en tiempo del profesional evaluador. Uno de los modos más populares de abordaje son los test breves, que permiten hacer un cribado para detectar a las personas con riesgo de padecer un trastorno de estas características. Así, es posible seleccionar a las personas que son susceptibles de necesitar las evaluaciones neuropsicológicas o las pruebas de biomarcadores más costosas. Se tratan de cuestionarios con una duración máxima de 20 minutos en los que se hace una valoración de varios dominios cognitivos como es el caso del Mini-Mental State Examination (MMSE), el más popular, u otros como el Montreal Cognitive Assessment (MoCA) y el Addenbrooke's Cognitive Examination Revised (ACE-R), o bien se valora un único dominio relevante, en el caso de la EA, memoria, como en el Memory Impairment Screening (MIS, De Roeck et al., 2019).

Una vertiente más reciente del cribado de demencias recae sobre el uso de nuevas tecnologías. En primer lugar, la forma más sencilla es la realización de evaluaciones neuropsicológicas a distancia a través de teléfono o videollamada, forma que ha demostrado tener la misma validez que las evaluaciones presenciales (Hunter et al., 2021)

y que presenta aspectos positivos como el de llegar más fácilmente a poblaciones aisladas como las áreas rurales (Barth et al., 2018). Otra forma, es la administración de pruebas a través de ordenadores, que de forma automatizada arrojan resultados muy similares a los test tradicionales realizados en papel (Chan et al., 2018). Una última forma destacable es la fusión de la informática y las tecnologías de comunicación con la salud, en lo que conocemos con términos como *e-Health* y *Digital Health*. Aunque estos términos no son nuevos y ya vienen recogidos en la resolución de la OMS WHA58.28 (Organización Mundial de la Salud, 2005), han cobrado especial relevancia a partir de los confinamientos de 2020. Estos términos comprenden ámbitos como la asistencia sanitaria virtual y registros informáticos y electrónicos de salud. A través de la recogida de diversos datos del usuario, se han desarrollado herramientas de cribado tanto de demencias como de otros aspectos de la salud de los mayores.

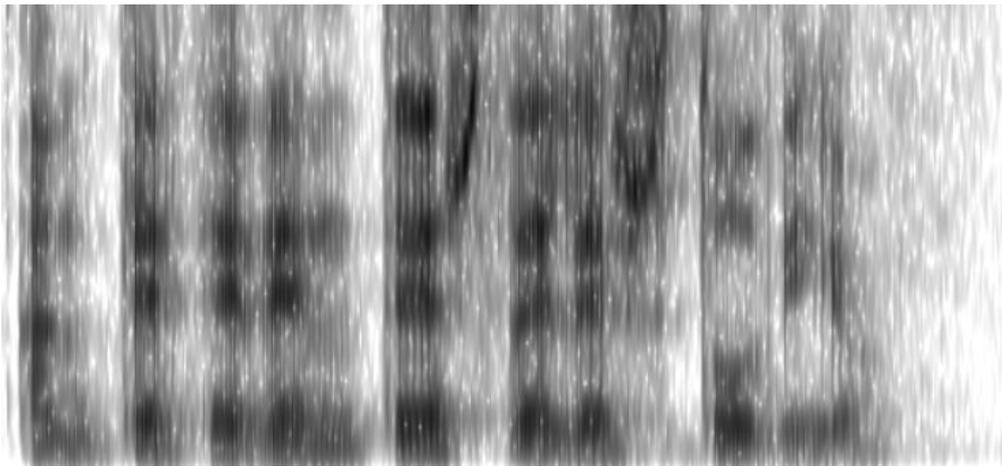
Uno de estos registros, en el cual vamos a enmarcar esta tesis, es el de la señal del habla. Así, como veremos, en los últimos quince años hemos participado en el desarrollo de tecnologías de análisis automático del lenguaje oral que permiten caracterizar e identificar patologías a través de los cambios que se producen en el habla de una persona. El estudio de los cambios en el lenguaje oral debido al desarrollo de la enfermedad de Alzheimer tiene un largo recorrido. El lenguaje sufre diferentes alteraciones a lo largo del desarrollo de la EA. La primera que se hace evidente es la anomia y, como consecuencia, un déficit en la fluidez verbal de la persona (Weiner et al., 2008). En las fases leves, los pacientes muestran problemas de comprensión y de producción, su anomia se agrava y su lenguaje se caracteriza por parafasias, circunloquios y falta de contenido. La producción verbal se vuelve ininteligible cuando aparecen otros síntomas como la disartria, la ecolalia, la palilalia y la falta de coherencia en las últimas fases de la enfermedad. Por

último, la etapa final se caracteriza por el mutismo y un grave déficit de comprensión, lo que impide cualquier tipo de interacción social (Kempler, 1995).

El desarrollo de las técnicas de análisis del habla ha permitido explorar la producción oral en los pacientes con EA. Estas técnicas tratan de aprovechar las consecuencias conductuales de los déficits cognitivos en la ejecución vocal del paciente para poder identificar los procesos dañados y, consecuentemente, la patología. Según los estudios, el habla en la EA se caracteriza por ser lenta y fluctuante, monótona, con temblores y vacilaciones, menor intensidad y pérdida del control de la fluidez (Pulido et al., 2020). Todos estos cambios se deberían a las dificultades para encontrar la palabra adecuada, y otras alteraciones en los procesos implicados en la planificación del lenguaje y en la forma en que las palabras deben asociarse entre sí semántica y sintácticamente. La creciente popularidad de este sistema se debe a que la producción vocálica es una de las primeras funciones afectadas por procesos neurodegenerativos debido a alteraciones sutiles de la planificación del habla y el control motor, alteraciones presentes en los estados preclínicos de la EA. El análisis automático del habla permite detectar dichas alteraciones vocálicas de forma rápida y eficaz en cuestión de minutos.

Partiendo de estas ideas generales, en este estudio pretendemos profundizar en el conocimiento sobre las alteraciones del habla en personas mayores con DCL y/o EA y utilizar este conocimiento de forma práctica para proponer herramientas que permitan el cribado de la enfermedad. Plantearemos cuáles son los parámetros de habla relevantes, a qué se deben sus cambios, y cómo podemos combinarlos en algoritmos de cálculo útiles para detectar procesos patológicos en el envejecimiento.

2. Antecedentes del estudio



2. Antecedentes del estudio

Hablar es aparentemente una conducta sencilla, de fácil aprendizaje, casi sin errores en su uso, y con un procesamiento muy rápido. Sin embargo, hablar requiere de la coordinación de muchos sistemas de procesamiento que la convierten en un acto cognitivo tremendamente complejo. Hablar requiere integrar esta información acústica auditiva con la de otras modalidades sensoriales como la vista en la lectura o el tacto en el braille. Requiere también planificar representaciones mentales basadas en el pensamiento, la emoción y la intención comunicativa. Además, hablar implica la planificación sintáctica, que entrelaza lo que se va a decir con el cómo decirlo en códigos o reglas compartidos con otros hablantes de una lengua. Requiere buscar en la memoria conceptos, palabras, morfemas, sonidos silábicos y producirlos con una prosodia adecuada. Finalmente implica la coordinación motora de todos los músculos que intervienen en la producción física del sonido que se implementarán en los sonidos fonológicos del lenguaje humano. Toda esta complejidad se expresa en una única señal acústica que puede ser analizada desde todos los ámbitos lingüísticos propuestos.

Uno de los más recientes e innovadores procedimientos que se vienen utilizando es el de sistemas automáticos de análisis mediante software informático como por ejemplo el Praat, el openSMILE toolkit, el Geneva Minimalistic Acoustic Parameter Set (eGeMAPS) y muchos otros. Estos programas permiten obtener de forma rápida cientos de parámetros que son, en esencia, la medición de propiedades matemáticas de la onda sonora que están relacionadas con diferentes dimensiones del habla.

Estas medidas del habla son parámetros que el oído humano no puede percibir de forma natural, ya que percibe la voz como un todo (sonido, habla y lenguaje), lo cual es útil para hacer la comunicación eficiente, pero nos impide distinguir cada componente de

la voz. Afortunadamente, los avances tecnológicos en el campo del análisis automático del habla han permitido la extracción objetiva de estos parámetros mediante un método basado en la teoría de la fuente-filtro (Fant, 1970). Según esta teoría, la señal acústica del habla se entiende como una señal de sonido fuente producida en la glotis que es filtrada a través de las resonancias con los órganos articulatorios del tracto vocal colocados en distintas posiciones o configuraciones. Es decir, la señal acústica, generada en las cuerdas vocales y compuesta por frecuencias y armónicos, se moldea en el tracto vocal a través de la prosodia en una secuencia temporalmente organizada dando lugar a formantes o claves acústicas fonéticamente reconocibles junto con el ruido vocal. Esta técnica se centra en el análisis acústico de los formantes y el tono que son percibidos como rasgos perceptivos por el sistema auditivo y procesados como rasgos prosódicos suprasegmentales del habla: entonación, ritmo y acento.

En el análisis automático del habla se abordan tanto los elementos segmentales, esto es, los fonemas o sílabas que podrán verse reflejados por ejemplo en el análisis de formantes o frecuencias características del sonido, como los suprasegmentales o prosódicos temporales. Las características segmentales han sido tradicionalmente las más utilizadas en el estudio foniátricos de las patologías físicas del habla. Sin embargo, es la prosodia el elemento que suscita más investigaciones cuando hablamos de desarrollo del habla con trastornos en procesos cognitivos o, en sus deterioros, como puede ser la demencia. La prosodia hace referencia a las propiedades fonéticas y fonológicas del discurso que no se deben a la elección de los elementos acústicos del léxico, sino del modo en que estos elementos se relacionan entre sí semántica o sintácticamente, cómo se agrupan rítmicamente, dónde pone el énfasis el hablante o qué tipo de acto comunicativo codifica (Ladd, 2008; Wagner & Watson, 2010). También dará información contextual, tanto del contexto gramatical dentro del discurso, contribuyendo al significado y a cómo

se establecen los turnos en la conversación, como del contexto personal del hablante reflejando su actitud y estado emocional (Cole, 2015).

En este apartado presentaremos las bases teóricas y prácticas que sustentan este programa de investigación. Comenzaremos con una introducción al método del análisis automático del habla, seguido por un repaso a su utilidad como herramienta de cribado cognitivo, introduciremos las alteraciones del lenguaje asociadas a la demencia tipo Alzheimer que supondrán los procesos lingüísticos que provocan los cambios en el habla y finalizaremos el apartado con una revisión sistemática sobre el uso del análisis automático del habla en el estudio del deterioro cognitivo leve y la enfermedad de Alzheimer.

2.1 Producción y análisis del habla

Esta primera sección de la introducción tiene como objetivo exponer los fundamentos de la producción del habla sobre las que se sustenta el análisis acústico. Presentaremos los distintos elementos que se examinan, cuáles son los parámetros de habla fundamentales y cómo se extraen a partir de una señal de sonido. Además, introduciremos su uso en el estudio de procesos cognitivos y expondremos el debate sobre su posible consideración como biomarcadores.

El análisis del habla es el proceso de extraer información relevante a partir de la señal de sonido producida al hablar. Esa señal de sonido contiene información lingüística que el emisor trata de transmitir y que se ha codificado mediante una serie de movimientos de órganos articuladores. Producir habla es, entonces, emitir sonidos articulados que conocemos como palabras, que además entrelazamos de forma que tengan sentido en un discurso. Es un proceso complejo con múltiples sistemas implicados, empezando por el sistema nervioso participando mediante el procesamiento de la información a transmitir, que además controla los órganos fonatorios encargados de producir el sonido, hasta la intervención de sistemas de percepción como el propio oído en forma de retroalimentación. Ese sonido producido implica la expulsión de aire de los pulmones, que, aunque habitualmente se expulsa en silencio, mediante la configuración del tracto vocal podemos cambiar sus cualidades logrando modelar las características de su onda.

El sonido es efectivamente una onda de presión producida por la vibración de un cuerpo que en el caso de la voz humana se transmite a través del aire. Al expulsar el aire de los pulmones, la vibración de las cuerdas vocales producirá la vibración de las partículas del medio produciendo las ondas de sonido, que podemos definir a través de dos características fundamentales: la longitud y la energía. La longitud de onda es la

distancia entre dos puntos equivalentes de la onda. Esto determinará la frecuencia, que es la distancia entre dos puntos parejos y seguidos de la onda partido por el tiempo que se tarda en producir, y se medirá en hertzios o ciclos por segundo. En la Figura 1, a la izquierda, encontramos representada una onda con el total de su longitud, en la que la línea discontinua marca los puntos equivalentes. Longitudes de onda más largas significará frecuencias más bajas, es decir, sonidos más graves; y las más cortas producirán frecuencias más altas o sonidos agudos. En la misma Figura 1, en la parte derecha, se representa una onda de 10 Hz, más grave, y la de la derecha, de 20 Hz, más aguda.

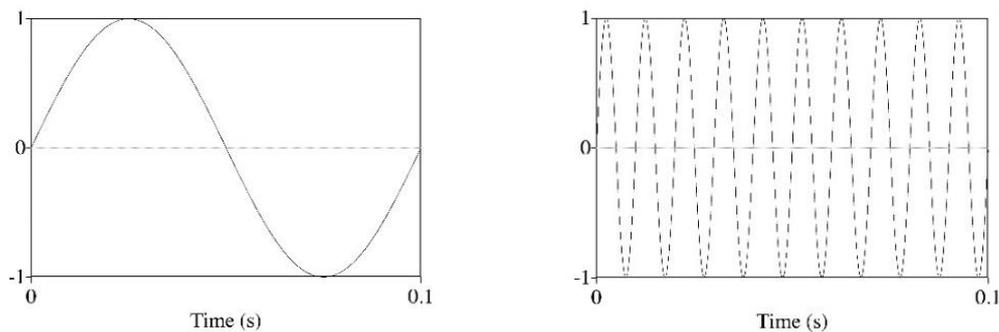


Figura 1. “Ondas sonoras periódicas simples con la misma amplitud y con diferente frecuencia” por Joaquim Llisterri licencia bajo BY CC 4.0

La otra característica es la energía, definida como magnitud que hay entre el punto mínimo y el máximo de la onda, de modo que cuanto más amplitud haya, percibiremos mayor intensidad, que habitualmente se mide en decibelios. En la figura 2 observamos ondas de 100 Hz con amplitudes distintas.

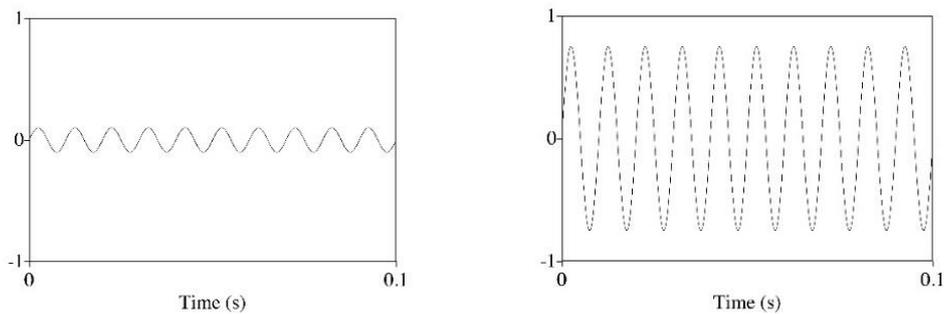


Figura 2. “Ondas sonoras periódicas simples con la misma frecuencia (100 Hz) y con diferente amplitud” por Joaquim Llisterri licencia bajo BY CC 4.0

La voz humana, además, es un sonido complejo, es decir, formado por la superposición de numerosas de estas ondas, y su interrelación dará lugar a toda la variedad de sonidos que producimos. Será la medición de estas dos características y muchas otras derivadas, las que se abordan mediante el análisis del habla dando lugar a decenas de parámetros. A partir de este punto nos referiremos a grupos de parámetros basados en diferentes dimensiones del habla, cuyas bases describiremos posteriormente:

- Parámetros de duración y temporales: se incluirán aquellas características del habla que afecten al tiempo de emisión del mensaje del hablante.
- Parámetros de ritmo: serán aquellos que se refieran a la organización y distribución del habla en su tiempo correspondiente.
- Parámetros de intensidad: son aquellos que se refieren al volumen y sus variaciones.
- Parámetros de frecuencia: incluiremos parámetros derivados de la frecuencia y las medidas de su variación. Además, esta categoría contendrá parámetros de perturbación de la vibración de la glotis.
- Parámetros del análisis del espectro: serán aquellos que miden el ajuste de las frecuencias y la intensidad.

Para generar habla, podemos distinguir dos fases. La primera fase tiene que ver con una serie de procesos mentales de planificación del mensaje. Esto incluye seleccionar la información semántica que el hablante tiene intención de transmitir. Además, debe convertir ese mensaje en una serie de palabras, cada una de las cuales está formada por una secuencia de fonemas y deben estar ordenadas gramaticalmente de forma que tengan sentido. Esto dará lugar además a una frase que debe ser pronunciada con una prosodia concreta con la duración, entonación e intensidad deseadas. En la segunda fase, se produce una serie de señales neuromusculares con el objetivo de controlar la musculatura implicada en la articulación de los sonidos. Estas señales provocan las configuraciones determinadas que darán lugar a una señal acústica continua que es lo que conocemos como habla (Docio-Fernandez & Garcia-Mateo, 2009).

En la producción de habla, desde un punto de vista fisiológico, van a participar varios sistemas que forman el aparato fonador en los que vamos a distinguir los pulmones, la glotis o abertura superior de la laringe, donde se encuentran las cuerdas vocales, y el tracto articulatorio formado por la cavidad bucal y nasal. Cada uno de estos componentes tendrán una función en la producción del habla. En primer lugar, los pulmones son la fuente de aire durante el habla. En la glotis, las cuerdas vocales pueden vibrar de forma periódica dando lugar a los sonidos sonoros como los de las vocales, o tensarse e interrumpir el paso de aire creando los sonidos sordos como los de algunas consonantes. A continuación, el sonido viajará al tracto vocal y/o nasal pasando por el velo del paladar, si este permite el paso al tracto nasal se producirán sonidos nasales como el de la *n*. De otro modo, pasará al tracto vocal donde encontraremos otros resonadores como el paladar, los dientes, o la lengua, que funcionará como un articulador flexible posicionándose en relación a los otros dos. Finalmente, la forma de los labios, más o menos redondeada o abierta, determina también el paso del aire afectando al sonido de las vocales o

produciendo consonantes como la *p*. En la Figura 3 podemos ver un esquema del aparato fonador y los elementos del habla con los que se relaciona cada parte.

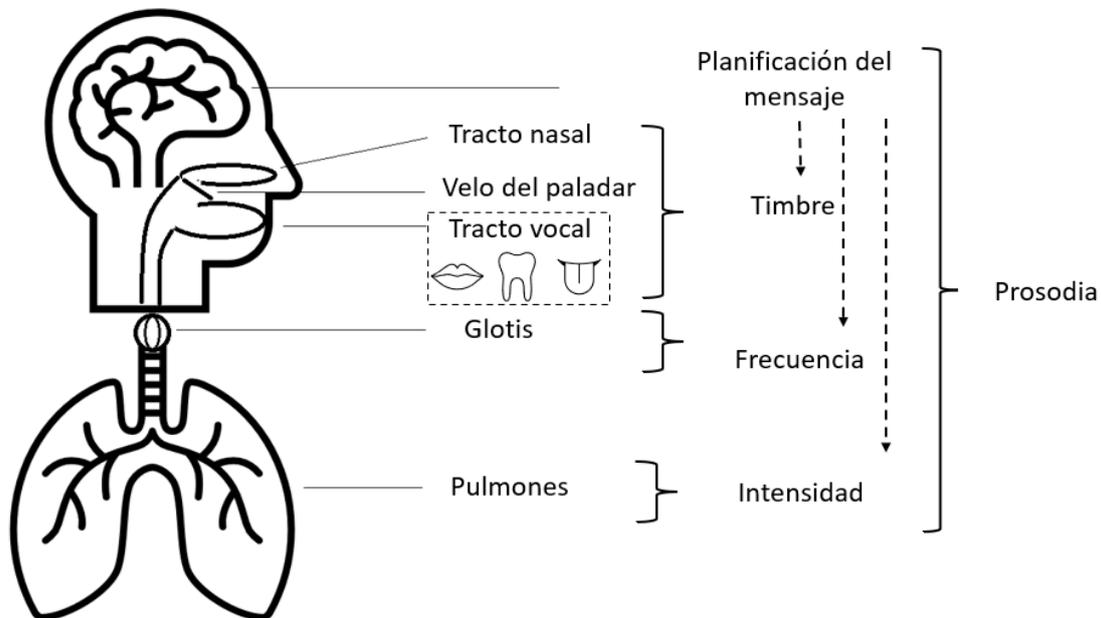


Figura 3. Representación del aparato fonador y su relación con parámetros del habla.

Dentro de los sonidos del habla, como hemos visto, podemos distinguir los sonoros y los sordos. Hemos apuntado que lo que determinará cuál se produce es la acción de las cuerdas vocales. Los sonidos sonoros, como los de las vocales o algunas consonantes como la *m* o la *d*, se producen cuando hay vibración de las cuerdas vocales, y los sordos, como el de la *t*, cuando esa vibración sufre un retardo. Al producir los sonidos sonoros, el ritmo de la vibración de las cuerdas vocales determina lo que llamamos frecuencia fundamental o formante 0 (F_0); este es el factor que más contribuye a la percepción del tono del habla y se conoce así porque establece la línea base de los armónicos. Los armónicos son componentes de una onda sonora compleja, como la del habla, y se considera que la frecuencia de los armónicos es un múltiplo entero de la frecuencia fundamental. Por ejemplo, suponiendo que la F_0 sea de 100 Hz, el segundo armónico será el resultado de multiplicar la frecuencia por dos, es decir, 200 Hz. Para

obtener el tercer armónico se multiplicará por 3, obteniendo 300 Hz, y así sucesivamente. Una vez que la onda empieza a resonar contra los órganos articulatorios, su configuración determinará cuáles de estos armónicos resuenan y adquieren mayor amplitud y cuáles se atenúan. A los armónicos resultantes, que además serán los típicos de las vocales los llamamos formantes y se representarán, según se alejen de la línea base, como primer formante F1, segundo formante F2, y siguiendo F3, F4 y F5. Ilustramos los formantes en la Figura 4 a través de un espectrograma. Los espectrogramas son una de las formas más comunes de representar el habla, se trata de un modelado gráfico del espectro del sonido. Un espectrograma es un diagrama que muestra las amplitudes relativas de las diferentes frecuencias componentes de un sonido. Así, las zonas más oscuras significan mayor energía para ese tiempo y frecuencia. En el ejemplo de la figura 4 vemos la representación del sonido de las vocales en español en su variante europea. Las zonas más oscuras que se aprecian como líneas horizontales se corresponderían con los formantes. Indicado con líneas verdes podemos ver el F1 de cada vocal, y con líneas naranjas el F2. Además, representado mediante una línea azul observamos la F0.

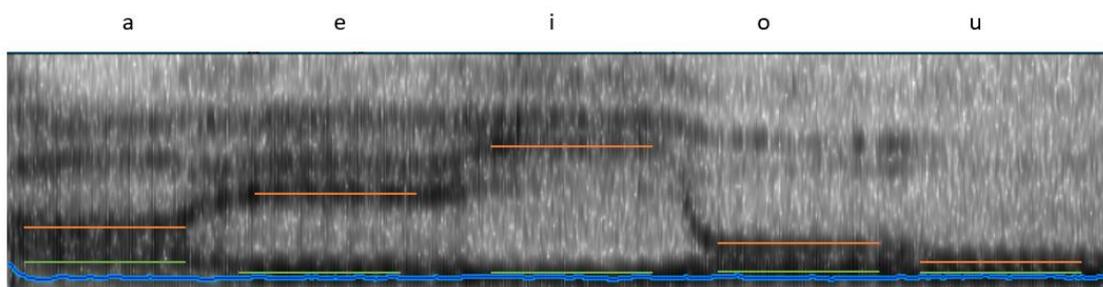


Figura 4. Espectrograma de las cinco vocales del español en su variante europea.

Los formantes, además de por su frecuencia y amplitud, se caracterizan por su ancho de banda (al que haremos referencia como *bandwidth* o B1, B2, etc.), cuanto menor sea el ancho de banda, más selectivo es el armónico y se percibirá su frecuencia más claramente (Murtró, 2010). El sonido de cada fonema, además, vendrá determinado por

el de los adyacentes en lo que llamamos coarticulación, es decir, al producir cadenas de fonemas, los movimientos de los diferentes articuladores se afectan mutuamente, y la articulación del fonema actual estará condicionada por cuál es el anterior y el siguiente.

Un elemento medido habitualmente en el análisis del habla es la perturbación de la vibración de la glotis. De entre los parámetros de este tipo destacan el jitter, el shimmer, la ratio de armónicos-ruido (HNR) y las roturas de voz. El jitter es la variación de la frecuencia entre dos pulsos glotales consecutivos y tendría que ver con la falta de control de la vibración de las cuerdas vocales. En cambio, el shimmer hace referencia a la variación de la amplitud y se relacionaría con una reducción de la resistencia de la glotis al paso del aire. El HNR se trata de la relación entre los componentes periódicos y aperiódicos de un segmento, es decir, de la proporción de sonido con vibración de las cuerdas vocales y ruido glótico. Un HNR bajo se relacionaría con disfonías. Por último, las roturas de voz medirían la falta de capacidad para mantener una fonación estable, de modo que se considera una rotura de voz cuando se da una distancia anormal entre dos pulsos glotales (Teixeira et al., 2013). Todas estas medidas se suelen considerar dentro de lo que llamamos calidad de voz y que determina el timbre del hablante. Otro modo de medir calidad de voz es mediante el análisis del espectro de sonido, como el de la Figura 4. El espectro evidencia la relación de las resonancias entre sí, es decir, cómo amplifican o atenúan los armónicos, en una interacción entre las frecuencias y la amplitud. En este sentido, la calidad de voz, más que con elementos concretos del espectro, tiene que ver con su distribución general, y viene determinada por las configuraciones y ajustes laríngeos y de las cavidades supraglóticas.

La calidad de voz es probablemente la dimensión más abstracta del habla y también la que más variedad de medidas arroja en los diversos estudios. Aquí nos vamos a centrar en los parámetros de análisis del espectro que utilizaremos en esta investigación.

Algunos de los parámetros más utilizados se refieren a los momentos espectrales o la distribución de la energía de las frecuencias en el espectro, como el centro de gravedad y la asimetría. El centro de gravedad del espectro es una medida de la altura media de las frecuencias de un espectro. Por su parte, la asimetría sería la diferencia entre la forma del espectro por debajo del centro de gravedad y la forma por encima de ese punto. Los valores positivos de asimetría indicarán que hay mayor predominio de la energía en las zonas de frecuencias bajas, mientras que valores negativos indican mayor energía en las frecuencias altas. En la Figura 5 podemos observar dos ondas con sus frecuencias representadas en el eje de abscisas y la amplitud en el de ordenadas. La representación de la parte superior se corresponde con una onda simple de 1000Hz, al tratarse de una única onda la media de sus frecuencias será también 1000Hz, hemos marcado el centro de gravedad con una línea roja. En la parte de abajo tenemos la representación de una fricativa, la *s*, que al ser una onda compleja muestra una gran distribución. Concretamente, es una onda con centro de gravedad en 2304,60 Hz y una asimetría de 1,47. Por tanto, lo que vemos es que la intensidad en las ondas las frecuencias bajas (a la izquierda de la línea roja) es mayor.

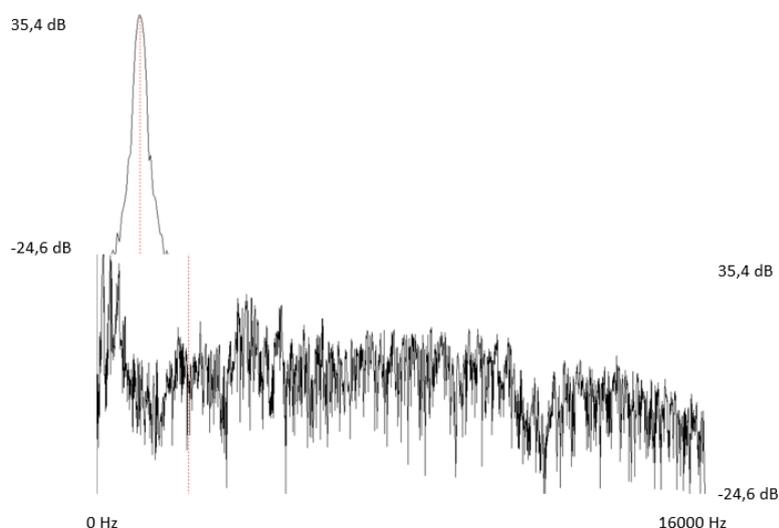


Figura 5. Representación de la distribución de la intensidad y las frecuencias de una onda simple y una compleja.

Por otro lado, tenemos el Índice Acústico de Calidad Vocal (acoustic voice quality index, AVQI) que es, en realidad, una combinación de seis parámetros de calidad de voz que correlacionan con la gravedad de una disfonía (Maryn & Weenink, 2015). La energía de banda (Band Energy, BE), es la energía en cada una de una colección de bandas de frecuencia utilizando bandas de 500 Hz superpuestas entre sí (Surendran & Levow, 2008), en el espectrograma de la Figura 6 podemos ver cada una de las bandas de 500 Hz marcadas por líneas discontinuas. Y, por último, el análisis del espectro medio a largo plazo (Long-term average spectrum, LTAS) permite ver la distribución de la frecuencia media de la energía en una muestra de habla, siendo esta última una medida que se suele considerar más fiable porque tiene en cuenta información del total de la duración del audio (Boullosa & Pérez Ruíz, 1984).

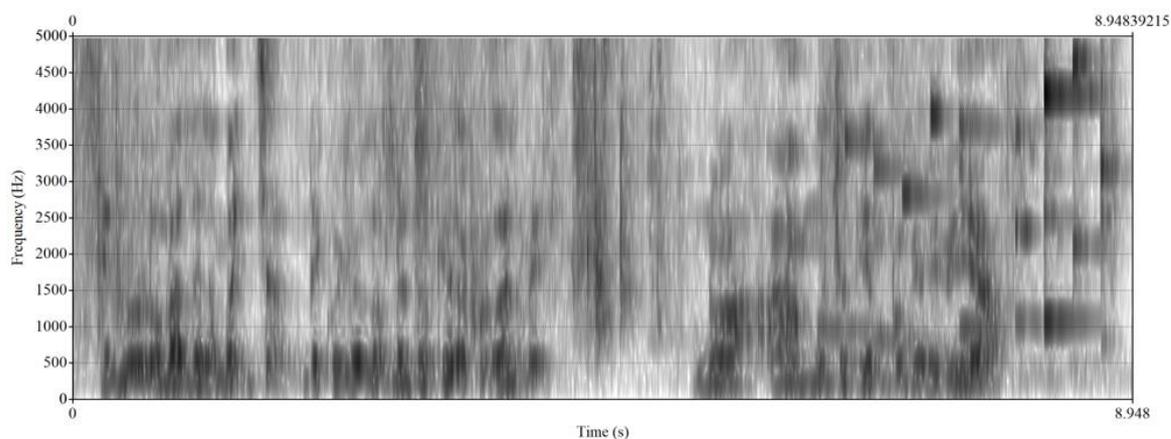


Figura 6. Espectrograma con bandas de 500 Hz.

Estos son los ingredientes básicos de la producción de voz que determinarán la producción de fonemas, pero serán las relaciones establecidas entre esos elementos para formar sílabas, palabras y frases completas las que nos van a permitir caracterizar el habla. Hasta ahora, por tanto, hemos descrito el análisis de los elementos segmentales del habla que implican únicamente fonemas. Los niveles que afectan a más de un segmento o fonema se denominan suprasegmentales y en ellos se incluyen las sílabas y palabras, pero

también los elementos de la prosodia: acento, tono, melodía, ritmo, pausas y velocidad. Desde el punto de vista del análisis del habla, habitualmente los interpretaremos a través de parámetros de duración, de fluidez y ritmo, y que tienen que ver con la dimensión temporal, así como melódicos y de la calidad de voz o timbre, relacionados más estrechamente con la frecuencia.

Los parámetros de duración hacen referencia a la cantidad de tiempo que tarda el hablante en emitir un mensaje. Es común hablar del tiempo total que tarda por ejemplo en leer un texto o describir una lámina. Dentro de ese tiempo se pueden incluir elementos durante los cuales se emite sonido, o segmentos con voz, y también momentos en los que no, o segmentos silenciosos o sin voz. Una forma de medir los segmentos con voz será el tiempo de fonación definido como la duración en segundos de los núcleos intra e inter-silábicos sin contar las pausas. Las pausas, son uno de los parámetros temporales por excelencia en el análisis del habla. Una pausa es una interrupción breve en la emisión de habla, que, en el análisis del habla se mide como un momento en el que la intensidad del habla es indistinguible de la intensidad del ruido de fondo. De las pausas es habitual medir su duración, variabilidad y cantidad.

Dentro de los parámetros de fluidez y ritmo debemos mencionar los que tienen que ver con la medición de las sílabas. Una sílaba es una cadena corta de segmentos o sonidos que está compuesta por un núcleo, habitualmente una vocal, y puede ir precedido o seguido de otros sonidos, habitualmente consonantes. Cada sílaba es el resultado de un solo impulso de energía muscular. Una sílaba habitualmente va a estar compuesta de una parte inicial de intensidad creciente llamada ataque, el núcleo o parte de mayor amplitud sonora, y una parte decreciente en intensidad conocida como coda. Aunque mediante el análisis del habla no se pueden identificar las palabras dado que la emisión es continua y no hay un marcador de cambio, a no ser que por algún motivo se produzca una pausa que

las preceda, las sílabas sí se pueden detectar midiendo esas variaciones de la intensidad y del espectro alrededor del núcleo. De este modo, se puede identificar el número de sílabas producidas, cuánto duran y la variabilidad de su duración. En la Figura 7, la línea amarilla representa la intensidad y sus variaciones a partir de las cuales se estima el núcleo de las sílabas. Las sílabas, además, se pueden utilizar como medida del ritmo de la producción de habla. Un parámetro comúnmente usado con este fin es el índice de variabilidad por pares (Pairwise Variability Index, PVI), una medida que se concibe como la media de las diferencias en la duración de dos intervalos vocálicos seguidos. De ello se extrae una medida bruta (rPVI), de la que se obtiene también una medida normalizada (nPVI) dividiendo el rPVI entre la duración media de los pares (Grabe & Low, 2008). El PVI se utiliza normalmente como indicador de la ritmicidad de una lengua. El español, de hecho, se considera una lengua con un ritmo muy regular, de modo que la alteración del PVI en un hablante se considera como un indicador patológico.

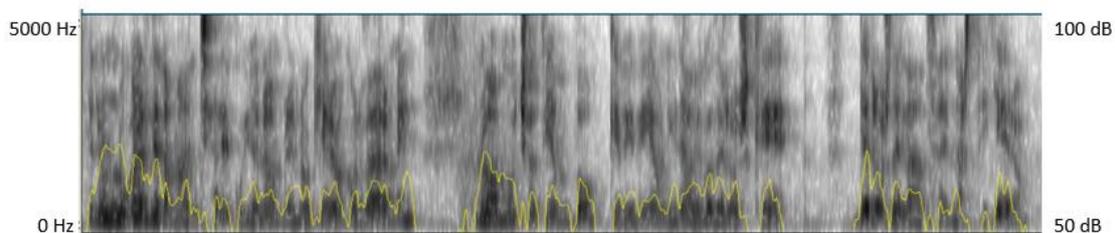


Figura 7. Espectrograma de un hablante leyendo el principio del Quijote.

A partir de la medida de las sílabas se pueden obtener otros indicadores generales del ritmo y la fluidez del habla como la velocidad de habla y de articulación. La velocidad de habla se mide como el número de sílabas producidas partido por el tiempo de habla sin las pausas. La velocidad de articulación, por su parte, se obtiene dividiendo el número de sílabas entre el tiempo de fonación.

Una dimensión más de la prosodia que podemos caracterizar con el análisis del habla es la melodía o el acento y la entonación. Esto hace referencia a las variaciones de la frecuencia fundamental a lo largo del tiempo. Se puede medir a través de los picos y valles del tono. Un ejemplo de esta clase de parámetros son los que miden la trayectoria del tono, bien dentro de los núcleos silábicos, conocido como TrajIntra, o entre ellos, en cuyo caso se suele denominar como TrajInter.

Como hemos visto, el objetivo de este apartado era introducir las bases de la producción de habla y cómo podemos analizar las características físicas de la onda de sonido producido para extraer información lingüística. Hemos introducido así algunos de los conceptos más relevantes en los estudios del habla y que serán usados en este estudio. Como recapitulación, en la Figura 8 se han ordenado los parámetros recordando la clasificación propuesta al inicio.

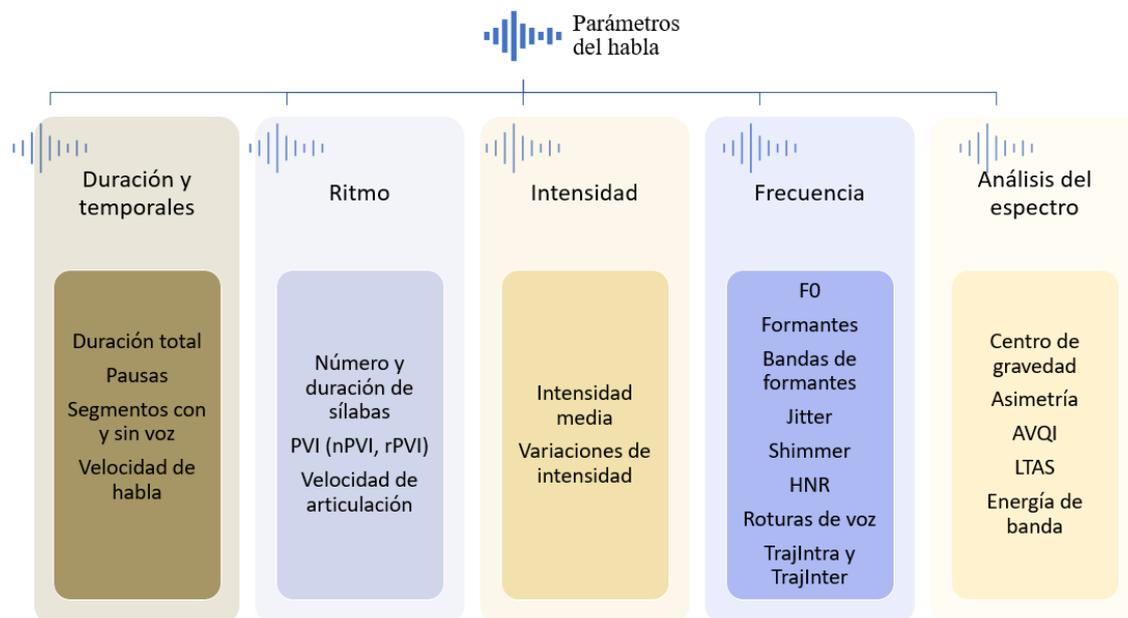


Figura 8. Parámetros y su clasificación.

Además, como hemos mencionado anteriormente utilizaremos el término calidad de voz para referirnos a parámetros de perturbación de la F0 y los del análisis del espectro.

El análisis del habla y los procesos cognitivos

Hasta aquí hemos planteado la producción de habla en un sentido fisiológico para comprender qué es lo que permite que podamos analizarla a través de las propiedades físicas del sonido. De ello se desprende que las características del sonido dependen de configuraciones musculares. En este sentido, la información contenida en la señal acústica constituye una medida comportamental objetiva que se puede medir en los parámetros extraídos a través de los softwares de análisis del habla. Pero, además, parece razonable argumentar que esas configuraciones estarán determinadas por señales neuromusculares que dependen en último término de la planificación motora cerebral (Correia et al., 2020; Tremblay, Deschamps, et al., 2019). Y, por otro lado, que hay parámetros que no dependen de esas configuraciones, sino de las intenciones y capacidades del hablante. Así, en el habla se plasmarían los procesos cognitivos necesarios para producirla. Esta ha sido siempre una asunción teórica basada en los procesos esenciales para emitir una prosodia adecuada. Pero más allá de las argumentaciones teóricas, existen estudios que exploran empíricamente la relación entre el rendimiento en determinados procesos y la ejecución del habla (Cohen et al., 2016; Suzuki, 2021).

Hay diversas aproximaciones a la medición de procesos cognitivos a través del análisis automático del habla. Por un lado, se pueden encontrar trabajos en los que se relaciona un parámetro con el rendimiento en pruebas cognitivas. Por ejemplo, se ha comprobado que la fluidez verbal semántica muestra una relación con parámetros como la F0, las pausas y la variabilidad de la intensidad (Cohen et al., 2016). Una variante de estos estudios es la que utiliza el rendimiento en las pruebas neuropsicológicas para explicar el desempeño en el habla de los aprendices de una segunda lengua. Así, podemos encontrar trabajos que indican que el rendimiento en memoria fonológica y asociativa

predice el número de pausas al hablar en la segunda lengua (Granena & Yilmaz, 2019; Suzuki, 2021), o que la memoria de trabajo predice la ejecución de parámetros de fluidez (Nielson & DeKeyser, 2019). Otro bloque de trabajos sería el de los que manipulan el material o la tarea de producción de habla. Aquí incluiríamos las investigaciones sobre los cambios en la velocidad del habla y los parámetros espectrales al aumentar la carga cognitiva, es decir, comprometer los recursos de la memoria de trabajo (Dahl & Stepp, 2021; MacPherson et al., 2017). Tendríamos además las investigaciones en las que se relaciona la prosodia con el contenido del discurso, de las que pueden servir como ejemplo las que relacionan menor velocidad de habla con el uso de palabras de menor frecuencia y estructuras sintácticas más complejas (Cohen Priva, 2017).

Una cuestión de importancia es que estas explicaciones no pasan de ser exploratorias y podemos encontrar variedad de resultados para los mismos parámetros en función de diversos condicionantes como el tipo de muestra (jóvenes, mayores, con o sin patologías, etc.) o el tipo de tarea utilizada en el análisis (habla espontánea, lectura, etc.). Además, un parámetro puede encontrarse explicado por varios procesos y no va a tener una explicación única. En definitiva, se trata de costelaciones de parámetros que se relacionan con procesos concretos. Trataremos de arrojar un poco de luz sobre la relación hallada entre procesos y algunos de los parámetros más comunes en la literatura.

Los parámetros de duración son algunos de los más estudiados. De entre ellos, las pausas se han relacionado muy frecuentemente con diversos procesos alterados y dificultades en la producción de lenguaje. Las pausas aparecen como un mecanismo compensatorio ante la alteración de un proceso habitual que puede ser de búsqueda en el sistema léxico (Corley & Stewart, 2008; Gayraud et al., 2011), de planificación sintáctica o elaboración del discurso (Bortfeld et al., 2001; Grande et al., 2012) o provocado por un aumento de las demandas de la tarea. Como por ejemplo, cuando se trata de relatar

un recuerdo, las pausas aparecerían si hay alteraciones en la memoria (Pistono et al., 2016). Otros aspectos, como la velocidad de habla, pueden estar influidos también por la función ejecutiva (De Looze et al., 2021; Suzuki, 2021) o procesos atencionales (De Looze et al., 2018). La capacidad de la memoria de trabajo y la función ejecutiva influyen en la habilidad de planificar oraciones afectando a su longitud y complejidad, lo que además altera la velocidad de habla y la aparición de pausas y la longitud de las mismas (Singh et al., 2001).

Los parámetros de intensidad y su variabilidad se han relacionado con una falta de expresividad asociada a variaciones de los procesos afectivos (Zhang et al., 2020) y a una prosodia emocional alterada (Del Carmen Pérez-Sánchez et al., 2021). También se ha relacionado con la complejidad de las tareas, lo que compromete la atención y la memoria de trabajo (Abur et al., 2021). La complejidad también puede provocar cambios en otros elementos como la velocidad de articulación y la frecuencia fundamental (MacPherson, 2019). Se han observado cambios también en las bandas de los formantes debidos a una alta carga cognitiva (Huttunen et al., 2011), aunque otros estudios señalan que estos cambios tendrían más que ver con dificultades en la denominación (Thomas et al., 2020), o el control motor de los músculos implicados en la articulación (Gómez-Vilda et al., 2019). Los efectos de la carga cognitiva sobre el habla han suscitado numerosas investigaciones, habitualmente comprometiendo los recursos cognitivos mediante tareas duales, y han encontrado, fundamentalmente, una afectación de la producción del habla reflejada en parámetros de calidad de voz. Se han hallado cambios en la distribución de las frecuencias en el espectro y la presencia de mayor ruido vocal (Le et al., 2011; Van Segbroeck et al., 2014; Yap et al., 2011).

En definitiva, a través del análisis automático del habla se pueden evaluar diversos aspectos del desempeño cognitivo y el estado emocional de una persona. Sin embargo,

aún es un campo joven con mucho que explorar. De hecho, la mayoría de los estudios han pasado por alto este paso intermedio y se han lanzado a su uso en la detección de estados patológicos proponiendo su uso como un mero biomarcador biológico sin comprender, en la mayoría de las ocasiones, los procesos que se están midiendo.

El habla como biomarcador digital

Un biomarcador se puede definir como un indicador objetivo del estado de un paciente observado desde fuera, sistema que pueden medirse de forma precisa y reproducible (Strimbu & Tavel, 2010). En contraste con los síntomas, los biomarcadores no son necesariamente un indicador de salud o enfermedad, sino que pueden estar mostrando procesos normales, como el caso del pulso. El habla, como se ha argumentado, requiere la coordinación de varios procesos cognitivos y motores, que son reflejados en una única señal que puede ser analizada objetivamente, ofreciendo información del estado mental y funcional del hablante. Es en este sentido en el que se propone que los parámetros de habla pueden ser considerados biomarcadores. Además, como estas características del habla constituyen datos de comportamiento recogidos a través de dispositivos digitales, pueden considerarse biomarcadores digitales.

Muchos investigadores se han lanzado a considerar como biomarcadores los parámetros que han identificado en sus investigaciones sobre el habla en determinadas patologías, incluso denominándolos como tal (Kourtis et al., 2019; Poellabauer et al., 2015; Zhang et al., 2020). Sin embargo, no hay actualmente ningún parámetro del habla reconocido como biomarcador por ningún organismo de salud nacional o internacional (Fagherazzi et al., 2021). En esta sección nos basaremos en las propuestas de Robin et al., (2020), Fagherazzi et al., (2021) y Ramanarayanan et al., (2022) para exponer si,

efectivamente, los parámetros del habla cumplen los requisitos necesarios para ser considerados biomarcadores y qué pasos debemos dar los investigadores en el campo del habla para consolidarlos como tal.

En el contexto del habla, un biomarcador sería un parámetro o conjunto de parámetros que se asocian con un resultado clínico y puede ser usado bien para detectar un estado patológico, o para monitorizar y clasificar su gravedad y los estadios de la alteración. Lógicamente, el biomarcador del habla debe tener las mismas propiedades que los clásicos. Así, se considera que para identificar un buen biomarcador debe evaluarse su relevancia y su validez. La relevancia es la capacidad del biomarcador para proporcionar información clínicamente relevante. La validez hace referencia a que el biomarcador se pueda medir objetivamente de forma precisa y reproducible, y que además sea útil en distintas poblaciones.

En cuanto a la relevancia, la cantidad de estudios que exponen la relación entre el habla y diversos estados patológicos es abrumadora. Se ha hallado que puede ser un método fiable para detectar enfermedades neurodegenerativas (Pulido et al., 2020), tal y como es el objeto principal de esta tesis. Para detectar el estado de ánimo, mostrando buena correlación con la gravedad de los síntomas de depresión (Zhang et al., 2020) o monitorizando los estados de un trastorno bipolar (Faurholt-Jepsen et al., 2016). Y además, midiendo afecciones físicas como enfermedades cardiovasculares (Maor et al., 2018) o la artritis reumatoide (Kosztyła-Hojna et al., 2015). En el siguiente apartado, se expondrá de forma exhaustiva un abanico de afecciones relacionadas con la salud mental en las que el habla ha demostrado una gran relevancia.

En cuanto a la validez, es donde más dudas puede plantear el habla como biomarcador. Parece razonable después de lo expuesto que el habla es medible de forma objetiva y precisa. Además, hay estudios que prueban que es reproducible y estable en el

tiempo (Cohen et al., 2012). Sin embargo, puede plantear dudas la validez externa entendida como su utilidad en distintas poblaciones. El habla es influenciada, lógicamente, por las características del idioma del hablante. Ya hemos señalado, por ejemplo, que los idiomas tienen ritmos distintos, pero también acentos, melodías y entonaciones y fonemas. A lo que podemos añadir la disparidad de tipos de habla según las tareas (sostener fonemas, ritmo diadococinético, habla espontánea, lectura, etc.). Estos condicionantes suponen un reto que deberá ser afrontado mediante la creación de procedimientos estandarizados y, como proponen Fagherazzi et al., (2021), muy probablemente los biomarcadores basados en el habla se limitarán a un idioma concreto o a grupos específicos de la población.

Teniendo en cuenta las limitaciones señaladas, parece haber cierto consenso en que con el habla nos encontramos ante un potencial biomarcador. Por ello hay un esfuerzo por llegar a líneas de trabajo comunes que permitan su generalización. Describiremos el proceso en tres fases generales: verificación, validación analítica y validación clínica (ver figura 9).

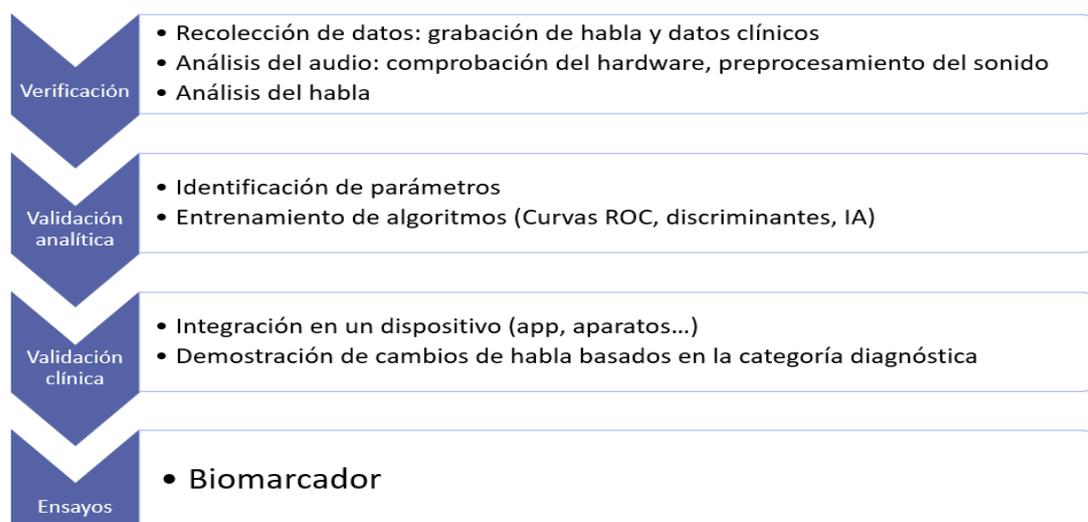


Figura 9. Proceso de validación de un parámetro del habla como posible biomarcador.

En primer lugar, la verificación es el proceso de valorar el hardware de recogida de audio, si tiene las características adecuadas para obtener unos datos con una calidad aceptable, y el preprocesamiento del audio. Hoy en día, la calidad de grabación de la mayoría de dispositivos móviles arroja un audio de suficiente calidad (Manfredi et al., 2017), pero se debe tener en cuenta que al ser una tecnología pensada para la evaluación remota, desconocemos las condiciones en las que se realizará tal audio. Por ello habitualmente se recomienda el tratamiento del audio para eliminar ruido o diversas medidas de corrección destinadas a calibrar el audio sin que pierda fiabilidad. Teniendo esto en mente, el primer paso para identificar un parámetro del habla susceptible de ser un biomarcador será el de recoger las correspondientes grabaciones y datos clínicos de los participantes.

La validación analítica es la fase en la que se comprueba que las medidas del habla efectivamente son sensibles al fenómeno estudiado. En este punto se trata de identificar los parámetros del habla que muestran una alteración en las personas con la patología objetivo, y de ellos, seleccionar los que contribuyan a distinguir a los afectados. Para ello habitualmente se usan métodos de reducción de dimensiones, análisis discriminantes o técnicas de aprendizaje automático.

Por último, la validación clínica requerirá la comprobación de la capacidad de los parámetros para diagnosticar la presencia o ausencia de la patología, o para monitorizar sus cambios en el tiempo. Es decir, requiere una comprobación de la efectividad de los datos obtenidos con la replicación de los resultados en una muestra independiente que permita generalizarlos a la práctica clínica a través de ensayos.

Es en esta última fase donde no se ha dado el paso aún que permita hablar de biomarcadores reconocidos. Veremos en la siguiente sección estudios que demuestran la relevancia y validez de diversos parámetros para una amplia variedad de trastornos que

van acompañados en general de algoritmos con sus correspondientes validaciones cruzadas y pruebas de test-retest, pero no se han llevado a cabo ensayos independientes fuera de los laboratorios. De hecho, Ramanarayanan et al., (2022) señalan que posiblemente en los ensayos con población general, quedaría comprometida la relevancia clínica de los parámetros del habla. Indican que los estudios en laboratorio utilizan un diseño de grupo control vs grupo patológico, y es probable que las medidas encontradas en estas condiciones sean sensibles, pero no específicas para una única entidad patológica. Esto da importancia a otro de los que podemos identificar como un gran reto en el estudio del análisis del habla, que es el problema de los diagnósticos diferenciales.

2.2 El Análisis Automático del Habla como herramienta de cribado en salud mental

Como hemos visto, el análisis del habla puede utilizarse entonces como una medida directa de la capacidad de procesamiento (Cohen et al., 2015), además de ser útil para explorar la posible alteración de los procesos implicados en la producción de lenguaje (Meilán et al., 2018; Pistono et al., 2019). Así, no es de extrañar que en la última década haya surgido un gran número de estudios que tratan de encontrar disrupciones en la señal del habla relacionadas con el inicio o desarrollo de diferentes patologías. El procedimiento habitual de este tipo de estudios consiste, en primer lugar, en identificar el parámetro o parámetros acústicos característicos de la patología objetivo, basándose en estudios de cohortes. A continuación, este parámetro o una combinación de ellos, se implementa en algoritmos obtenidos mediante aprendizaje automático o análisis discriminantes. En los últimos años, hemos asistido a la evolución desde una primera aproximación exploratoria al habla (Hoffmann et al., 2010; Vanello et al., 2012), hasta un rápido cambio hacia estudios de detección automática (Alghowinem et al., 2013; López-de-Ipiña, Alonso, Solé-Casals, et al., 2013). El objetivo final sería introducir este algoritmo en un dispositivo electrónico que pueda utilizarse en el ámbito clínico, aunque este es un paso que queda por dar en la mayoría de las patologías.

El interés por esta tecnología está creciendo debido a varios factores. Por un lado, el interés social: es evidente la necesidad de contar con pruebas de cribado que permitan dar una respuesta temprana y rápida a la creciente demanda de atención a la salud mental. Por otro lado, el procedimiento es rápido, sencillo y barato y, por tanto, está al alcance de un gran número de usuarios potenciales. El análisis del habla podría implementarse en aplicaciones como las que todos utilizamos en nuestros smartphones o dispositivos

similares. De este modo, el usuario utilizaría el micrófono para realizar cualquier tarea que implique el lenguaje hablado y la señal acústica resultante se enviaría a una nube donde se realiza el análisis computacional correspondiente, devolviendo un resultado tanto al paciente como al clínico de referencia. Esto facilitaría tanto el diagnóstico como el seguimiento de la evolución del trastorno. Ya existen apps de este tipo desarrolladas en los últimos años con diferentes fines, como por ejemplo, detectar variaciones en la voz vinculadas al estrés crónico o a determinados estados afectivos, y recomendar la búsqueda de estrategias de afrontamiento cuando sea necesario a través de la monitorización de las variaciones en el habla (Braun et al., 2016), o para la detección de la demencia (Hall et al., 2019).

Otra cuestión interesante es que, frente a la evaluación habitual basada en el autoinforme y las entrevistas clínicas -y por tanto sujeta a los sesgos tanto del paciente como del terapeuta-, se propone el análisis automático del habla como una medida objetiva (Kourtis et al., 2019), estable a lo largo del tiempo (Cohen et al., 2012), y con alta correlación con la gravedad de los síntomas (Zhang et al., 2020). Lo que significa que podría ser útil para monitorizar los cambios en la evolución de la enfermedad. De hecho, los estudios sobre los trastornos del estado de ánimo han demostrado que se puede detectar un cambio en los patrones de habla cuando comienza el problema, seguido de una vuelta a la línea base cuando el trastorno remite. Incluso pueden observarse cambios de estado dentro del mismo trastorno, como en el caso del trastorno bipolar, entre las fases maníaca y depresiva (Flanagan et al., 2021).

En este apartado, el objetivo es exponer una serie de ámbitos que demuestran la utilidad del análisis automático del habla en la detección y caracterización de diversas patologías. Nos centraremos en trastornos que producen un cambio en el habla ya sea por la afectación o el deterioro del sistema nervioso, o por una alteración de la capacidad de

procesamiento cognitivo o en la expresión de la emoción debido a trastornos mentales, del comportamiento o del neurodesarrollo, según la clasificación de la CIE-11 (OMS, 2022). No abordaremos, sin embargo, el estudio de las características del habla en determinadas patologías -síndrome de Down (Albertini et al., 2010), trastornos de la voz (Hegde et al., 2019), tiroidectomía (Lang et al., 2016)- o el habla como herramienta diagnóstica de afecciones físicas -apnea del sueño (Solé-Casals et al., 2014)-, ya que se ve afectada por motivos distintos a los mencionados, aunque la aparición de estudios de este tipo avala la utilidad del análisis acústico del habla en el ámbito clínico.

2.2.1 Análisis del habla en las enfermedades del sistema nervioso

Enfermedad de Alzheimer

La enfermedad de Alzheimer es la causa más común de demencia. Las personas que sufren esta enfermedad muestran alteraciones en el ritmo del habla, un aumento de las roturas de voz, mayor frecuencia y longitud de las pausas y una velocidad de habla más baja en comparación con grupos control formados por mayores sanos. También se pueden observar alteraciones en parámetros acústicos y de la amplitud como la Frecuencia Fundamental, los formantes y el shimmer, la ratio entre armónicos y ruido y el Promedio del espectro medio a largo plazo (LTAS, Martínez-Nicolás et al., 2021; Pulido et al., 2020). En lo que se refiere a la detección de la enfermedad de Alzheimer, podemos identificar dos corrientes principales: bien utilizar tan solo unos pocos parámetros característicos, o conjuntos amplios de parámetros seleccionados mediante técnicas de aprendizaje automático. Ambos métodos han logrado altos porcentajes de precisión a la hora de distinguir entre personas en fases tempranas de la enfermedad o

personas de grupo control, comprendidos en la mayoría de estudios entre el 80 y el 97% (Meilán et al., 2014; Nasrolahzadeh et al., 2018).

Deterioro Cognitivo Leve

El trastorno neurocognitivo leve, más comúnmente conocido como deterioro cognitivo leve, es una alteración del desempeño en uno o más dominios cognitivos, que no es lo suficientemente grave como para interferir significativamente con la independencia de la persona y las actividades de la vida diaria. El habla de las personas con DCL se caracteriza por la aparición de pausas, una menor velocidad de habla y por una alteración de la F0 y los formantes característicos de las vocales (Martínez-Nicolás et al., 2021). Los estudios de detección del deterioro cognitivo leve suelen mostrar una precisión de entre el 75 y el 80% (König et al., 2015a; Themistocleous et al., 2018). En este caso, se pueden encontrar enfoques interesantes que tratan de combinar parámetros acústicos con otras medidas lingüísticas o conductuales, consiguiendo aumentar la precisión hasta el 85% (Fraser et al., 2019; Gosztolya et al., 2019).

Enfermedad de Parkinson

La enfermedad de Parkinson es una enfermedad degenerativa progresiva caracterizada por la bradiquinesia y otros síntomas como temblor, rigidez o inestabilidad postural. Una de las manifestaciones comunes de esta enfermedad es a través de trastornos del habla (específicamente disartria hipocinética), por lo que no es sorprendente que los estudios del habla de personas con Parkinson sean clásicos (Zwirner et al., 1991). Estos estudios han hallado un habla caracterizada por un aumento la variabilidad de la frecuencia (Jitter), de la intensidad de sonido (Shimmer), de la ratio entre Ruido y

Harmónicos (NHR), de la variabilidad de la F0 (Chiaramonte & Bonfiglio, 2020a), y de forma menos frecuente, en el número de pausas (Rusz et al., 2011). Los estudios de clasificación muestran una precisión que va desde el 85 al 99% (Almeida et al., 2019; Braga et al., 2019; Solana-Lavalle & Rosas-Romero, 2021), resultado que además es consistente en diferentes lenguas (Orozco-Arroyave et al., 2016).

Enfermedad de Huntington

La enfermedad de Huntington es un trastorno neurodegenerativo que puede causar corea, demencia y cambios de la personalidad. Un meta-análisis (Chan et al., 2019) concluyó que estas personas muestran una reducción del ritmo de producción de sílabas, aumento de las roturas de voz, mayor variación del shimmer y del tono, y una reducción del HNR. Estos cambios, además, han sido detectados tanto en personas con la enfermedad, como en aquellas con su expresión genética, consideradas como pacientes en los que aún no se han manifestado los síntomas debilitantes. Hay estudios que diferencian entre personas sanas y personas con la enfermedad de Huntington con una precisión del 99%, utilizando únicamente parámetros acústicos (Guimarães et al., 2020).

Traumatismo craneoencefálico

Las lesiones por traumatismo craneoencefálico son un daño infligido a los tejidos cerebrales producido por un golpe en la cabeza. Se ha detectado una alteración de algunos parámetros del habla entre una y cuatro horas después del traumatismo. Serían medidas relacionadas con la F0 y el control motor del habla (Poellabauer et al., 2015), manifestadas en una menor velocidad de habla y mayor duración de las sílabas (Banks et al., 2021). En una de las primeras aproximaciones a este campo, Falcone et al. (2013)

podieron detectar atletas con traumatismo craneoencefálico leve con una precisión del 98% analizando únicamente sonidos de las vocales.

Otros trastornos

Aunque no son entidades nosológicas propiamente dichas, hay muchos trabajos que exploran las consecuencias de la disartria y la apraxia del habla. Estas alteraciones se suelen considerar como síntomas de ciertos trastornos, algunos de los cuales se han comentado anteriormente, como la enfermedad de Parkinson.

Tanto la esclerosis múltiple (Noffs et al., 2018) como la esclerosis lateral amiotrófica (Chiaramonte & Bonfiglio, 2020b) provocan disartria en una mayoría de los enfermos, que se manifiesta a través de la inestabilidad del shimmer y el jitter, así como una menor velocidad de habla y un mayor número de pausas. Por otro lado, los estudios sobre la disartria derivada de la parálisis cerebral muestran que las personas que la padecen presentan un tiempo de fonación significativamente mayor y un pico espectral medio (mean spectral peak) inferior al de los hablantes sanos, una medida de la producción de consonantes fricativas que se refiere a la frecuencia que divide el espectro de manera que las frecuencias de la mitad superior sean iguales a las de la mitad inferior. (Hernandez et al., 2019).

A través del análisis acústico, la apraxia del habla puede distinguirse de la afasia y de otros síntomas posteriores a un accidente cerebrovascular con una precisión del 90% teniendo en cuenta el Índice de Variabilidad Rítmica entre Pares normalizado (nPVI), la proporción de errores de distorsión, la variabilidad del inicio del tiempo de sonoridad (VOT) y el espectro de la envolvente de la amplitud modulada (*amplitude envelope modulation spectrum*), una representación de la distribución de energía en las

fluctuaciones de amplitud a través de frecuencias designadas (Basilakos, 2018; Basilakos et al., 2017). En la misma línea, la velocidad del habla y de las sílabas puede distinguir de forma fiable a las personas con apraxia progresiva primaria del habla de aquellas con afasia progresiva primaria (Duffy et al., 2017).

2.2.2. Análisis del habla en los trastornos mentales, del comportamiento y del neurodesarrollo

Depresión

Los trastornos depresivos son un grupo de trastornos del estado de ánimo caracterizados por la tristeza y la pérdida de placer, que puede ir acompañado de otros síntomas cognitivos, comportamentales o neurovegetativos. El habla en las personas con depresión se caracteriza por menor variabilidad del jitter y el shimmer, menor intensidad y pausas más frecuentes y largas, por lo que su habla se percibe como lenta y plana.

En este sentido, se pueden utilizar conjuntos de parámetros acústicos y prosódicos para caracterizar la gravedad de los síntomas depresivos e incluso la ideación suicida (Zhang et al., 2020). En cuanto a los estudios de cribado, se ha comprobado que los parámetros de amplitud pueden clasificar a las personas con depresión mayor y personas de grupo control sano con una precisión del 78% (Alghowinem et al., 2013), y con Coeficientes Cepstrales en las Frecuencias de Mel (Mel Frequency Cepstral Coefficients, MFCCs), que es una medida del análisis del espectro basado en frecuencias a una escala logarítmica, se obtiene una precisión del 81.9% (Taguchi et al., 2018). Según nuestra búsqueda en la literatura, los estudios con mayor precisión utilizan conjuntos de parámetros acústicos y prosódicos seleccionados mediante métodos de aprendizaje

automático, clasificando alrededor del 90% de los pacientes correctamente (Espinola et al., 2021; Higuchi et al., 2018).

Cabe destacar que la mayoría de las patologías que estamos tratando en este apartado no se analizan por separado según el género, sin embargo, en el caso de la depresión, se han observado algunas diferencias. En algunos de los estudios, el habla sólo sirve para predecir la gravedad de los síntomas depresivos en el caso de los hombres (Hashim et al., 2017). E incluso en aquellos estudios que consiguen detectar la depresión en ambos géneros, dicha detección es más precisa en los varones -80,30%- que en las mujeres -75,96%- (Jiang et al., 2017).

Por otro lado, hay que tener en cuenta que los estudios mencionados anteriormente se realizaron con muestras de adultos. Sin embargo, la depresión en los mayores se ha investigado por separado. Dada la alta prevalencia de este trastorno en la población mayor de 65 años y la interacción que tiene con la demencia, es de esperar que el problema se aborde de manera específica en esta población. A este respecto, Smith et al. (2020) consiguieron diferenciar a los mayores con depresión moderada o severa de los que mostraban síntomas leves o ausencia de síntomas, con un 92% de precisión. De igual forma que en la población general, estos estudios se han realizado dividiendo la muestra por género, encontrando así un mejor rendimiento de detección utilizando parámetros espectrales y relacionadas con la amplitud para los hombres -86%-, y características relacionadas con la prosodia para las mujeres -77%-. (Lee et al., 2021).

Trastorno bipolar

El trastorno bipolar es un trastorno episódico en el que se alternan episodios o síntomas maníacos, mixtos o hipomaníacos, con otros en los que se muestran síntomas

depresivos. En las fases de manía, el habla tiende a mostrar un tono alto, mientras que, en las fases depresivas, el tono es más bajo. Por otro lado la fase de eutimia o normalidad, se caracteriza por menor variabilidad del jitter (Vanello et al., 2012). Mediante análisis del habla se puede distinguir entre las fases de eutimia y depresiva con un área bajo la curva de 0,78, y entre las fases maníacas o mixtas y la de eutimia con una de 0,89 (Faurholt-Jepsen et al., 2016).

Por tanto, en el trastorno bipolar, el esfuerzo se ha puesto en el seguimiento de los cambios entre episodios, en lugar de limitarse a diagnosticar el trastorno. Este enfoque ha permitido el desarrollo de una variedad de medidas conductuales ecológicas. En este sentido, es común el uso de smartphones para recoger datos, no sólo del habla. Así, Grünerbl et al. (2014) combinaron los parámetros de habla con otros indicadores comportamentales que incluían la interacción social, el movimiento físico y los patrones de desplazamiento, logrando una precisión del 97% en la detección de cambios entre fases.

Esquizofrenia

La esquizofrenia se caracteriza por distorsiones del pensamiento, la percepción, cognición, intención, afecto y comportamiento. Desde el inicio del estudio de este trastorno se ha descrito un patrón de habla atípico. Se han encontrado alteraciones en el tono medio, la variabilidad del tono, la variabilidad de la intensidad, la velocidad del habla y el número y duración de las pausas. Sin embargo, un meta-análisis reciente de Parola et al. (2020) revela que solo la duración de las pausas muestra evidencias consistentes con un tamaño del efecto robusto. Este mismo meta-análisis encontró cinco estudios que intentaban clasificar a personas con esquizofrenia y controles sanos, con

precisiones que oscilaban entre el 76,5% y el 87,5%. Diferentes enfoques indican que, si bien los parámetros temporales del habla sirven para diferenciar entre pacientes y controles (Martínez-Sánchez et al., 2015), los acústicos podrían utilizarse para distinguir entre pacientes con síntomas predominantemente positivos de aquellos que muestran principalmente síntomas negativos. La composición de los formantes (Covington et al., 2012), la amplitud del sonido y diferentes parámetros de calidad de voz se relacionarían con síntomas negativos, mientras que el porcentaje y número de los segmentos del habla con voz se relacionarían con los positivos (Boer et al., 2021).

Trastornos del espectro autista

El autismo es un trastorno del neurodesarrollo caracterizado por déficits en la interacción social y la comunicación, y por patrones de comportamiento restringidos, repetitivos e inflexibles. Dado que una de sus alteraciones características tiene que ver con el lenguaje, son ya clásicos los estudios sobre la prosodia en el autismo. Sin embargo, como muestran Fusaroli et al. (2017) en un meta-análisis, no es hasta 2009 cuando podemos observar un verdadero auge del diagnóstico a través del habla. Los estudios abordados en ese meta-análisis que buscan parámetros distintivos del habla encuentran algunas diferencias notables en las medidas relacionadas con el tono, pero no tan a favor de las medidas relacionadas con la intensidad o la calidad de la voz. Los estudios que pretenden clasificar el síndrome mediante aprendizaje automático obtienen precisiones de entre el 70 y el 95%. Cabe destacar que, aunque la muestra suele estar formada por niños de diferentes edades, hay estudios que consiguen diagnósticos precisos con niños pequeños de entre 18 y 36 meses analizando únicamente sus vocalizaciones pre-verbales (Santos et al., 2013). Al igual que en otros campos, los estudios sobre el autismo han intentado integrar el análisis acústico y el lingüístico, como es el caso de Gong et al.

(2018), en el que utilizaron una *app* para recoger el habla de los niños tanto en tareas estructuradas como en la comunicación cotidiana. Aunque combinaron medidas lingüísticas y acústicas, encontraron que estas últimas eran las que más contribuían a la clasificación, obteniendo una precisión del 88,9%.

Otros trastornos

Podemos encontrar trabajos sobre otros trastornos, que, aunque menos estudiados, han arrojado resultados prometedores, como son la ansiedad generalizada, el trastorno obsesivo-compulsivo, el trastorno de estrés postraumático, la anorexia y la bulimia (Low et al., 2020). Algunos estudios muestran que las personas con ansiedad generalizada producen habla con una F0 más alta, mayores valores en jitter y shimmer, una velocidad de habla mayor y con menos pausas (Özseven et al., 2018). Sin embargo, hasta donde sabemos, no hay estudios que intenten detectar este trastorno a través del habla. Del mismo modo, en el trastorno obsesivo-compulsivo se observa un aumento del jitter y de la velocidad de habla (Cassol et al., 2010), pero tampoco hemos encontrado estudios de clasificación. Asimismo, mediante el uso de 18 marcadores que indican un habla lenta y monótona, es posible diferenciar entre veteranos de guerra que han desarrollado un trastorno de estrés postraumático de otros que no, con un 89% de precisión (Marmar et al., 2019).

Cabe destacar que, en el caso de la anorexia y la bulimia nerviosa, si bien hay estudios que encuentran cambios en el habla de las personas que padecen estos trastornos, estos cambios no se deberían a una alteración de procesos cognitivos subyacentes, sino a una serie de consecuencias derivadas del propio trastorno. Esto es, por un lado, las alteraciones de la laringe producidas por los vómitos, en el caso de la bulimia. Y por otro,

si la anorexia comienza a una edad temprana, podría producir alteraciones en el desarrollo de la laringe durante la pubertad (Low et al., 2020).

2.2.3 Críticas y perspectiva de futuro

Como se ha visto, hay pruebas sólidas de que el análisis automático del habla puede utilizarse para detectar con éxito diversos trastornos mentales. Los prometedores resultados en algunas de estas patologías han llevado a los investigadores a probar su uso en muchas otras, siendo un campo en desarrollo constante y en el que se buscan nuevas aplicaciones potenciales. Como muestra, en los apartados anteriores se han incluido estudios que encuentran alteraciones características en el habla, aunque todavía no se hayan dado los pasos necesarios para permitir la detección de esos trastornos. Pero más allá de esto, incluso aquellos en los que las pruebas son sólidas y se propone el análisis del habla como una herramienta de *screening*, la transición del laboratorio a la clínica es más problemática. Aventuramos que esto cambiará pronto y que en poco tiempo aparecerán validaciones clínicas que avalen su utilidad práctica, ya que su aplicación como herramienta de diagnóstico y monitorización está ya ampliamente respaldada en el contexto de la investigación. Junto a ello, hay que señalar que este método tiene una gran aceptación entre los potenciales usuarios debido a su sencillez (Arevian et al., 2020), lo que nos lleva de nuevo a la idea de que los dispositivos electrónicos pueden ser y serán grandes aliados en la atención a la salud mental.

En este punto debemos hacer notar una serie de limitaciones que son comunes en muchos de los estudios independientemente del trastorno, y que de hecho, han sido criticados en varias revisiones (Fusaroli et al., 2017; Martínez-Nicolás et al., 2021; Parola et al., 2020).

En primer lugar, muchos de los estudios utilizan muestras pequeñas, y a menudo, heterogéneas. En segundo lugar, los métodos de recogida de datos y de extracción de parámetros del habla son muy variables, lo que dificulta la comparación entre ellos. Además, pocos estudios se rigen por un enfoque teórico, sino que siguen un enfoque observacional y exploratorio en el que se alteran los rasgos del habla, sin dar información sobre las razones de esa alteración. Por último, en la misma línea, la mayoría de los estudios utilizan inteligencia artificial u otros métodos de selección automática de parámetros. Esto, aunque no es un problema en sí mismo, puede convertirse en uno cuando no se especifican los parámetros que se seleccionan, sino que los autores se refieren a conjuntos opacos de variables y se limitan a informar de un resultado de precisión de la clasificación. Esta práctica dificulta enormemente la replicación y el consenso y, por tanto, el progreso de todo el campo.

Uno de los retos fundamentales a partir de ahora es considerar la inmensa complejidad y variabilidad de las personas con trastornos mentales y neurológicos. La mayoría de los estudios abordan la detección de patologías puras, pero la comorbilidad es habitual en la práctica clínica. Aunque algunos estudios recientes están empezando a seguir este enfoque, tratando de detectar diferentes síntomas neuropsiquiátricos dentro de los pacientes con deterioro cognitivo leve (König et al., 2021), la realidad es que todavía estamos en las primeras etapas de estos nuevos enfoques. Por tanto, si queremos que esta herramienta sea de verdadera utilidad en la práctica clínica, será necesario integrar en el mismo análisis la posibilidad de detectar diferentes patologías, sentando las bases para un diagnóstico diferencial.

En resumen, en este apartado hemos realizado una breve revisión de la literatura sobre la detección de trastornos mentales mediante el análisis automático del habla. En la última década, el campo ha evolucionado desde los primeros enfoques exploratorios, en

busca de parámetros del habla alterados, hasta el uso actual de éstos para posibles diagnósticos. Esta herramienta puede detectar una gran variedad de trastornos, incluidos los del sistema nervioso, así como los mentales, del estado de ánimo y del comportamiento. Se trata de un campo en rápida evolución que sigue abriendo posibles usos. Los retos que tenemos por delante se basan principalmente en convertir esta herramienta en un aliado para la atención a la salud mental, utilizando nuestras tecnologías cotidianas, y teniendo en cuenta las comorbilidades que pueda presentar cada paciente.

2.3 El lenguaje en la Demencia tipo Alzheimer

En esta sección se expondrán los déficits del lenguaje característicos de la demencia tipo Alzheimer, dado que serán estas alteraciones las que, en último término, provocan los cambios en el habla de estos pacientes. Comenzaremos describiendo el curso general y posteriormente se pondrá especial atención a los cambios sutiles asociados a las primeras etapas ya que son las de mayor interés para la detección del trastorno.

Si bien el déficit de memoria episódica ha sido siempre considerado el síntoma típico de la EA, esta enfermedad provoca alteraciones en otros dominios, destacando la función ejecutiva y el lenguaje. En este sentido, mientras en el proceso de envejecimiento normalizado parece haber una relativa preservación de las aptitudes lingüísticas, ya en 1907 Aloisius Alzheimer (Alzheimer et al., 1995) estableció que mientras los trastornos de memoria en los individuos con EA difieren cuantitativamente en su ejecución, lo que dificulta establecer puntos de corte, los deterioros en los rasgos del lenguaje muestran diferencias cualitativas y caracterizan tempranamente la enfermedad. Además, los cambios cualitativos en el lenguaje permiten comprender mejor las fases de la enfermedad.

La importancia del déficit del lenguaje en la EA es sobradamente reconocida y se encuentra entre sus criterios diagnósticos (McKhann et al., 2011). Esto ha llevado a describir el cuadro de la afasia progresiva primaria (APP) en su variante logopénica, caracterizado por déficits en la recuperación de palabras, y considerado como habitualmente causado por la EA. Siendo la APP uno de los síndromes asociados y característicos de la EA (Dubois et al., 2014). Aunque otros estudios, sin embargo, consideran que los déficits del lenguaje característicos de la EA difieren cuantitativa y cualitativamente de los de la APP logopénica (Ahmed et al., 2012; Blair et al., 2007; Taler

& Phillips, 2008). La cuestión de fondo es la dificultad en consensuar cuáles son los síntomas del lenguaje asociados a la EA, una enfermedad caracterizada por una gran heterogeneidad. Y, concretamente, qué niveles del lenguaje se ven afectados y cuales se conservan.

Se ha considerado normalmente que el léxico y la pragmática son los dominios más afectados, mientras que los niveles estructurales como la sintaxis, la morfología y la fonética, se conservan hasta etapas más avanzadas de la enfermedad. Sin embargo, recientes investigaciones ponen también esto en duda, y se sugiere que los niveles léxico y pragmático son simplemente más fácilmente detectables (Ivanova et al., 2017). Esto se debería a que clásicamente los estudios sobre EA se han centrado en el nivel de la palabra, dejando de lado tanto sus componentes como su uso en el contexto de la comunicación (Rochon et al., 2018).

En las fases más tempranas de la EA el déficit por excelencia es la anomia. Aparecen dificultades para realizar tareas estructuradas que requieren denominación o producción de palabras. En el lenguaje espontáneo, es más frecuente el uso de palabras baúl, vacías semánticamente, a modo de estrategia de compensación para mantener la fluidez sacrificando el contenido (Kempler, 1995). La comprensión está relativamente preservada, excepto para el lenguaje abstracto. En esta fase, la habilidad comunicativa del enfermo sería lo suficientemente buena como para rendir adecuadamente en la mayoría de las situaciones sociales, aunque tendría problemas para seguir conversaciones complejas. Cuando el deterioro pasa de leve a moderado, se hacen más evidentes las dificultades tanto en la producción como en la comprensión. La afasia anómica empeora y se hacen comunes las parafasias semánticas y los circunloquios. El discurso en esta etapa es pobre de contenido, con dificultad para mantener un tema concreto, pero la sintaxis producida es relativamente correcta. Socialmente, el paciente ya muestra graves

dificultades para atender a las demandas comunicativas y se produce un retraimiento. En fases más avanzadas, cuando el deterioro es grave, la producción es totalmente vacía, incoherente e ininteligible debido a la disartria. Se caracterizaría por manifestar ecolalia, palilalia o mutismo. La comprensión es nula y el enfermo es incapaz de comunicarse de ningún modo.

Nos centraremos ahora en cada uno de los dominios indicados anteriormente: acceso léxico-semántico, pragmática, sintaxis, morfología y fonética. En las fases más tempranas, el primer signo de un déficit del lenguaje es la alteración del acceso léxico-semántico. Este proceso implica la búsqueda de un concepto en la mente, la activación de posibles candidatos fonológicos y la selección del apropiado. Este déficit se hace evidente en tareas como las de decisión léxica en las que discriminar entre palabras reales y no reales. Pruebas en las que estos pacientes muestran mayores tiempos de respuesta (Szatloczki et al., 2015); en tareas de denominación, en las que ante la imposibilidad de emitir la respuesta adecuada los pacientes producen parafasias semánticas (Silagi et al., 2015); y en tareas de fluidez verbal. Las tareas de fluidez verbal consisten en pedir al participante que produzca el mayor número de palabras posible en un tiempo determinado utilizando una pista, que puede ser una categoría semántica en el caso de la fluidez verbal semántica, o un fonema, en el caso de la fluidez verbal fonológica. Estas tareas tienen una implicación de la función ejecutiva, dado que el participante debe buscar y organizar las respuestas adecuadas, mantener las anteriores para evitar la repetición e inhibir las respuestas inapropiadas. Además, ambos tipos de tarea difieren en los procesos implicados, mientras que la fonológica se basa en representaciones léxicas, la semántica lo hace en asociaciones de significado con un superordenado. En la demencia tipo Alzheimer se puede observar un deterioro diferencial en estas tareas, de modo que la semántica muestra mayor alteración que la fonológica (Gordon & Chen, 2022; Hodges

et al., 1996), lo que sugiere una disrupción de la memoria semántica. Más allá de los resultados en pruebas neuropsicológicas, el deterioro del acceso léxico-semántico tiene consecuencias en el uso habitual del lenguaje. Como resultado, las personas con EA producen parafasias semánticas, crean neologismos y producen un discurso vago, incoherente e impreciso (Ferris & Farlow, 2013). Además, en conversación tienden a producir respuestas más cortas y con menos unidades de información que los mayores sanos, además de producir más errores de tipo semántico (Croisile et al., 1996). Muestran también dificultades para responder a preguntas abiertas y aportar información nueva, debido principalmente a las deficiencias de comprensión y de acceso léxico-semántico (Rousseaux et al., 2010).

Una cuestión polémica es la de explicar la causa de estos déficits. Se han propuesto dos teorías principales. La primera achacaría las alteraciones a un fallo en la organización y estructura de las redes semánticas. Así, la degradación de las áreas cerebrales asociadas al almacenamiento y la representación de los conceptos y sus atributos provocaría una pérdida real de ese conocimiento. Por otro lado, la otra teoría defiende que la causa es un fallo en la recuperación de la información de una red semántica conservada debido a las deficiencias en los procesos ejecutivos implicados en el acceso a dichas representaciones (Rogers & Friedman, 2008). Podemos encontrar evidencia a favor y en contra de cada una de las teorías. El distinto rendimiento de los pacientes con EA en las tareas de fluidez fonológica y semántica apoya la idea de que pueden estar experimentando una pérdida en la organización de la memoria semántica más que sufrir dificultades en su recuperación (Haugrud et al., 2011; Henry et al., 2004). Además, en tareas de denominación, existe una relación entre la capacidad de un paciente con EA para nombrar un objeto y su conocimiento de ese mismo objeto. Cuando deben definir un objeto a cuyo nombre no pueden acceder, sus descripciones se empobrecen, proporcionando menos atributos e

incluso perdiendo características fundamentales (Garrard et al., 2005; Hodges et al., 1996). Por otro lado, la teoría que achaca el déficit a la función ejecutiva está respaldada por estudios basados en el paradigma del priming semántico. Esta postura se basa en la idea de que la difusión de la activación en la red semántica requiere conexiones intactas dentro del sistema. Si la red desaparece, no debería ser posible primar la palabra meta que hay que encontrar con una previa relacionada. Diferentes estudios encuentran consistentemente un priming semántico intacto en pacientes con EA (Bell et al., 2001; Chertkow et al., 1994; Nakamura et al., 2000), lo que apoya que el problema se debería al deterioro de los mecanismos de recuperación. Una forma de resolver la cuestión sería demostrar la pérdida completa de un concepto, lo que implicaría que la persona no puede acceder a la palabra de forma consistente en diferentes tareas y momentos. Hodges et al., (1992) mostraron que cuando los pacientes no son capaces de utilizar un ítem en una tarea, es probable que tampoco puedan hacerlo en otras tareas que utilicen ese mismo ítem como objetivo, lo que apoyaría la degradación del almacenamiento. No obstante, no se puede ignorar la evidencia mencionada de una red semántica intacta. Por lo tanto, parece probable que ambas deficiencias se produzcan y tengan consecuencias en el rendimiento de los pacientes con EA. Una aproximación conciliadora que se ha propuesto afirma que, en etapas tempranas, el déficit de recuperación podría ser el causante de las dificultades a las que luego se suma la degradación (Salehi et al., 2017) provocando, a largo plazo, una anomia total.

En cuanto al nivel de la pragmática, este dominio va más allá del significado de las palabras e incluye la comprensión de información contextual como puede ser la interpretación de la intención del hablante y de la situación comunicativa. Las alteraciones de los pacientes de EA en las fases más tempranas incluyen dificultad para entender refranes o el uso de formas abstractas de lenguaje. A medida que avanza el curso de la

enfermedad desarrollarán dificultades para iniciar la comunicación y respetar los turnos de conversación, seguir la conversación de forma coherente, corregir sus fallos lingüísticos o utilizar el lenguaje de forma adecuada socialmente mediante el uso de fórmulas de cortesía (Sakin, 2021). Estos déficits se ven fuertemente influidos por las alteraciones de la semántica, la función ejecutiva y la memoria episódica, aunque se ha comprobado que los pacientes de EA muestran dificultades para entender los modismos y las peticiones indirectas incluso en ausencia de déficits semánticos evidentes (Luzzi et al., 2020).

En cuanto a los niveles preservados en las primeras etapas, la sintaxis está relativamente conservada y es correcta y coherente (Mueller et al., 2016). Sin embargo, empeora durante el curso de la enfermedad, produciendo enunciados sintácticamente más simples, más cortos, con menos oraciones subordinadas y evitando el uso de construcciones complejas como las oraciones pasivas (Kemper et al., 1993). Por su parte, en el nivel morfológico, las personas con EA no muestran diferencias con respecto a otros mayores sin la enfermedad (Nikolaev et al., 2019). En cuanto al nivel fonético y fonológico, la opinión dominante es que tanto el procesamiento fonológico como la producción fonética están conservados durante las primeras fases de la enfermedad y efectivamente muchos estudios reportan que no hay deterioro (Lira et al., 2011). Sin embargo, esta posición es cada vez más cuestionada. Ya se consideran clásicos los estudios que demuestran la presencia de déficits en el acceso a representaciones fonológicas y que, de hecho, predijeron las alteraciones en los sistemas motores de articulación (Croot et al., 2000). En los afectados por la EA se ha descrito el alargamiento de sílabas, vocales y consonantes, sustituciones y adiciones de fonemas, y transposiciones; alteraciones que no se observan entre los mayores sanos (Cera et al., 2018). En lectura oral, se ha indicado que los errores fonológicos de sustitución, omisión

o adición de fonemas son, de hecho, comunes ante palabras de baja frecuencia (Del Carmen Pérez-Sánchez et al., 2021), sin embargo, también se señala que serían menos graves en comparación a la afectación del acceso al léxico al principio de la enfermedad y que empeoran con su avance (Cuetos et al., 2005). A nivel de producción fonética, también hay estudios recientes que demuestran sutiles alteraciones tempranas en la producción motora de fonemas (Nasrolahzadeh, Mohammadpoori, et al., 2016; Themistocleous et al., 2018).

2.4 El habla en el deterioro cognitivo leve y la enfermedad de Alzheimer: características y detección

Como hemos visto, el análisis del habla ha adquirido en los últimos tiempos un respaldo extenso en cuanto a su utilidad en el cribado de diversas patologías y estados. Nos centraremos en este apartado en el fundamento para aplicar esta técnica al estudio del Deterioro Cognitivo Leve (DCL) y la Enfermedad de Alzheimer (EA). En esta sección se mostrarán los resultados de una revisión sistemática sobre la caracterización del habla de las personas con DCL y EA y los estudios que se han realizado acerca de su detección y propuesta como herramienta de cribado.

Distinguir las primeras fases de la EA del deterioro cognitivo asociado al envejecimiento normal es un reto, ya que una parte importante de los pacientes con EA son asintomáticos durante las fases preclínicas del proceso patológico, que se cree que dura aproximadamente 17 años (Jansen et al., 2015; Villemagne et al., 2013) hasta que compromete la cognición de la persona. Sin embargo, durante la larga fase preclínica asintomática -y los progresivos cambios patológicos subyacentes (Jack et al., 2010)-, los parámetros del habla asociados a las funciones cognitivas pueden anticipar las manifestaciones clínicas de la demencia y por tanto pueden ser útiles en el diagnóstico precoz de la EA y en el desarrollo y evaluación de estrategias preventivas y terapéuticas. Los estudios sobre el lenguaje en la demencia han tendido a centrarse en lo que el paciente dice, más que en cómo lo dice, ignorando así los cambios en los procesos automáticos del lenguaje -por ejemplo, el desarrollo del habla- a lo largo de la enfermedad.

En el estudio del habla cuando se trata de población mayor, es importante tener en cuenta que no sólo se verá alterado por trastornos y cambios cognitivos, sino que también varía a lo largo de la vida debido a los procesos normales de envejecimiento. Por

ejemplo, los primeros estudios sobre la eficacia del control articulatorio durante los movimientos del habla en personas mayores concluyeron que su rendimiento es peor que el de los jóvenes. Una razón que se aduce es que la precisión de la amplitud de los movimientos tiende a disminuir durante el envejecimiento, lo que repercute en la exactitud y los parámetros temporales del habla (Ballard et al., 2001). Estudios más recientes se han centrado en un conjunto característico de rasgos clínicos que aparecen en las voces de las personas mayores, conocido como presbifonía. Se trata de una alteración de la voz, debida al proceso de envejecimiento, causada por los cambios anatómicos y fisiológicos en la laringe y el tracto vocal de las personas mayores, y por las dificultades que surgen en el control de los parámetros acústicos (Alonso et al., 2001), y otros factores mecánicos, estructurales y hormonales (Bruzzi et al., 2017). Estos cambios explican la disfonía característica entre las personas mayores que produce una reducción del rango vocal de frecuencias utilizado, una disminución de la frecuencia fundamental (F0) en las voces femeninas (desde niveles normales en torno a 248 Hz hasta 175 Hz), y un aumento de la F0 en las voces masculinas (desde 110 Hz hasta 135-160 Hz). Además, aparece una mayor variabilidad en el jitter y shimmer, se reduce la resonancia del sonido vocálico y hay más pausas en el habla (Linville, 2004).

En las personas con enfermedad neurodegenerativa, la prosodia es uno de los parámetros más estudiados, dado que suelen estar alterados diversos procesos cognitivos que se reflejarían en la señal acústica. Estaríamos hablando de elementos como el ritmo del habla, es decir, las pausas, los acentos o la velocidad del habla, junto con otros parámetros relacionados con las medidas temporales y acústicas de la voz, como el ritmo articulatorio, la intensidad, el tiempo de emisión y las frecuencias (variaciones en las frecuencias de la señal, el timbre o la estructura de los formantes).

En la historia del análisis del lenguaje en las personas con EA, son clásicos los estudios de principios de siglo que extraían parámetros prosódicos de forma manual. Singh et al., (2001) utilizaron una tarea de lectura oral con la que concluyeron que la velocidad de habla, la duración media de las pausas y el tiempo de fonación discriminan con precisión entre las personas mayores sanas y las que padecen EA. Estudios de esta naturaleza sentaron las bases posteriores del análisis automático del habla. En 2006, Salhi and Cherif desarrollaron un nuevo software de procesamiento del habla y lo utilizaron para analizar audios grabados por personas con diferentes patologías. Uno de ellos pertenecía a un paciente con EA cuya habla estaba distorsionada en los formantes F2 y F3. Este hallazgo abrió el camino a varios estudios sobre el análisis automático del habla en la EA. Desde ese primer análisis, los estudios que exploran las consecuencias conductuales en la ejecución vocal debidas a cambios sutiles en los procesos del lenguaje han ido ganando popularidad. Desde 2010, cuando se publicó el primer estudio que abordaba directamente el tema (Hoffmann et al., 2010), se ha producido un número creciente de estudios que tratan de identificar los cambios en el habla de los mayores afectados por la EA y, posteriormente, el abordaje de personas con DCL. La mayoría de estos estudios se centran en la identificación de los parámetros característicos del habla en este grupo de población, o los utilizan para discriminar entre los mayores sanos y los afectados por la enfermedad. Sin embargo, los resultados son heterogéneos debido a la variedad de métodos utilizados, a que influyen otras características físicas y sociales y a las diferentes características del habla. Los programas de análisis de habla extraen decenas de parámetros, muchos de ellos interrelacionados, pero cuesta obtener una imagen clara de los aspectos del habla que caracterizan uno u otro grupo de población. Aunque los parámetros del habla en los que se encuentran alteraciones se atribuyen a cambios cognitivos, no está claro cuáles son ni cómo cambian. También hay una amplia

gama de tareas diferentes utilizadas para elicitarse el lenguaje oral y, en este mismo sentido, esto podría afectar al resultado obtenido, ya que las características del habla podrían depender también de los procesos específicos implicados en la tarea de su producción.

Para profundizar en este campo haremos una revisión sistemática de los trabajos realizados en la última década y analizaremos críticamente la calidad y el riesgo de sesgo de la literatura marcando como inicio el anteriormente mencionado primer artículo de Hoffman. Abordaremos qué parámetros son los relacionados en la EA y el DCL y si son útiles en su diagnóstico, qué tareas se han utilizado más y cómo han evolucionado los abordajes del tema. Para ello proponemos las siguientes preguntas de investigación:

1. ¿Qué rasgos caracterizan el habla de las personas con EA y DCL?
2. ¿Es el análisis automático del habla un método fiable para evaluar la EA y el DCL?
3. ¿Cuál es la tarea más útil en estos dos objetivos para elicitarse el lenguaje oral y analizarlo?

Criterios de elegibilidad

Se han incluido únicamente artículos de revistas con revisión por pares que aplicaran técnicas de análisis de la señal acústica del habla a las voces de personas con EA o DCL. Los artículos debían incluir al menos un grupo de EA o un grupo de DCL, así como un grupo de control sano (CS) formado por personas mayores. También se incluyeron artículos que exploraban otros trastornos neurodegenerativos pero que proporcionaban resultados para estos dos grupos. Debían incluir la grabación de enunciados hablados y su estudio mediante técnicas de análisis automático del habla. Se excluyeron los trabajos que sólo contenían parámetros del habla extraídos manualmente.

Sólo se incluyeron los estudios cuyo objetivo era explorar las características del habla o diagnosticar la EA o el DCL. Los resultados consisten en parámetros descriptivos del habla alterados en personas con DCL o EA con respecto a una población sana de edad avanzada, o en la precisión del diagnóstico para distinguir entre estos grupos.

Los criterios de exclusión fueron los siguientes 1) estudios que no utilizaran técnicas de análisis automático de la señal del habla; 2) estudios sin un grupo de EA y DCL; 3) estudios en un idioma distinto del inglés o el español; 4) estudios no publicados en una revista revisada por pares; 5) estudios que fueran revisiones narrativas o sistemáticas; y 6) estudios de casos.

Método de localización e identificación de estudios

La búsqueda se realizó en las bases de datos PubMed, CINAHL y PsychINFO. La última búsqueda se realizó el 28 de abril de 2022. Se incluyeron los mismos términos en todas las bases de datos (*speech features OR acoustic features OR voice OR speech analysis OR acoustic analysis OR speech signal OR spoken language OR speech production OR spontaneous speech OR connected speech OR speech acoustics OR automatic spontaneous speech analysis OR ASSA*) AND (*Alzheimer OR Mild Cognitive Impairment*).

Se recopilaron los resultados y se eliminaron los duplicados. A continuación, dos revisores (el autor y un compañero del GIR en Neurofisiología, Cognición y Conducta) revisamos de forma independiente los títulos y resúmenes de los estudios obtenidos. Esta fase requirió la lectura del texto completo varias veces, siendo la razón principal que el método de análisis del habla no estaba claramente descrito en el resumen. La concordancia entre los revisores fue de 0,992, $\kappa = 0,993$. Los desacuerdos entre los

revisores se resolvieron mediante discusión. Se eliminaron los artículos que no cumplían los criterios de inclusión. Para realizar esta revisión se siguió la declaración PRISMA (Moher et al., 2009).

Evaluación de la calidad y extracción de datos

La calidad metodológica y el riesgo de sesgo de los estudios seleccionados se evaluaron mediante dos listas de verificación: La lista de comprobación JBI (Moola et al., 2017) para los estudios observacionales transversales, y la herramienta QUADAS-2 (Whiting et al., 2011) para la evaluación de la calidad de los estudios de precisión diagnóstica.

En todos los artículos se extrajo el tamaño de la muestra y el método para elicitar el habla. Uno de los objetivos de esta revisión es identificar los rasgos relevantes que caracterizan las voces de las personas con EA o DCL. Por lo tanto, se buscaron diferencias significativas en los valores de los parámetros rítmicos y acústicos entre los controles sanos y esos grupos. Cuando los estudios pretendían clasificar a los participantes en EA o DCL según la información proporcionada por su habla, se seleccionaron los valores de precisión siempre que fue posible para comparar la eficacia de los diferentes métodos. Si no se proporcionaron valores de precisión, se seleccionó el valor más relevante. Cuando se proporcionó más de un valor de precisión, se extrajo el más bajo y el más alto.

Estudios extraídos

El proceso de búsqueda se ha resumido en la Figura 10 mediante un diagrama de flujo PRISMA. Tras la búsqueda se obtuvieron un total de 2296 estudios; se eliminaron

385 duplicados. Tras el cribado por título y resumen, se seleccionaron 94 estudios para una revisión de texto completo, y 36 de ellos se eliminaron tras aplicar los criterios de exclusión. Por tanto, finalmente se incluyeron un total de 57 artículos. La razón más común para la exclusión fue la falta de un análisis automático de la señal del habla, ya que muchos estudios se centraron en el análisis del discurso y del contenido. Otro motivo importante fue el uso de esta técnica para patologías que no fueran EA o DCL, como el Parkinson o la esquizofrenia.

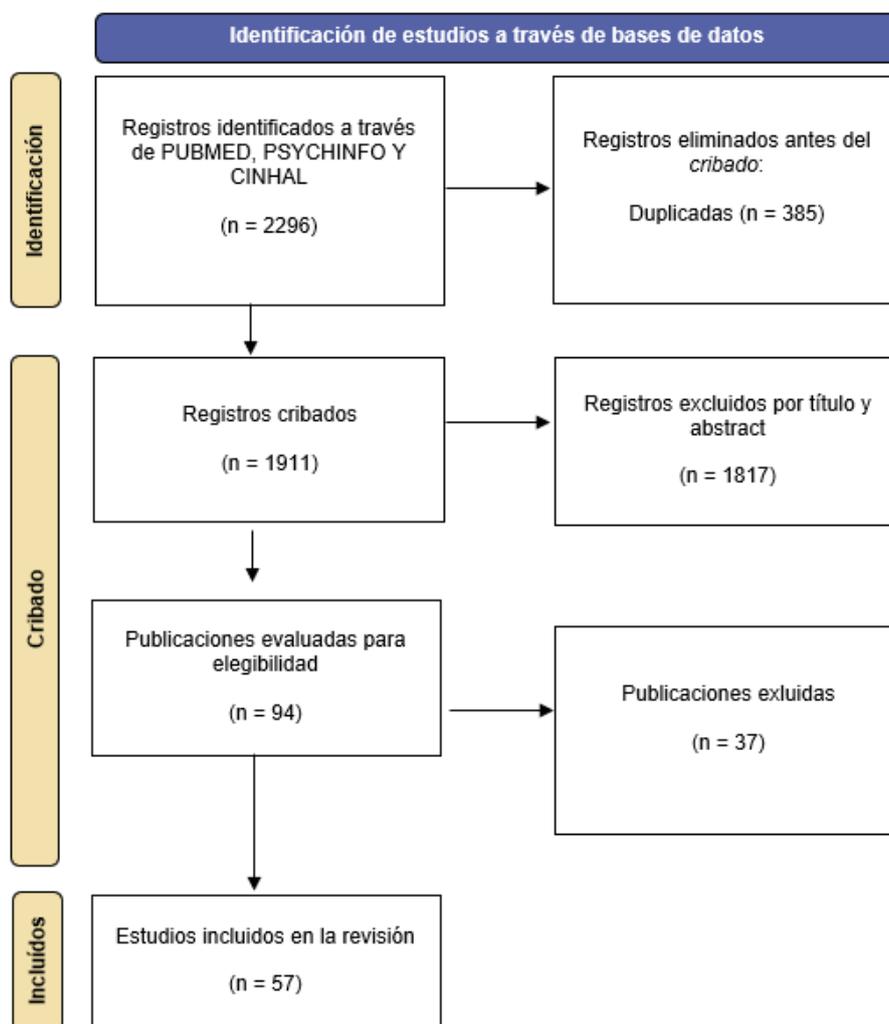


Figura 10. Diagrama de flujo PRISMA del proceso seguido para seleccionar los artículos incluidos en la revisión.

Entre los estudios seleccionados, uno examinó únicamente pacientes con EA, otro examinó únicamente DCL, 27 compararon grupos sanos de control y EA, ocho compararon grupos de control y DCL, y 20 compararon grupos de control, DCL y EA.

Los métodos utilizados para elicitación del habla son diversos, y muchos estudios utilizan más de uno. Los más comunes son las conversaciones estructuradas, las tareas de lectura y las tareas estandarizadas, como la descripción de imágenes (especialmente la lámina "*Cookie Theft*" del test de Boston) o las tareas de fluidez verbal. Otras tareas que se utilizan ocasionalmente son recordar acontecimientos o vídeos y hablar sobre la vida cotidiana. En las Tablas 1 y 2 se puede encontrar información más detallada sobre el tamaño de las muestras y el método de obtención de las señales del habla.

Se identificaron dos tendencias principales en los estudios seleccionados: 17 de ellos comparaban las voces de los grupos para encontrar rasgos característicos, describirlos y explicar las diferencias; 41 trataban de clasificar la muestra en su grupo correspondiente utilizando los parámetros del habla; es decir, desarrollando una herramienta para el diagnóstico del DCL y la EA. Para simplificar la información, se han organizado los estudios en dos tablas según sus objetivos. La tabla 1 contiene los estudios descriptivos, y la tabla 2, los prescriptivos.

Tabla 1. Estudios descriptivos del habla de las personas con DCL o EA.

ARTICLE	SAMPLE	TASK	MAIN FINDING
(Hoffmann et al., 2010)	CS (15) EA (30)	Lenguaje espontáneo: explicar por qué están en la clínica, recuerdo de eventos y actividades de la vida diaria	Mayor tiempo de habla y de fonación; más roturas de voz / pausas >30 ms; menor velocidad de habla y de articulación
(Horley et al., 2010)	CS (20) EA (20)	Repetición de frases con entonación emocional y lectura	Diferencias en F0 (solo al expresar sorpresa o felicidad) F0 dt y velocidad de habla (en lectura, pero no en repetición)
(Martínez-Sánchez et al., 2012)	CS (17) EA (25)	Lectura oral	F0, F0 dt, tiempo de fonación, proporción de pausas, % de segmentos sin voz, roturas de voz
(Nasrolahzadeh, Mohammadpoory, et al., 2016)	CS (30) EA (30)	Contar historias personales y conversación	Personas con EA muestran menos variaciones en la señal acústica del habla
(Nasrolahzadeh, Mohammadpoori, et al., 2016)	CS (30) EA (30)	Contar historias personales y conversación	Las señales del habla espontánea de los pacientes con EA son menos caóticas y no lineales que las de los sujetos sanos usando análisis espectral de orden superior.
(Beltrami et al., 2018)	CS (48) Deterioro cognitivo (48: 16 aDCL, 16 mdDCL y 16 tDem)	Describir una imagen compleja, un día típico de trabajo y recordar el último sueño	Los parámetros acústicos y rítmicos pueden diferenciar entre DCL multidominio, Demencia temprana y Control. Algunos parámetros acústicos discriminan entre Control y aDCL. Diferencias en el centroide espectral; tasa de pausa estandarizada; tasa de fonación transformada; duración de los segmentos de habla; duración de los segmentos sin voz.

(Meilán et al., 2012)	EA (21)	Lectura oral	El porcentaje de segmentos sin voz explica una parte significativa de la varianza en las puntuaciones obtenidas en pruebas neuropsicológicas de los pacientes con EA
(Meilan et al., 2018)	CS (102), DCL (38) EA (42)	Lectura oral	Las tareas de fluidez verbal semántica y fonética explican el 30,1% de la varianza del porcentaje de segmentos sin voz y el 26,4 del porcentaje de roturas de voz.
(De Looze et al., 2018)	CS (36) DCL (16) EA (18)	Lectura de frases de diferentes longitudes y complejidad sintáctica.	Cambios en la segmentación y el tiempo del habla al leer oraciones con alta demanda cognitiva pueden ser marcadores del DCL y la EA como consecuencia de los déficits en la memoria de trabajo y la atención.
(Qiao et al., 2020)	CS (24) DCL (20) EA (20)	Lámina del robo de galletas del Boston	Siete parámetros se correlacionan con la función cognitiva. Una regresión por pasos mostró que la duración máxima y media de los segmentos sin voz, el porcentaje de la duración del silencio y la duración mínima de los segmentos fraseales explican el 47,8% de la variación de la puntuación del MMSE.
(Meilán et al., 2020)	ndDCL (73) DCL preEA (13)	Lectura oral	La duración y el tiempo de fonación, el número de pausas y varios parámetros de frecuencia e intensidad diferencian a las personas con DCL que desarrollan Alzheimer de las que no.
(Themistocleous et al., 2020)	CS (29) DCL (26)	Lámina del robo de galletas del Boston	H1-A3, Cepstral Peak Prominence, Centro de gravedad, Shimmer, duración media de sílabas, velocidad de articulación.
(Vincze et al., 2021)	CS (25) DCL (25) EA (25)	Recuerdo inmediato y demorado de un vídeo, recuerdo del día anterior	El número de pausas y su duración y la velocidad de habla diferencian a CS, DCL y EA en tareas de memoria.

(Del Carmen Pérez-Sánchez et al., 2021)	CS (20) EA (20)	Lectura de un texto con frases declarativas, exclamativas e interrogativas.	Las alteraciones dependen del tipo de frase. Aparecen más en las interrogativas y exclamativas. Tienen que ver con el número de pausas y la F0, duración e intensidad de las sílabas.
(Yeung et al., 2021)	CS (10) DCL (10) EA (10)	Lámina del robo de galletas del Boston	Los parámetros de pausas se relacionan con dificultades para acceder a palabras, los espectrales con la incoherencia del discurso y los errores.
(Pastoriza-Domínguez et al., 2022)	CS (29) DCLcodificación (28) DCLrecuperación (29) EA (26)	Lámina del Bilingual Aphasia Test.	Las personas con DCL con fallos de codificación producen más pausas y con más variabilidad que las que muestran DCL con fallos de recuperación.

Estudios descriptivos sobre el habla de los pacientes con EA y DCL

Como se ha indicado anteriormente, el primer estudio sobre este tema fue publicado en 2010 por Hoffmann et al., (2010). Utilizaron el software PRAAT para analizar el habla espontánea de un grupo de control de personas mayores sanas y de un grupo de EA. En este estudio se encontraron varias características afectadas por la enfermedad, como más roturas de voz, mayor tiempo de habla y de fonación, y menor velocidad de habla y de articulación. Ese mismo año, otro estudio realizado por Horley et al., (2010) exploró la prosodia emocional en pacientes con EA, encontrando diferencias en la F0 al expresar sorpresa o felicidad, en comparación con un grupo de control. Los pacientes con EA tenían una expresión prosódica deteriorada cuando intentaban imitar el habla emocional. Hallazgo replicado en estudios más recientes (Del Carmen Pérez-Sánchez et al., 2021). Además, otros aspectos alterados del habla se fueron sumando en sucesivos estudios como cambios en las pausas y la F0, diferencias en la frecuencia de las roturas de voz y de los segmentos de habla sin voz (Martínez-Sánchez et al., 2012) y menos variaciones en la señal del habla (Nasrolahzadeh, Mohammadpoori, et al., 2016; Nasrolahzadeh, Mohammadpoory, et al., 2016).

En cuanto a las personas con DCL, los esfuerzos se han centrado principalmente en la detección, como se señalará en la siguiente sección. Sin embargo, hay varios trabajos que hacen una aproximación descriptiva, encontrando diferencias con respecto a CS en el número, duración y variabilidad de las pausas del habla (Vincze et al., 2021) y en parámetros espectrales, duración de las sílabas y velocidad de articulación (Themistocleous et al., 2020). Dentro de las personas con DCL, se han encontrado rasgos que caracterizan diferentes subtipos de deterioro, como diferencias entre el DCL amnésico y el DCL multidominio (Beltrami et al., 2018) en rasgos rítmicos como el PVI, y especialmente en la duración de los segmentos de silencio y habla, de fonación y de

pausa, y el centroide espectral, que hace referencia al punto del espectro donde se acumula la mayoría de la energía de sonido. En un estudio, Meilán et al., (2020) han tratado de encontrar dos perfiles diferentes dentro de las personas con DCL, y han planteado la hipótesis de que diferentes características del habla permiten diferenciar entre las personas con DCL que desarrollarán EA y aquellas que no sufrirán un mayor deterioro del propio del DCL. En una aproximación similar, Pastoriza-Domínguez et al., (2022) han demostrado diferencias en la longitud y variabilidad de las pausas en aquellos DCL amnésico que muestran fallos de codificación de memoria frente a otros con fallos de recuperación, considerando que los primeros son más proclives a evolucionar a EA.

Esta búsqueda descubrió seis artículos que se proponían explicar los cambios en términos de procesos cognitivos. Dos de ellos pretendían explorar si una serie de variables neuropsicológicas de acceso al léxico podían predecir los cambios en los rasgos del habla. Meilán et al., (2012) han encontrado que una tarea de fluidez verbal fonológica explica el 46% de la varianza en los segmentos sin voz. En otro estudio, los mismos autores (Meilán et al., (2018) han informado de que las tareas de fluidez semántica y verbal podrían explicar el 30,1% de la varianza en la proporción de segmentos sin voz y el 26,4% del porcentaje en las pausas vocales, es decir, la proporción del tiempo de pausa con respecto al tiempo de habla. De Looze et al., (2018) han utilizado oraciones que difieren en longitud y complejidad sintáctica, encontrando que sus déficits de habla se relacionan con déficits de memoria de trabajo y atención en el caso de personas con DCL, y con déficits de lenguaje en pacientes con EA. Yeung et al., (2021) han correlacionado los parámetros de pausas con dificultades en el acceso léxico-semántico, y asociaron los parámetros espectrales con marcadores de incoherencia y errores en el discurso. Finalmente, un estudio señala la relación del habla con el estado cognitivo medido a través del MMSE.

El de Qiao et al., (2020) ha encontrado que siete parámetros explican el 47,8% de la variación en la puntuación del test.

Estudios de cribado de la EA y el DCL

Tres años después de la publicación del primer artículo sobre el análisis automático del habla de los pacientes con EA, se publicaron casi simultáneamente los tres primeros artículos sobre la clasificación de controles sanos y personas con EA a partir de su señal de habla. López-de-Ipiña, Alonso, Solé-Casals, et al., (2013) analizaron una tarea de habla espontánea logrando una precisión del 97,7%. Utilizaron dos conjuntos de parámetros que incluían rasgos acústicos, de calidad de voz, de duración, prosódicos y paralingüísticos. El mismo grupo registró un 93,79% de precisión utilizando solo dos características: porcentaje de segmentos de habla con y sin voz (López-de-Ipiña, Alonso, Travieso, et al., 2013). Martínez-Sánchez et al., (2013) utilizaron una tarea de lectura oral y obtuvieron un 80% de precisión mediante la velocidad de habla y de articulación. El mismo grupo (Meilán et al., 2014) elevó ese valor al 84,4% utilizando cuatro parámetros: porcentaje de roturas de voz, número de roturas de voz, número de periodos o pulsos glotales, HNR y shimmer apq3 (parámetro que hace referencia a la variación de amplitud de un pulso glotal con respecto al inmediatamente anterior y posterior). Tras estos primeros estudios se desarrollarían diferentes enfoques en los que se experimentaría con las tareas para elicitación del lenguaje oral y los métodos de selección de parámetros discriminatorios. Por un lado, veremos aproximaciones que utilizan uno o unos pocos parámetros. Por ejemplo Khodabakhsh et al., (2015) utilizaron el ratio de silencios obteniendo un 83,5%, o Martínez-Sánchez et al., (2017) que utilizaron la desviación típica de la duración de las sílabas, con un 87% de precisión. Por otro lado, será lo más habitual utilizar métodos de aprendizaje automático para intentar aumentar la precisión (Chien

Tabla 2. Estudios de cribado de DCL y EA.

Artículo	Muestra	Tarea	Parámetros	Valor de precisión
(López-de-Ipiña, Alonso, Travieso, et al., 2013)	CS (20) EA (20)	Contar historias personales y conversación	Combinación de dos sets de parámetros: Análisis emocional del habla (parámetros acústicos, de calidad de voz y de duración) y Temperatura emocional (parámetros prosódicos y paralingüísticos)	EA: 75.2 - 97.7%
(López-de-Ipiña, Alonso, Solé-Casals, et al., 2013)	CS (20) EA (20)	Contar historias personales y conversación	% segmentos con voz, % segmentos sin voz	EA: 83.7 - 93.79%
(Martínez-Sánchez et al., 2013)	CS (35) EA (35)	Lectura oral	Punto de corte para velocidad de habla: 3.08; sílabas por Segundo 4.27.	EA: 80%
(Meilán et al., 2014)	CS (36) EA (30)	Lectura oral	Porcentaje de roturas de voz, periodos de voz, roturas de voz y shimmer (apq3)	EA: 84.4%
(Khodabakhsh & Demiroglu, 2015)	CS (27) EA (27)	Conversación	Combinación de 13 parámetros (relacionados con actividad vocal, articulación y ritmo)	EA: 75.5 – 94.3%
(Khodabakhsh et al., 2015)	CS (51) EA (28)	Conversación	Porcentaje de segmentos sin voz	EA: 78.5 - 83.5%
(López-de-Ipiña, Alonso, et al., 2015)	CS (20) EA (20)	Contar historias personales y conversación	Combinación de parámetros	EA: 96.89%
(López-de-Ipiña, Solé-Casals, et al., 2015)	CS (20) EA (20)	Contar historias personales y conversación	Selección automática de parámetros del habla espontánea y de los valores máximos, mínimos, de varianza, de desviación estándar y mediana para la señal completa y la señal con voz	EA: 87,30 - 92.43%

(Martínez-Sánchez et al., 2017)	CS (82) EA (45)	Lectura oral	Desviación típica de la duración de las sílabas	EA: 87%
(Nasrolahzadeh et al., 2018)	CS (30) EA (30)	Contar historias personales y conversación	Análisis espectral de orden superior	EA: 94.18 - 97.71%
(König et al., 2015)	CS (15) DCL (23) EA (26)	Contar hacia atrás, repetición de frases, descripción de imagen, fluidez verbal.	Combinación de parámetros significativos extraídos de varias tareas	DCL: 79% EA: 87% DCL vs EA: 80%
López-de-Ipiña et al. (2018a)	CS (187) DCL (38)	Fluidez verbal semántica (animales)	Se extraen un total de 920 parámetros. Los mejores resultados se obtienen con un conjunto de 25 características.	DCL: 92 - 95%
López-de-Ipiña et al. (2018b)	Tres muestras: Fluidez verbal (187 CS y 38 DCL); descripción de lámina (12 CS y 6 EA); lenguaje espontáneo (50 CS y 20 EA)	Fluidez verbal semántica (CS vs DCL) Descripción de lámina y lenguaje espontáneo (CS vs EA)	Selección de los parámetros más relevantes de cada comparación	DCL: 73% EA: 89 - 95% (usando la tarea de lenguaje espontáneo)
Kato, Homma, & Sakuma (2018)	CS (91) DCL (91) EA (91)	Responder un cuestionario indicando lugar de nacimiento (T1), nombre de su colegio de primaria (T2), orientación temporal (Q2),	SPCIR (speech prosody-based cognitive impairment rating; 128 parámetros)	EA: 74.7 - 89.5% (usando Q2) DCL: 70.9 - 76.4% (usando Q6)

		repetición de tres números hacia atrás (Q6)		
(Themistocleous et al., 2018)	CS (30) DCL (25)	Lectura (se extrajeron las vocales de las frases)	Formantes, F0, duración de las vocales	DCL: 75 - 83%
(Toth et al., 2018)	CS (36) DCL (48)	Recuerdo inmediato y demorado de un vídeo	Combinación de parámetros (duración, velocidad de habla, velocidad de articulación y pausas)	DCL: 75%
(Fraser, Meltzer, et al., 2016)	CS (97) EA (167)	Lámina del robo de galletas del Boston	Combinación de entre 35 y 50 parámetros acústicos, semánticos y sintácticos.	EA: 78.72 - 81.92%
(Fraser et al., 2019)	CS (29) DCL (26)	Lámina del robo de galletas del Boston y lectura	Parámetros acústicos + seguimiento ocular + lingüísticos	DCL: 41 - 83%
(Gosztolya et al., 2019)	CS (25) DCL (25) EA (25)	Recuerdo inmediato y demorado de un vídeo	Set 1: parámetros acústicos (velocidad de habla y número y duración de pausas) Set 2: parámetros acústicos + lingüísticos	Set 1: 74-82% Set 2: 80-86%
(König, Satt, et al., 2018)	SCI (56) DCL (44) DV (38) EA (27)	Fluidez, descripción de lámina, contar hacia atrás y lenguaje espontáneo	Diferentes combinaciones de parámetros extraídos para cada comparación.	SCI vs. EA = 92% SCI vs. VD = 92% SCI vs. DCL = 86% DCL vs. EA = 86%.
(Martínez-Sánchez et al., 2018)	CS (98) EA (47)	Lectura	Edad, amplitud mínima, diferencia de amplitud máxima, media y desviación estándar del NHR; asimetría; desviación estándar en el primer formante; ancho de banda del formante 3; desviación estándar del Acoustic Voice Quality Index; variabilidad del tono; nPVI	EA: 92.4%

Al-Hameed et al. (2019)	CS con quejas de memoria (15), Enfermedades neurodegenerativas (15: EA 10, aDCL 2, demencia frontotemporal 2, VD 1)	Conversación	Varios sets de parámetros (el mejor utiliza nueve)	Enfermedades neurodegenerativas: 81 – 92%
(Chien et al., 2019)	CS (30) EA (30)	Respuestas a pruebas neuropsicológicas	Secuencia de características (una representación de varios elementos del habla)	EA: AUC .838
(Nagumo et al., 2020)	CS (6343) DCL (1601) global cognitive impairment (367) DCL + GCI (468)	Pronunciamientos vocales, trabalenguas, ritmo diadococinético, frases cortas	Set parámetros temporales y acústicos	DCL: AUC .61
(O'Malley et al., 2020)	CS (15) EA (15) DCL (15) ndDCL (15)	Responder preguntas de un asistente virtual	Combinación de parámetros acústicos, léxicos y de análisis de conversación	EA y DCL vs ndDCL y C: 77-87% EA vs DCL vs ndDCL y C: 65% EA vs DCL vs ndDCL vs C: 60%
(Thomas et al., 2020)	CS (64) Sin Dem (69) Dem inicio (61)	Respuestas en una evaluación neuropsicológica	Set1: (solo acústico) Rango intercuartílico de F0, MFCCs, Ventana de F2, pendiente espectral Set2: acústico + datos demográficos	CS y sin Dem vs Dem Set1: 75,8

	Dem leve (62) Dem moderada (39) Dem Severa (25)		Set3: acústico + demográficos + lingüísticos	Set2: 85,5 Set3: 94,2
(Chlasta & Wolk, 2020)	CS (78) EA (78)	Lámina del robo de galletas del Boston	Combinación de parámetros acústicos	EA: 54,2 – 62,5%
(Haulcy & Glass, 2020)	CS (78) EA (78)	Lámina del robo de galletas del Boston	i-vector y x-vector	EA: 45,8 – 65,7%
(Jang et al., 2021)	CS (83) Dem (DCL + EA)	Lámina del robo de galletas del Boston, lectura y recuerdo	Combinación de parámetros acústicos (MFCCs), lingüísticos y seguimiento ocular	Dem: 74-78%
(C. De Looze et al., 2021)	CS (12) DCL (10) EA (10)	Interacción con cuidador mientras se hacen tareas con el juego tangram	Parámetros temporales de los turnos de palabra, pausas y velocidad de habla.	82,47 – 84,17
(Balagopalan et al., 2021)	CS (78) EA (78)	Lámina del robo de galletas del Boston	No hay parámetros acústicos seleccionados en el algoritmo final.	-
(Lindsay et al., 2021)	CS (78) EA (76)	Lámina del robo de galletas del Boston	No hay parámetros acústicos seleccionados en el algoritmo final.	-
(Mahajan & Baths, 2021)	CS (82) EA (82)	Lámina del robo de galletas del Boston	Set de parámetros acústicos extraídos con eGEMAPS, emobase y ComParE.	EA: 52,8-72,9%
(Shimoda et al., 2021)	CS (99) EA (24)	Habla espontánea por teléfono	Segmentos sin voz, duración, media y variabilidad del tono y la intensidad, parámetros espectrales	EA: ,863 - ,893 (AUC)
(Martinc et al., 2021)	CS (78) EA (78)	Lámina del robo de galletas del Boston	Set de parámetros acústicos y lingüísticos	EA: 91,67%

(Tang et al., 2021)	CS (80) DCL (80)	Conversación semi-estructurada	Set 1: acústicos MFCCs Set 2: Acústicos + lingüísticos	DCL Set1: 65% Set2: 82,7%
(Yamada, Shinkawa, Kobayashi, Caggiano, et al., 2021)	CS (47) DCL (45) EA (26)	contar hacia atrás, restar, fluidez verbal fonológica y semántica, y Lámina del robo de galletas del Boston	Set 1: parámetros acústicos (MFCCs y prosodia). Set 2: parámetros acústicos, dibujo y marcha	CS vs DCL vs EA Set 1: 81,9% Set 2: 93%
(Wang et al., 2021)	CS (60) DCL (50)	Descripción de imagen, fluidez verbal semántica y repetición de frases	Combinación de parámetros acústicos, sintácticos, semánticos y léxicos.	DCL: 94-96%
(Yamada, Shinkawa, Kobayashi, Nishimura, et al., 2021)	CS (39) DCL (37)	Preguntas sobre actividades de la vida diaria, respuestas a test neuropsicológicos	MFCCs, pausas, variabilidad del tono, velocidad de habla, jitter, shimmer, tiempo de fonación, duración.	DCL: 86,4-86,8%
(Momeni & Rahmani, 2021)	CS (10) EA (11)	Conversación	Parámetros de modulación temporal, espectral y de frecuencias.	EA: 99,99%
(Xue et al., 2021)	CS (291) DCL (309) EA (223)	Evaluación neuropsicológica	MFCCs	EA: 57,8-74% CS+DCL vs EA: 55,5-67,5%

et al., 2019; Khodabakhsh & Demiroglu, 2015; López-de-Ipiña, Alonso-Hernández, et al., 2015; López-de-Ipiña, Solé-Casals, et al., 2015; Shimoda et al., 2021; Xue et al., 2021). Los estudios que logran mayores precisiones utilizan este método, siendo los de (Nasrolahzadeh et al., 2018) con un 97,71% utilizando un análisis espectral de orden superior con una tarea de habla espontánea, y de Momeni y Rahmani, (2021) con un 99,99% utilizando parámetros de modulación temporal, espectral y de frecuencias. Este último, sin embargo, podría tratarse de un efecto sobredimensionado debido al avanzado estado de deterioro de los participantes.

En 2015, se publicó el primer estudio que utilizó el análisis automático del habla para identificar el DCL. König et al., (2015) compararon las voces de adultos mayores sanos, personas con DCL y pacientes con EA. Extrajeron parámetros que mostraban diferencias significativas en varias tareas y obtuvieron la mejor combinación mediante aprendizaje automático. Esto dio lugar a una precisión del 79% para el DCL y del 87% para la EA. Además, pudieron diferenciar a las personas con DCL de las que tenían EA con un 80% de precisión. Otros trabajos han venido obteniendo resultados similares con diferentes parámetros prosódicos (Kato et al., 2018), estructuras de formantes (Themistocleous et al., 2018), o parámetros de duración, velocidad de habla y articulación y pausas (Toth et al., 2018). López-de-Ipiña, Martínez-de-Lizarduy, Calvo, Beitia, et al., (2018) y López-de-Ipiña, Martínez-de-Lizarduy, Calvo, Mekyska, et al., (2018) también utilizaron combinaciones de parámetros con una precisión del 73-95% para DCL, logrando el mejor resultado. Otras aproximaciones incluyen la de Al-Hameed et al., (2019) en la que reunieron a un grupo con varios trastornos neurodegenerativos, entre ellos la EA y DCL amnésico, y los clasificaron correctamente con un 92% de precisión. Uno de los pocos estudios que trataron de diferenciar entre CS, DCL y EA con un único algoritmo es el de De Looze et al., (2021) en el que consiguen un 84,17% usando

parámetros temporales de los turnos de palabra, pausas y velocidad de habla. El estudio con la mayor muestra hasta la fecha es también el más decepcionante en cuanto a resultados. Con más de 8.000 participantes, sólo consiguieron un AUC de 0,61, aunque como los propios autores indican esto podría deberse a una clasificación poco fiable de la muestra. Sin embargo, defienden que el análisis del habla es una buena medida de la gravedad de la deficiencia (Nagumo et al., 2020).

Otro abordaje es el de estudios que combinan el análisis automático del habla con otras técnicas en busca de mejores resultados. Uno de los enfoques más comunes es el de añadir otra clase de análisis del lenguaje, principalmente parámetros semánticos, sintácticos y léxicos. Por ejemplo, un estudio de Gosztolya et al. (2019) ha mejorado la clasificación de los grupos de control, DCL y EA, pasando de un 74-82% de precisión al utilizar solo características acústicas a un 80-86% al añadir parámetros lingüísticos. En otros estudios se obtienen resultados muy similares (Fraser, Meltzer, et al., 2016; O'Malley et al., 2020; Tang et al., 2021). Los más destacados serían los de Wang et al., (2021) con un 96% en CN frente a DCL y Thomas et al., (2020) que llega al 94% sumando datos sociodemográficos. Otras técnicas novedosas serían el seguimiento ocular (Fraser et al., 2019; Jang et al., 2021) y parámetros del dibujo y de la marcha (Yamada, Shinkawa, Kobayashi, Caggiano, et al., 2021) llegando a un 93% de precisión para detectar DCL.

Por último, hay otro tipo de estudios significativos entre los que pretenden predecir la EA y el DCL. A partir de 2018, el conocimiento acumulado permitió dar un paso más con la aparición de estudios que programan el análisis automático del habla en dispositivos o aplicaciones que los clínicos podrían utilizar como herramienta de cribado. En esta búsqueda se han encontrado tres aproximaciones al respecto: la primera de ellas, de König et al., (2018), utiliza una aplicación móvil para discriminar entre personas con deterioro cognitivo subjetivo (DCI), DCL, demencia vascular (DV) y EA, con una

precisión que va del 86% al 92%. También con una aplicación, Yamada et al., (2021b) diferencian el DCL con un 86% de precisión utilizando parámetros de MFCCs (análisis del espectro), pausas, variabilidad del tono, velocidad de habla, jitter, shimmer, tiempo de fonación y duración. Por otro lado, Martínez-Sánchez et al., (2018) presentaron un dispositivo tecnológico que procesa un algoritmo computacional que utiliza nueve parámetros para detectar la EA con una precisión del 92,4%.

En este apartado es necesario señalar una iniciativa publicada en el año 2020 conocida como el ADReSS Challenge (Luz et al., 2020). Aunque debido a los criterios de inclusión del estudio original no fue incluido en los resultados de la revisión, ha dado lugar a diversas publicaciones posteriores. En ese documento, los autores proponían usar un conjunto de grabaciones correspondientes a la lámina del robo de galletas del Boston de 78 sujetos del grupo CN y 78 AD que ponían a disposición de otros investigadores con el objetivo de avanzar en el diagnóstico de Alzheimer mediante el lenguaje. En su publicación, los autores proporcionaban como línea base de la que partir los resultados de su propia clasificación, obteniendo un 62,5% utilizando parámetros acústicos. A partir de este reto, se han publicado diversos trabajos que sí entran dentro de la revisión. Algunos, obtienen resultados muy similares, que mejoran en poco la precisión obtenida por Luz (Chlasta & Wołk, 2020; Haulcy & Glass, 2020; Mahajan & Baths, 2021). De hecho, dos de los trabajos son muy desalentadores para el análisis del habla. Balagopalan et al., (2021) utilizan aprendizaje automático combinando parámetros lingüísticos y acústicos, pero a pesar de encontrar diferencias en algunos como las pausas o algunos coeficientes MFCCs, ninguno entra en el algoritmo de clasificación final. Por otra parte, Lindsay et al., (2021) utilizan dos bases de datos, la del ADReSS Challenge para voces en inglés y otra en francés, extrayendo parámetros sintácticos, semánticos y del habla. Concluyen que ningún parámetro del habla es diferenciador. De hecho, argumentan que

las diferencias encontradas no son consistentes en los idiomas y que desaparecen tras correcciones estadísticas, y que, por tanto, nos encontraríamos ante artefactos. En este sentido, el trabajo de Mahajan y Baths, (2021), intenta dar una explicación a los pobres resultados de las investigaciones derivadas del ADReSS Challenge. Se basan en primer lugar en la falta de sujetos ya que el aprendizaje automático se beneficiaría de bases de datos más grandes y piensan que con una muestra como la que tienen, el enfoque a seguir en todo caso sería el de utilizar parámetros relevantes en la detección de EA, como los señalados en la literatura, en lugar de recurrir a selecciones automáticas. Se defienden además diciendo que las precisiones de otras bases de datos como las derivadas del DementiaBank, por encima del 90%, estarían hinchadas debido a la homogeneidad de las muestras. El único trabajo derivado del ADReSS challenge con resultados prometedores que se ha incluido en esta revisión es el de los propios autores convocantes, que combinando parámetros acústicos y lingüísticos, logran llegar a un 91,67% (Martinc et al., 2021).

Calidad metodológica de los estudios del habla en DCL y EA según los criterios del JBI y QUADAS-2

Los estudios descriptivos registraron una buena calidad en general, según la evaluación de la lista de comprobación para estudios transversales JBI. La evaluación específica de cada dimensión de la lista de comprobación como de riesgo de sesgo alto, bajo o incierto para cada estudio puede verse en la Figura 11, con un resumen de los resultados globales en la Figura 12. Más de la mitad de los estudios (62,5%) abordaron adecuadamente todas las cuestiones. Las preocupaciones más comunes surgieron de la definición de la muestra: dos estudios carecían de criterios de inclusión, y cuatro de ellos no describían la muestra de forma clara o suficiente en términos objetivos. Sin embargo,

la asignación a los grupos se realizó siguiendo los criterios NINCS-ADRDA o similares. Sólo dos estudios no identificaron o controlaron potenciales factores de confusión; el resto de ellos utilizó el emparejamiento de los participantes o correcciones en el análisis de los datos.

Para la evaluación de la calidad de los estudios de cribado se utilizó la herramienta QUADAS-2. La calidad de los estudios es aceptable o alta (ver figuras 13 y 14). Las principales fuentes de preocupación se encontraron en el ámbito de la selección de pacientes, al igual que en los estudios descriptivos. La razón más común para la identificación de un posible sesgo en los estudios evaluados fue la de no utilizar una secuencia consecutiva ni una selección aleatoria de la muestra, sino que se selecciona una muestra de pacientes y luego otra emparejada para el grupo de control. Aunque no es un defecto importante, plantea dudas sobre la generalización de los resultados. La mayoría de los estudios recibieron una calificación de sesgo bajo para el estándar de referencia, ya que se utilizaron pruebas adecuadas. En los casos marcados como poco claros, no se proporcionó suficiente información sobre las pruebas utilizadas. Cuando la calificación fue de alto riesgo de sesgo, las pruebas que utilizaron no fueron suficientes para asignar la categoría diagnóstica. La mayoría de los estudios recibieron una calificación de bajo riesgo de sesgo en lo que respecta a la inaplicabilidad, ya que la selección de pacientes, la prueba índice y el estándar de referencia coincidían con los objetivos y las preguntas de los estudios evaluados.

Estudio	P1	P2	P3	P4	P5	P6	P7	P8
Hoffman et al. (2010)	■	■	■	■	■	■	■	■
Horley et al. (2010)	■	■	■	■	■	■	■	■
Martínez- Sánchez et al. (2012)	■	■	■	■	■	■	■	■
Meilán et al. (2012)	■	■	■	■	■	■	■	■
Nasrolahzadeh et al (2016a)	■	■	■	■	■	■	■	■
Nasrolahzadeh (2016b)	■	■	■	■	■	■	■	■
Beltrami et al. (2018)	■	■	■	■	■	■	■	■
Meilán et al. (2018)	■	■	■	■	■	■	■	■
De Looze et al (2018)	■	■	■	■	■	■	■	■
Quiao et al. (2020)	■	■	■	■	■	■	■	■
Meilán et al. (2020)	■	■	■	■	■	■	■	■
Themistocleouss et al. (2020)	■	■	■	■	■	■	■	■
Pérez-Sánchez et al (2021)	■	■	■	■	■	■	■	■
Yeung et al (2021)	■	■	■	■	■	■	■	■
Pastoriza-Domínguez (2021)	■	■	■	■	■	■	■	■
Vincze et al (2020)	■	■	■	■	■	■	■	■

1. ¿Se han definido claramente los criterios de inclusión en la muestra?	■	Alto
2. ¿Se describieron detalladamente los sujetos del estudio y el entorno?	■	Incierto
3. ¿Se midió la exposición de forma válida y fiable?	■	Bajo
4. ¿Se utilizaron criterios objetivos y normalizados para medir la afección?	■	
5. ¿Se identificaron los factores de confusión?	■	
6. ¿Se indicaron las estrategias para hacer frente a los factores de confusión?	■	
7. ¿Se midieron los resultados de forma válida y fiable?	■	
8. ¿Se utilizó un análisis estadístico adecuado?	■	

Figura 11. Evaluación de la calidad de los estudios descriptivos mediante la lista de comprobación JBI, y su calificación es como riesgo de sesgo alto, bajo o incierto para cada ítem.

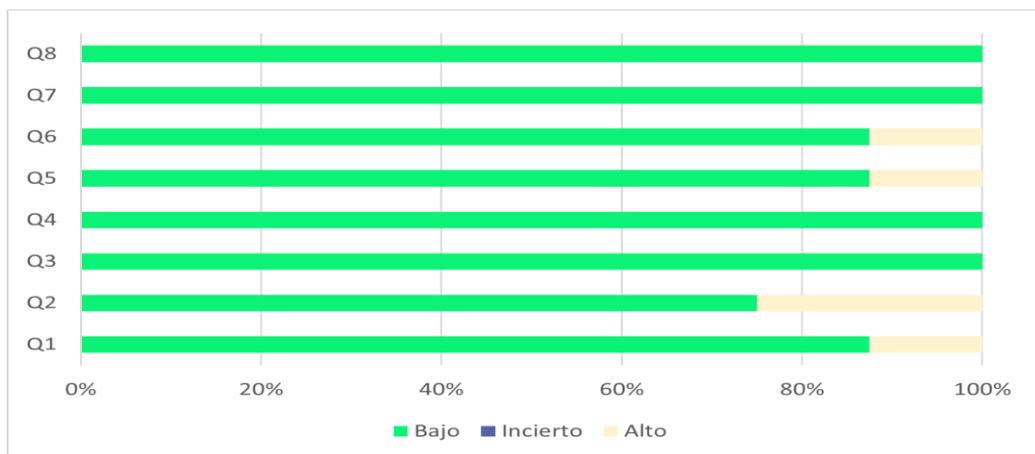


Figura 12. Proporción de estudios descriptivos con evaluación de riesgo de sesgo bajo, alto o incierto.

	Probabilidad de sesgos				Preocupación sobre la aplicabilidad		
	Selección de los sujetos	Prueba índice	Prueba de referencia	Flujo y tiempos	Selección de los pacientes	Prueba índice	Prueba de referencia
López-de-Ipiña, Alonso, Travieso, et al., (2013)							
López-de-Ipiña, Alonso, Solé-Casals, et al., (2013)							
Martínez-Sánchez et al., (2013)							
Meilán et al., (2014)							
Khodabakhsh & Demiroglu, (2015)							
Khodabakhsh et al., (2015)							
López-de-Ipiña, Alonso, et al., (2015)							
López-de-Ipiña, Solé-Casals, et al., (2015)							
Martínez-Sánchez et al., (2017)							
Nasrolahzadeh et al., (2018)							
König et al., (2015)							
López-de-Ipiña et al. (2018a)							
López-de-Ipiña et al. (2018b)							
Kato, Homma, & Sakuma (2018)							
Themistocleous et al., (2018)							
Toth et al., (2018)							
Fraser, Meltzer, et al., (2016)							
Fraser et al., (2019)							
Gosztolya et al., (2019)							

König, Satt, et al., (2018)							
Martínez-Sánchez et al., (2018)							
Al-Hameed et al. (2019)							
Chien et al., (2019)							
Nagumo et al., (2020)							
O'Malley et al. (2020)							
Thomas et al. (2020)							
Chlasta y Wolks (2020)							
Haulcy y Glass (2020)							
Jang et al. (2021)							
De Looze et al. (2021)							
Balagopalan et al. (2021)							
Lindsay et al. (2021)							
Mahajan y Baths (2021)							
Shimoda et al. (2021)							
Martinc et al. (2021)							
Tang et al. (2022)							
Yamada et al. (2021) b							
Wang et al. (2021)							
Yamada et al. (2021) a							
Momeni y Rahmani (2020)							
Xue et al (2021)							

 Bajo

 Alto

 Incierto

Figura 13. Evaluación de la calidad de los estudios de cribado mediante la lista de comprobación QUADAS-2 y su clasificación como riesgo de sesgo alto, bajo o incierto.

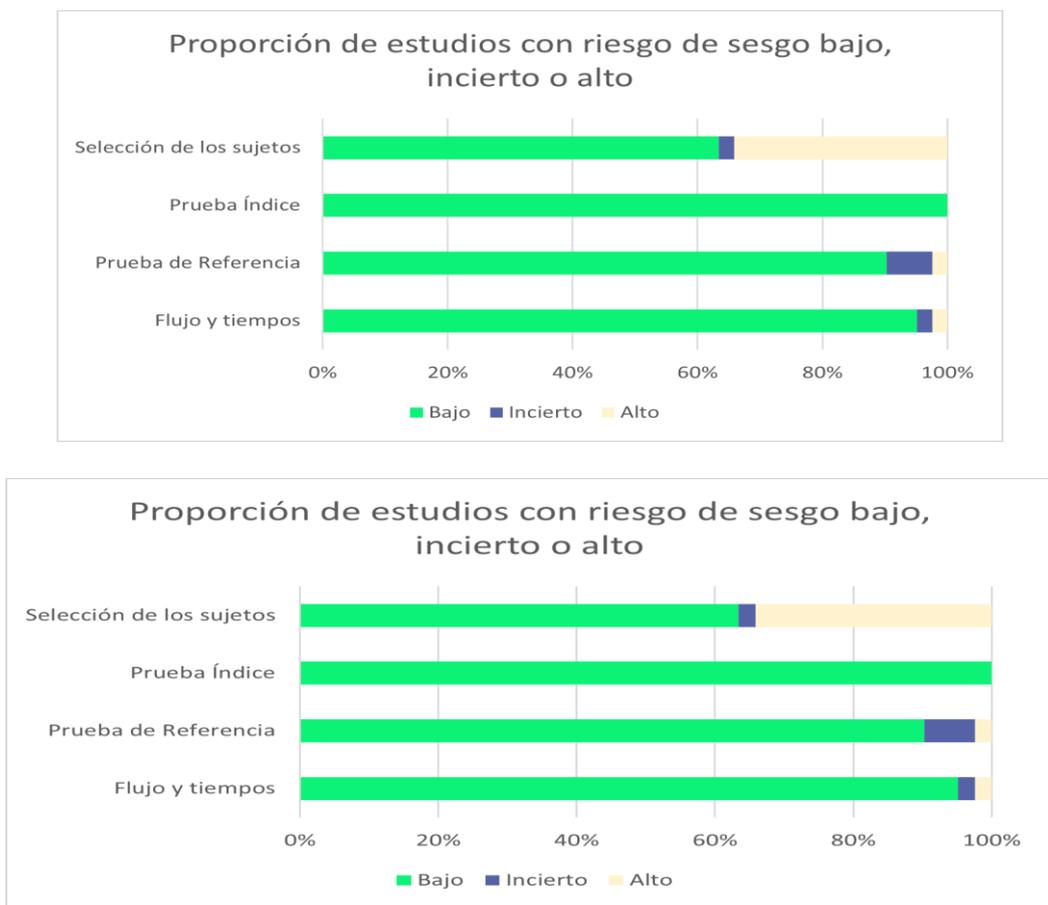


Figura 14. Proporción de los resultados de la evaluación de riesgos de sesgo según el Quadas-2.

En resumen, parece razonable afirmar que existen pruebas sólidas de que hay cambios en el habla de las personas mayores relacionados con el deterioro cognitivo y además se pueden utilizar para detectar el DCL y la EA, cuando se miden mediante el análisis automático de la señal acústica del habla. Las preocupaciones planteadas en relación con el riesgo de sesgo en los estudios fueron más frecuentes en los primeros años, y los estudios más recientes registran una mayor precisión, lo que en última instancia significa que se trata de un campo maduro y con una fundamentación sólida.

Discusión

El estudio del análisis automático del habla de las personas con EA y DCL ha ido evolucionando a lo largo de los años y ha ido creciendo en popularidad (ver figura 15). En un primer periodo, a partir de 2010, los esfuerzos se centraron en la extracción de parámetros que definieran las voces de los pacientes con EA. Estos estudios pronto llevarían a los primeros intentos, ya en 2013, de identificar la EA mediante el análisis de la señal acústica del habla. El éxito alcanzado por estos últimos estudios provocó un cambio en el enfoque. En 2015, aparecieron los primeros estudios sobre DCL, que aún coexistían con las investigaciones destinadas a perfeccionar el método para la EA y conseguir mejores resultados. Aunque menos preciso, el diagnóstico del DCL era prometedor, y se aplicarían métodos más complejos en la búsqueda de una mayor precisión. En 2018 se produjo un gran avance: se realizaron los primeros intentos de extrapolar los resultados del laboratorio a la práctica clínica. La inclusión del procedimiento en dispositivos y aplicaciones sin pérdida de efectividad supone la aparición de una herramienta de cribado rápida, no invasiva, fiable y barata. Siendo esto

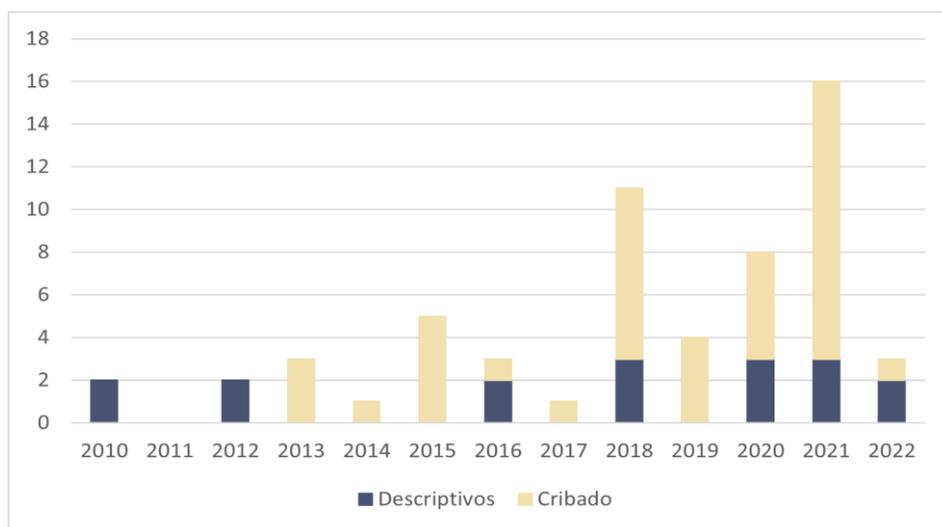


Figura 15. Número de artículos publicados descriptivos y de cribado obtenidos en la revisión.

un reflejo de los avances en la materia, ya que significa que está suficientemente desarrollada para dar el salto del laboratorio a la práctica clínica diaria.

Como se ha puesto de manifiesto, el análisis automático del habla puede detectar cambios sutiles en las voces de las personas que sufren procesos neurodegenerativos como el DCL y la EA. Un resumen de estos cambios puede encontrarse en la Tabla 3, que detalla los principales parámetros hallados en los estudios que muestran cambios para ambos grupos en comparación con personas mayores sanas. Esta tabla muestra qué cambios en los parámetros se asocian a DCL o a estadios más avanzados de demencia según los estudios revisados.

En la EA, los pacientes muestran cambios en varios parámetros temporales y acústicos, como la disminución de la velocidad de habla y articulación y el aumento del número de pausas (López-de-Ipiña, Alonso, Solé-Casals, et al., 2013; Toth et al., 2018). Se observa una mayor continuidad de los segmentos periódicos y armónicos para los adultos mayores sanos que para aquellos con DCL, que es aún menor en las personas con EA (König et al., 2015). Además, se pueden encontrar cambios en parámetros espectrales, como diversos cambios en MFCCs, una frecuencia fundamental menos variable y alteraciones en los formantes (Themistocleous et al., 2018), tanto en personas con DCL como con EA. En el habla de las personas con EA, se produce un aumento del número y la proporción de pausas (Martínez-Sánchez et al., 2012, 2013), y una menor velocidad de habla y articulación (Hoffmann et al., 2010; Horley et al., 2010; König et al., 2015a; Martínez-Sánchez et al., 2013), signos que van empeorando a lo largo de la enfermedad, lo que se traduce en un mayor tiempo de fonación (Hoffmann et al., 2010; König et al., 2015a; Meilán et al., 2012), Otras alteraciones se agravan, presentando más roturas de voz (Hoffmann et al., 2010; König et al., 2015a; Meilán et al., 2014), un mayor porcentaje de segmentos sin voz (López-de-Ipiña, Alonso, Travieso, et al., 2013; Meilán et al.,

2012), mayor variabilidad en los valores de autocorrelación o poca dependencia entre medidas de frecuencias sucesivas de voz (Meilán et al., 2014), mayores cambios en características espectrales (Fraser, Meltzer, et al., 2016), la frecuencia fundamental y una distorsión en los formantes (Khodabakhsh & Demiroglu, 2015; Meilán et al., 2014), y menor ratio entre ruido-armónicos (Meilán et al., 2014).

Tabla 3. Resumen de los parámetros alterados en las personas con DCL y EA.

	Deterioro cognitivo leve	Enfermedad de Alzheimer
Parámetros temporales y ritmo		
Tiempo de habla y de fonación	Mayor duración	Mayor duración.
Pausas	Aumento del número de pausas.	Aumento del número, proporción y duración de las pausas. Mayor variabilidad en la duración de las pausas.
Roturas de voz		Aumentan.
Segmentos sin voz		Aumentan su duración, explican una parte de la varianza del rendimiento en test neuropsicológicos
Velocidad	Menor velocidad de habla y articulación.	Menor velocidad de habla y articulación.
Prosodia emocional		Alteraciones en la expresión de sorpresa y felicidad, pero no en la tristeza.
Sílabas	Variabilidad de la duración.	Aumenta la desviación típica y la media de la duración de las sílabas.

		nPVI alterado.
Planificación fonológica		
Parámetros espectrales		Cambios en parámetros como centroides, MFCCs, centro de gravedad, LTAS.
Frecuencia fundamental	Alteración	Disminución de la media y la variabilidad, causando una prosodia plana.
Autocorrelación		Mayor variabilidad.
Formantes	Alteración de formantes. Mayor duración.	Distorsión de F1, F2 y F3. Cambios más lentos entre formantes.
AVQI		Aumenta la desviación típica.
Perturbación		
NHR		Menor ratio.
Continuidad de segmentos armónicos		Disminuye
Shimmer	Disminución de perturbaciones.	Disminución de las perturbaciones en 3 pulsos y 11 pulsos.
Intensidad		
Intensidad		Disminuye la intensidad mínima. Menor variabilidad.

Es difícil saber a partir de estos estudios si algunos parámetros tienen más peso que otros en el diagnóstico. Algunos estudios han explorado un único parámetro -ritmo de habla (80%), proporción de segmentos de silencio (78%), desviación típica de la duración de las sílabas (87%)- registrando peores resultados, aunque todavía buenos, en comparación con los estudios que utilizan grandes conjuntos de características. La

mayoría de estudios utilizan combinaciones de varios parámetros, mejorando la precisión por encima del 90%; sin embargo, ninguno de ellos explora las funciones individuales de estos parámetros. Tanto si se exploran individualmente como en conjuntos, parece que los parámetros más utilizados son los temporales. Cabe señalar una tendencia cada vez más común, en la que muchos estudios no informan de los parámetros discriminatorios y se refieren a conjuntos o grupos de parámetros, normalmente obtenidos tras utilizar métodos de aprendizaje automático en grandes sets con el fin de encontrar la combinación más potente para distinguir entre las entidades de diagnóstico. Aunque sirve para simplificar los resultados de artículos que a veces se refieren a decenas de parámetros, dificultan la replicación y el consenso.

El análisis automático del habla como herramienta para diagnosticar el DCL y la EA es el aspecto que más publicaciones ha suscitado. Los valores de precisión para la EA oscilan entre el 80% y el 99%, lo que supone una diferencia importante. Para facilitar la comparación con otras herramientas, consideraremos sólo los valores de los estudios con aplicaciones o dispositivos, dado que son los más parecidos a las herramientas clínicas. En estos casos, la precisión para la EA se sitúa en torno al 92%. Para el DCL, la mayoría de los valores oscilan entre el 73% y el 86%, o entre el 82 y el 86% en los estudios con aplicaciones. Los estudios que utilizan biomarcadores como el volumen del hipocampo medido con resonancia magnética (Chupin et al., 2009) o el trazado de amiloide mediante tomografía por emisión de positrones (Morris et al., 2016) registran precisiones superiores al 90% meses o incluso años antes del diagnóstico definitivo. Aunque los resultados de los estudios de biomarcadores son mejores porque pueden identificar la enfermedad antes de que aparezcan los síntomas, son extremadamente caros, consumen mucho tiempo y solo pueden realizarlos profesionales altamente especializados. Haciendo una comparación con otras herramientas de cribado, como pueden ser los

cuestionarios, encontramos que algunos de los más usados como el MMSE o el MoCA tienen AUCs de 0,778 a 0,842 (Pinto et al., 2019), lo cual es más cercano a las herramientas de *screening* basadas en análisis de habla. Aun así, el análisis del habla presenta ventajas como la rapidez y sencillez, y puede realizarse en dispositivos cotidianos como los teléfonos móviles, lo que lo convierte en una valiosa prueba de detección.

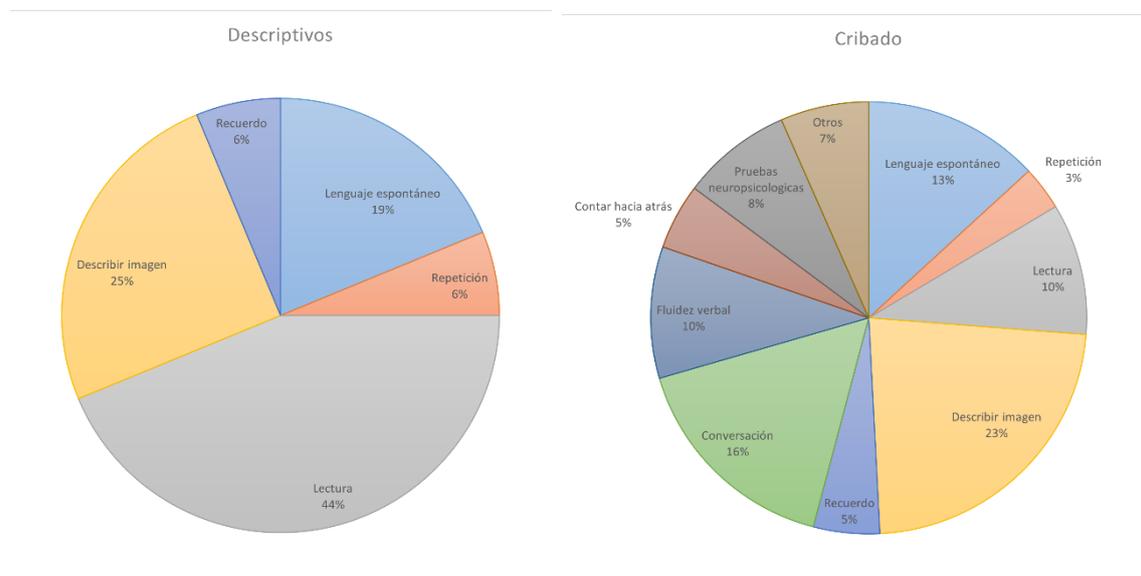


Figura 16. Tareas utilizadas para elicitar lenguaje en los estudios descriptivos y de cribado. Si un estudio utilizó varias tareas, se han incluido todas.

En cuanto al método utilizado para elicitar el lenguaje oral, es un reto establecer si existe una tarea más eficaz. La mayoría de los estudios exploran el habla espontánea y sólo unos pocos analizan tareas de lectura. En la Figura 16 se puede observar un resumen de las tareas utilizadas. Los resultados son bastante similares, y ambas tareas revelan las mismas características alteradas en la EA, como un mayor tiempo de habla, una menor velocidad de habla y un mayor número de pausas. En los estudios de detección de EA, los resultados son variables, y la mayoría de ellos oscilan entre el 80 y el 99% de precisión, independientemente del tipo de tarea. En el caso del DCL, la mayor precisión -95%- se alcanza mediante una tarea de fluidez verbal semántica (López-de-Ipiña, Martínez-de-Lizarduy, Calvo, Beitia, et al., 2018), lo que contrasta notablemente con el

rango que se encuentra normalmente utilizando otras tareas lingüísticas (80-85%). Sin embargo, no es posible afirmar que las tareas de fluidez verbal semántica sean las más eficaces, ya que algunos de esos estudios también utilizaron rasgos acústicos extraídos durante la ejecución de esa tarea. Se esperaría que las tareas más exigentes desde el punto de vista cognitivo provocaran mayores diferencias, tal como se señala en el único estudio que hace una comparación directa (Yamada, Shinkawa, Kobayashi, Nishimura, et al., 2021) pero no siempre parece ser así. A su vez, el uso de diferentes materiales puede ser útil a la hora de explorar los procesos cognitivos. El análisis del habla sirve como medida directa de la ejecución de ciertos procesos que dependerán del tipo de lenguaje. Por ejemplo, la lectura implica una prosodia característica (Dowhower, 1991) y se esperan diferencias con respecto a la prosodia del habla espontánea. Hay otros factores que pueden influir. Un ejemplo claro es la manipulación del material en el estudio de De Looze et al., (2018), en el que al aumentar la complejidad de las frases a leer por parte de las personas con EA se incrementa el número de pausas. Por lo tanto, el mejor enfoque en la actualidad parece ser el de realizar varias tareas y seleccionar diferentes parámetros de cada una para combinarlas en un algoritmo final más potente.

De esta revisión podemos extraer retos y líneas de futuro que pueden derivarse de la información expuesta, algunos de los cuales pretendemos afrontar en esta tesis. Una de las cuestiones de relevancia será el diagnóstico diferencial de las patologías neurodegenerativas, vertiente que hasta el momento ha generado muy pocos estudios y con muestras muy pequeñas. Una de las líneas fundamentales será la de detectar la EA en la etapa asintomática y la conversión de DCL a EA, dado que una de las tendencias que podemos observar es la de involucrar a pacientes en etapas más tempranas. En este sentido, la identificación de ciertos tipos de DCL (amnésico, multidominio) será de gran importancia, ya que constituyen un factor de riesgo para diferentes tipos de demencia.

Por lo tanto, los estudios longitudinales deberían ser una prioridad. Otra de las orientaciones previstas es explicar los cambios en el habla en términos cognitivos, un tema en el que inicialmente se avanzó poco, y que ha ido recobrando fuerza más recientemente. Estos cambios en el habla se suelen explicar principalmente por las alteraciones léxico-semánticas y el deterioro de los procesos ejecutivos y la atención. Cabe destacar que hay estudios que utilizan la extracción manual de parámetros como las pausas (Pistono et al., 2019), que también atribuyen a procesos léxico-semánticos, y se entienden como mecanismos compensatorios para mejorar la selección léxica y el recuerdo. Por último, el uso clínico de esta herramienta es otro gran paso. Como se ha señalado, el análisis automático del habla es un método de detección rápido, barato y fiable. Ya se han identificado varias iniciativas que utilizan pequeños dispositivos (Martínez-Sánchez et al., 2018) y aplicaciones móviles (König, Satt, et al., 2018; Yamada, Shinkawa, Kobayashi, Nishimura, et al., 2021), y no solo las incluidas en la revisión, sino también otras no publicadas en revistas revisadas por pares, como el estudio de Hall et al., (2019) utilizando una aplicación para tablet. Este es un objetivo importante, ya que estos formatos nos permiten transmitir la herramienta a cualquier profesional con un smartphone.

Conclusión

Con esta revisión se muestra que el análisis de la señal acústica del habla en personas con EA y DCL es una herramienta útil para detectar cambios sutiles en su lenguaje. En efecto, existen parámetros rítmicos y acústicos que caracterizan las voces de estos pacientes y que pueden utilizarse con gran éxito para diagnosticar su estado.

Constituye una herramienta eficaz, barata y fácil de usar que puede facilitar el cribado de las demencias.

3. Objetivos experimentales



3. Objetivos experimentales

El objetivo general de este estudio es explorar las funciones de los parámetros del habla alterados en personas mayores sanas, con DCL o EA, de modo que permita desarrollar nuevos y mejores métodos de cribado de esos estados. Atenderemos a varias de las líneas propuestas en la revisión del capítulo anterior, para lo que marcamos los siguientes objetivos específicos:

1. Investigar la causa de los cambios en el habla y sus correlatos cognitivos, observar de qué factores dependen y si pueden ser estimadores de la gravedad del deterioro.
 - 1.1. Comprobar la influencia de factores sociodemográficos y cognitivos sobre los cambios en el habla del mayor. *Estudio I.*
 - 1.2. Evaluar la influencia de diversos procesos cognitivos sobre los parámetros del habla. *Estudio II.*
2. Abordar el cribado de DCL y EA mediante diferentes tareas y observando perfiles concretos que permitan mejorar la detección temprana y la utilidad diagnóstica actual del análisis automático del habla.
 - 2.1. Explorar la existencia de perfiles de habla de personas con DCL que sean predictores del desarrollo de la demencia. *Estudio III.*
 - 2.2. Hallar combinaciones de parámetros óptimas para el cribado de DCL y EA. *Estudios IV y V.*
3. Hacer extensible el método del análisis de habla a nuevas patologías asociadas al envejecimiento como la depresión en mayores. *Estudio VI.*

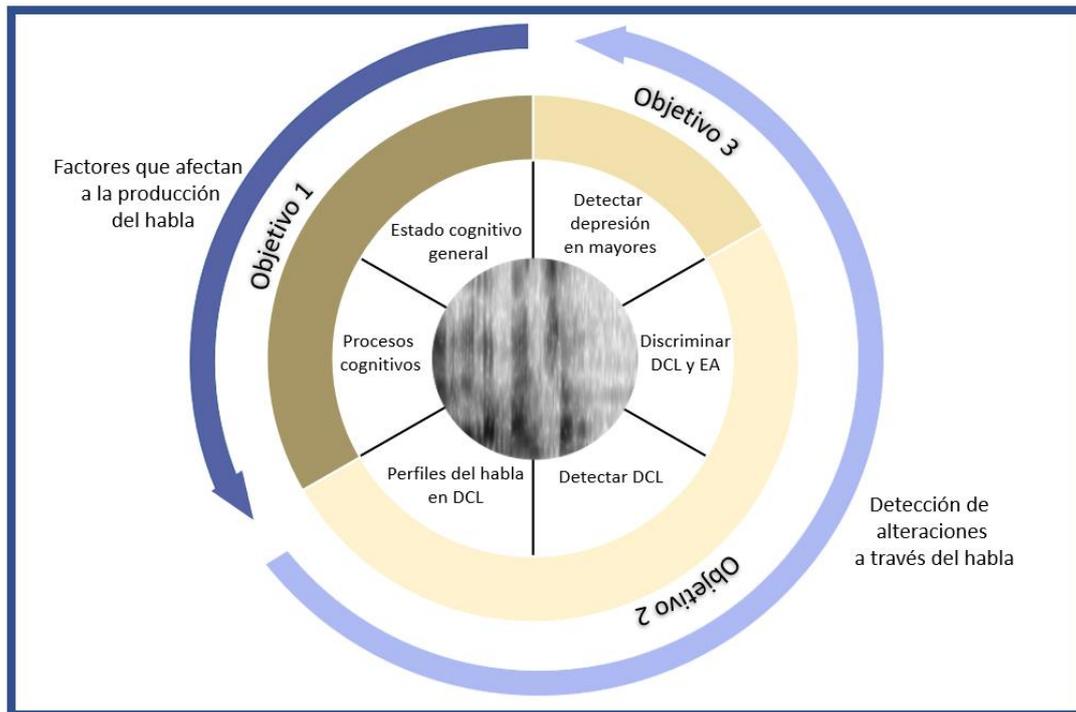
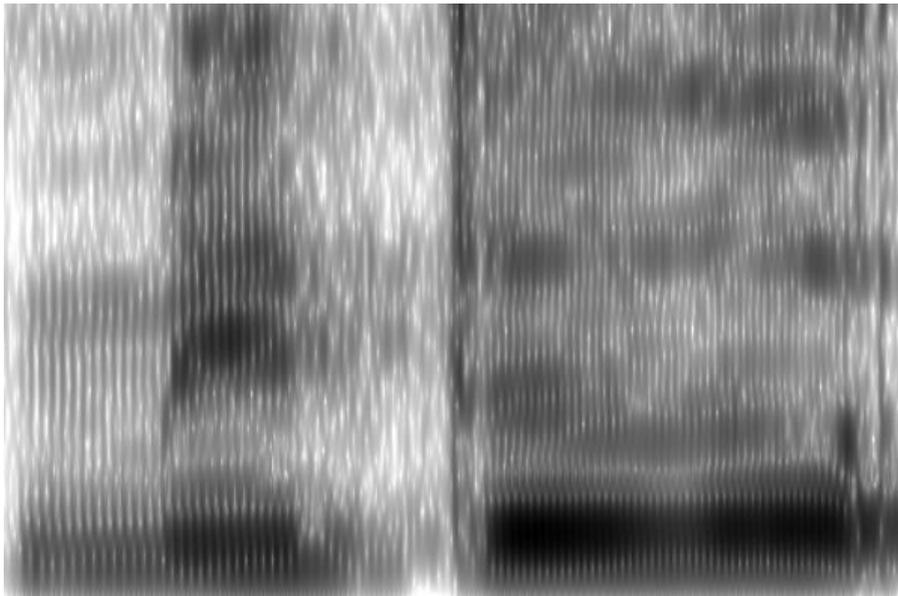


Figura 17. Diagrama de los objetivos de la investigación.

4. Método



4. Método

En este apartado presentamos los aspectos generales del método que serán concretados en cada estudio acorde a sus objetivos específicos. Para lograr el objetivo general de explorar y utilizar parámetros del habla como marcadores de procesos patológicos en mayores, será necesario recoger una muestra de esta población, que además debe ser evaluada para determinar su estado y su rendimiento en diversos dominios cognitivos, así como obtener una grabación de su voz mientras realiza diferentes tareas. En esta sección trataremos el proceso de obtención de la muestra, el procedimiento de su participación en el estudio y las herramientas utilizadas tanto en la evaluación como en el posterior análisis.

El proyecto cuenta con la aprobación del Comité de Ética del Centro de Referencia Estatal de Atención a Personas con Enfermedad de Alzheimer y otras Demencias de Salamanca, perteneciente al Ministerio de Derechos Sociales y Agenda 2030. El estudio se llevó a cabo de acuerdo con la Declaración de Helsinki y sus posteriores enmiendas, y con la normativa de la Unión Europea relativa a la investigación médica.

Participantes

La muestra para los estudios está compuesta por personas mayores de 55 años. Los participantes fueron voluntarios pertenecientes a diversos centros con los que hemos venido realizando convenios de colaboración: Centro de Referencia Estatal de Atención a Personas con Enfermedad de Alzheimer y otras Demencias, AFA Salamanca, Centro de Atención Primaria Miguel Armijo, residencias de mayores colaboradoras de la provincia de Salamanca (Jesán, Usera) y la Unidad de envejecimiento saludable y prevención de problemas cognitivos asociados al proceso de envejecimiento del Servicio

de Atención Psicológica de la Facultad de Psicología (USAL) e integrada ahora en la Unidad Municipal de Apoyo Psicosocial del ayuntamiento de Salamanca. A todos los participantes se les hizo una evaluación neuropsicológica completa a través de la cual se les asignó a los grupos de CS, DCL o EA. Para las personas con DCL se siguieron los criterios del International Working Group on Mild Cognitive Impairment (Winblad et al., 2004) y para las personas con EA los criterios NIA-AA (McKhann et al., 2011). EL diagnóstico de EA fue confirmado siempre por los servicios sanitarios públicos.

Los criterios de inclusión del estudio consistieron en haber firmado el consentimiento informado, ser hablantes nativos de español en su variedad europea y mayores de 55 años. Los criterios de exclusión fueron tener antecedentes personales de enfermedades del sistema nervioso central, abuso de alcohol o sustancias y la presencia de déficits sensoriales graves que pudieran impedir la administración de las pruebas cognitivas.

Materiales

Para las evaluaciones cognitivas se usó fundamentalmente la batería Dem-Detect que forma parte del proyecto NEURONORMA (Peña-Casanova et al., 2009) e incluye los datos normativos de población mayor española para las siguientes pruebas:

Pruebas de medición del estado mental general:

- Mini-Mental State Examination (MMSE) (Folstein, 1975) en su versión española (Blesa et al., 2001): se trata de una prueba de *screening* que se completa en unos 10 minutos y que incluye secciones que evalúan varias funciones cognitivas como la orientación temporal y espacial (2 ítems), la memoria inmediata y demorada (2 ítems), la memoria de trabajo (1 ítem),

el lenguaje (5 ítems) y la praxis (1 ítem). En el cribado de demencia, el MMSE tiene una sensibilidad de 0,87 y especificidad de 0,82.

Pruebas de medición de procesos cognitivos:

- Free and Cued Selective Reminding Test (FCSRT, Buschke, 1984): es una prueba de memoria en la que se muestra al participante un total de 16 palabras con una pista semántica asociada. Tras una prueba distractora se pide al participante que recuerde todas las palabras que pueda (recuerdo libre), tras lo que se proporciona la pista semántica para aquellas palabras que no ha logrado recordar (recuerdo facilitado). Este proceso se repite dos veces más con el objetivo de obtener una curva de aprendizaje. Tras finalizar, se realizan una serie de pruebas distractoras durante un periodo aproximado de 30 minutos y se realiza una última prueba de recuerdo demorado pidiéndole recordar las mismas palabras.
- Memory Impairment Screening (MIS, Buschke et al., 1999): se trata de una versión breve del FCSRT, los participantes tienen que repetir cuatro palabras que escuchan y asignar una clave semántica a cada una de ellas. Tras una tarea distractora, se pide al participante que recuerde las palabras meta. Se otorga puntuaciones dobles para las palabras recordadas sin ninguna pista y puntuaciones simples para palabras recordadas siguiendo una pista semántica.
- Trail Making Test (Reitan, 1986): es una prueba que explora múltiples dominios incluyendo atención, memoria de trabajo, función ejecutiva, velocidad visuomotora y rastreo visual. Consiste en dos partes. En la

primera el participante debe buscar y unir mediante líneas los números por orden de menor a mayor, centrándose en la atención y la coordinación motora. En la segunda, debe alternar los números con las letras en orden alfabético, lo que introduce un factor de flexibilidad.

- Boston Naming Test (Kaplan et al., 2001): es un test de denominación mediante confrontación visual de dibujos en blanco y negro. Si el participante comete un error, se le proporcionan pistas fonológicas o semánticas, de modo que se puede clasificar según los tipos de error.
- Test de Isaacs de fluidez verbal (Isaacs & Akhtar, 1972): consiste en solicitar al participante que diga en un minuto todas las palabras pertenecientes a una de las cuatro categorías siguientes: colores, animales, ciudades y frutas. La puntuación total de cada categoría es de 10 puntos, por lo que la puntuación total máxima es de 40 puntos. Todas las respuestas son anotadas, y los errores y duplicidades se restan de la cuenta final.
- Fluidez fonológica de la “p”: Se pide a los participantes que digan en un minuto todas las palabras que puedan empezar con la letra "p". Los errores, las duplicidades y los nombres propios se excluyen de la puntuación final.

Pruebas de medición del estado emocional:

- Test de Goldberg para depresión y ansiedad (Goldberg et al., 1988): es una prueba breve para determinar la presencia de síntomas depresivos o ansiosos o una combinación de ambos. Tiene dos subescalas de nueve preguntas orientadas a cada una de las condiciones. Las cuatro primeras

preguntas de cada subescala determinan si se debe continuar con el resto. Es decir, si se responde afirmativamente a dos de las cuatro primeras preguntas de la subescala de ansiedad, se debe continuar con el resto, si las respuestas son negativas, se pasa a aplicar la subescala de depresión.

Finalmente, pruebas de evaluación de informadores externos o cuidadores:

- IQCODE (Jorm & Jacomb, 1989): es un cuestionario que recoge la opinión del cuidador sobre los cambios cognitivos observados en el evaluado a lo largo de los últimos 10 años. De este modo se obtiene una valoración del participante con respecto a su propio nivel previo.
- ADFACS (Galasko et al., 1997): es un cuestionario para el cuidador en el que se pide estimar en una escala tipo likert el grado de capacidad del evaluado para realizar actividades instrumentales y actividades básicas de la vida diaria.
- Escala de Blessed (Blessed et al., 1968): se trata de una escala con tres partes valorando los cambios en la ejecución de las actividades de la vida diaria, en los hábitos y cambios conductuales del evaluado en un periodo de seis meses.

Para hacer las grabaciones se utilizó un iPad Air conectado a un micrófono Apogee MiC Plus (ver figura 18), tiene una cápsula con patrón polar cardioide, un rango de frecuencias de 20 Hz-20 kHz y 46 dB de ganancia de preamplificador de micrófono. Las grabaciones fueron realizadas en una sala insonorizada.



Figura 18. Izquierda: iPad Air y micrófono Apogee MiC Plus. Derecha: laboratorio del GIR en Neurofisiología, Cognición y Conducta situado en la Facultad de Psicología de la Universidad de Salamanca.

Para el análisis automático del habla se utilizó el software Praat versión 6.0 (Boersma y Weenink, 2022). Praat, además del análisis del habla, permite el tratamiento estadístico de datos, la edición y la manipulación de las señales de audio. Frente a otros métodos como los MFCCs basados en ventanas o segmentaciones de la onda, praat determina el tono mediante la autocorrelación, es decir, la correlación de la señal con una copia desfasada en el tiempo de sí misma, de forma que la estimación de la frecuencia es más precisa y resistente al ruido que en otros métodos. Además, se siguieron las recomendaciones de los autores del programa en cuanto al análisis de los rangos de F0 y se estableció un mínimo de 75 Hz para y un máximo de 300 para los hombres, y un mínimo de 100 Hz y máximo de 500 para las mujeres.

Procedimiento

En líneas generales se realizaron tres sesiones de evaluación neuropsicológica con cada participante. Las sesiones incluyeron una anamnesis completa, la evaluación de las actividades de la vida diaria y una evaluación cognitiva y del estado afectivo. La primera

sesión incluía la anamnesis, las pruebas de cribado y la grabación de voz. En las sesiones siguientes se realizaron pruebas más específicas y, en su caso, se incorporaron otras pruebas acordes a los objetivos de los estudios. Al finalizar las sesiones de evaluación, se entregó un informe a los participantes detallando los resultados y recomendaciones oportunas.

El procedimiento elegido incluye de forma transversal en la mayoría de los estudios llevados a cabo en esta tesis una tarea de lectura. Esta se basa en leer en voz alta las dos primeras frases de la novela de Miguel de Cervantes “El Ingenioso Hidalgo Don Quijote de la Mancha”:

“En un lugar de la Mancha, de cuyo nombre no quiero acordarme, no ha mucho tiempo que vivía un hidalgo de los de lanza en astillero, adarga antigua, rocín flaco y galgo corredor. Una olla de algo más vaca que carnero, salpicón las más noches, duelos y quebrantos los sábados, lantejas los viernes, algún palomino de añadidura los domingos, consumían las tres partes de su hacienda.”

Este texto ha sido seleccionado porque reúne varias características. Por un lado, muchos hablantes podrían reproducirlo de memoria, especialmente el comienzo del texto, dado que forma parte del conocimiento enciclopédico de la sociedad española europea. Por otro lado, el texto combina términos de alta frecuencia (tiempo, vaca, viernes) y de baja frecuencia (hidalgo, galgo, carnero) y forman enunciados sintácticamente complejos que incluyen proposiciones relativas y sujetos y objetos complejos. De este modo, el texto nos permite controlar la carga semántica, lo que es importante para el control y la supervisión del habla. El texto incluye elementos muy familiares y elementos muy poco familiares, comprometiendo la fluidez del hablante. Estudios anteriores sobre lectura en EA sugieren que no se ve afectada cuando se enfrentan a palabras de alta frecuencia, pero

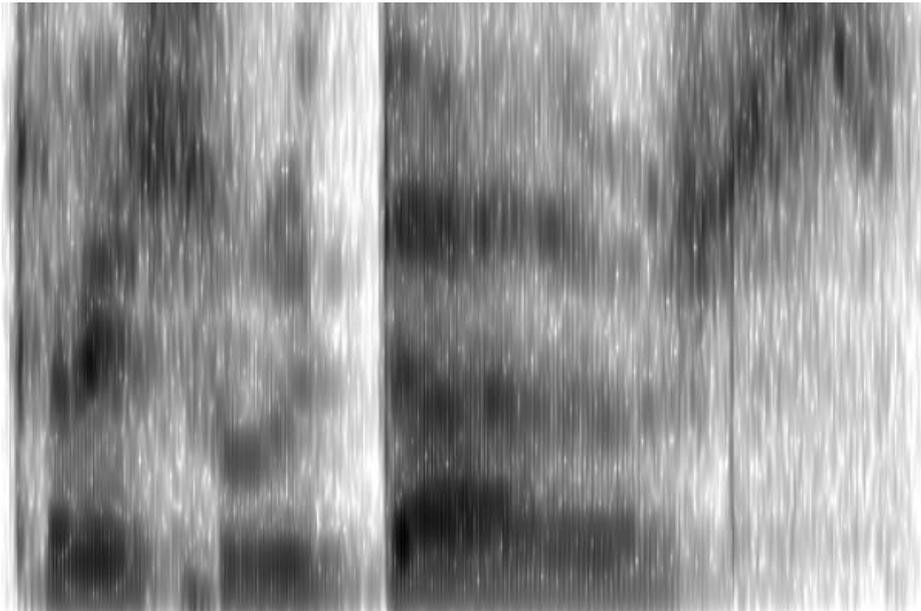
sí lo hace para las palabras con baja frecuencia (Benigas y Bourgeois 2011). Teniendo en cuenta que la conversión de DCL a EA está relacionada con la ralentización de la identificación de palabras durante la lectura (Massoud et al., 2002), es de suponer que la combinación de palabras conocidas y menos conocidas en un mismo texto puede convertirlo en una herramienta muy útil para poner de manifiesto las alteraciones léxicas y semánticas de la neurodegeneración.

Los participantes leen el texto en tamaño de 48 puntos desde una pantalla en una sala insonorizada. El micrófono se coloca a una distancia de entre 8 y 14 cm de la boca del participante con un ángulo aproximado de 45° para minimizar el ruido de la respiración.

Análisis estadístico

Todos los análisis estadísticos se realizaron con IBM SPSS Statistics for Windows, versión 25.0.

5. Resultados



Estudio I

Algunos cambios en el habla de los mayores se deben en realidad a cambios en su estado cognitivo

En este estudio indagaremos en las causas de los cambios en el habla de los mayores, argumentaremos qué diversos factores pueden influir en cómo se producen los cambios e intentaremos observar la evolución de algunos de ellos. En este sentido, los parámetros de habla habitualmente se entienden como marcadores de una alteración, sin tener en cuenta las características propias del hablante. La evolución de los dominios cognitivos en los adultos mayores puede ser muy variable, y se necesitan datos normativos para relacionar su rendimiento con su grupo de referencia (Strauss et al., 2006). Planteamos entonces que sería de interés establecer cómo interaccionan la edad, el género y el estado cognitivo en la producción de los parámetros de habla.

La edad provoca una serie de cambios en el habla relacionados con cuestiones fisiológicas. Ya se ha hecho mención anteriormente a la conocida como presbifonía, es decir, las características clínicas de la voz de los mayores, que hacen referencia a cambios que se deben a alteraciones anatómicas y fisiológicas en el tracto vocal y en la curvatura de la laringe, y que provocan dificultades en el control de los parámetros acústicos (Alonso et al., 2001; Hughes & Rhodes, 2010). Por lo general, se produce una reducción del rango vocal, una disminución de la frecuencia fundamental (F0) en las voces femeninas (de niveles medios en torno a 248Hz a 175Hz) y un aumento en las masculinas (de 110Hz a 135-160Hz), una mayor variación de las frecuencias (jitter) y de la amplitud en decibelios (shimmer), una reducción de la resonancia y un aumento de las pausas del habla (Linville, 2004).

Otra serie de cambios en el lenguaje expresivo de los mayores podrían deberse al rendimiento cognitivo. En este sentido cabe mencionar las dificultades para encontrar palabras, ya que son un síntoma especialmente relevante en las primeras fases de la EA. Sin embargo, las dificultades para encontrar palabras también se han descrito en otros grupos de personas mayores, tanto sanas como con DCL, como consecuencia del envejecimiento (Kavé et al., 2011; Zellner-Keller, 2007). A modo de ejemplo de cambios en el habla, los mayores sanos muestran una mayor frecuencia y duración de las pausas del habla, junto con la consiguiente disminución de la velocidad del habla (Schötz, 2005). Estos cambios, además, se interpretan como el resultado de las dificultades de acceso al léxico (Zellner-Keller, 2007) tal y como se interpretan en procesos neurodegenerativos. Algunos de estos parámetros han sido propuestos como indicadores de la edad del hablante. Es el caso de la velocidad del habla (Skoog Waller et al., 2015), los parámetros de calidad de voz y otras variables derivadas del análisis espectral, como el centro de gravedad y el LTAS. En el caso de los pacientes con DCL y EA, se ha encontrado que varias de estas características del habla están alteradas y se correlacionan con el rendimiento en medidas cognitivas (Pulido et al., 2020).

El beneficio de examinar específicamente el lenguaje oral y su relación con el estado cognitivo general de una persona parece estar respaldado por varios estudios que relacionan directamente el deterioro del lenguaje con el deterioro general (Ahmed et al., 2013; Luz et al., 2021; McKhann et al., 2011). Se sabe que los trastornos del habla y del lenguaje en los adultos mayores correlacionan con la gravedad del deterioro cognitivo (Kauhanen et al., 2000; Marinelli et al., 2017). Por ejemplo, los trastornos afásicos se relacionan con un mayor riesgo de desarrollar demencia, una mayor gravedad del deterioro o una progresión más rápida del mismo (O'Hara et al., 2002).

Una de las herramientas más utilizadas como referencia del estado cognitivo de los mayores es el MMSE. Es la prueba de cribado más popular, principalmente por su facilidad de uso (Ismail et al., 2010). Aunque el MMSE se ha utilizado para el cribado de la EA, ha sido más eficaz en el seguimiento de los cambios de grado y gravedad del estado cognitivo. Sin embargo, no está exento de críticas, ya que algunos autores destacan que muestra una menor sensibilidad a la hora de monitorizar los cambios en pacientes con demencia severa (Lancu & Olmer, 2006). o a la hora de diferenciar entre DCL y los estadios iniciales de la demencia (Delavaran et al., 2017). De hecho, la fase prodrómica o preclínica de la EA es conceptualmente similar tanto al DCL como al envejecimiento sano de las personas mayores, que muestran déficits relacionados con la edad sin afectación de las actividades de la vida diaria. Todos ellos suelen mostrar déficits en la memoria episódica, así como en la atención, el razonamiento, la percepción temporal y espacial y el lenguaje. Por ello se hace necesario establecer el grupo de referencia del paciente y diferenciar los cambios debidos a alteraciones de la cognición de aquellos debidos a la edad.

Teniendo en cuenta estos antecedentes, el objetivo principal de este estudio es explorar la relación entre los cambios en el lenguaje expresivo asociados a la edad y el grado de deterioro cognitivo relacionado con estos déficits. Por tanto, como objetivos específicos planteamos en primer lugar comprobar cómo cambian los parámetros del habla con la edad. En segundo lugar, comparar los parámetros de habla y su relación con tres grados de deterioro cognitivo establecido mediante el MMSE. Así, una de las principales novedades de este estudio frente a otros que han abordado el papel del habla en relación con la clasificación binaria de demencia frente a no demencia, aquí pretendemos evaluar el grado en que los parámetros del habla están relacionados con la gravedad del deterioro cognitivo. Para ello, seleccionaremos una serie de parámetros

temporales, de amplitud y de frecuencia habitualmente relacionados con el envejecimiento. Dado que algunos parámetros se alteran con el deterioro cognitivo, indagaremos en cuáles de ellos están vinculados en mayor medida a un deterioro biológico y, por tanto, dependen de la edad, y cuáles son el resultado de una disminución de los recursos cognitivos. Asimismo, exploraremos si la variable de género también influye en la evolución de alguno de estos parámetros.

Método

Participantes

La muestra de este estudio incluye 400 participantes mayores de 55 años. Para esta muestra tan numerosa se ha recopilado a todos aquellos participantes que el GIR en Neurofisiología, Cognición y Conducta ha obtenido a lo largo de los años, constituyendo un fin de ciclo que ha culminado posteriormente en la puesta a disposición libremente de esta base de datos para otros investigadores como parte del Dementia TalkBank (Ivanova et al., 2022). Esta muestra contiene participantes clasificados como CS (52,3%, MMSE media=28,24, dt=2,13), sujetos con DCL (24%, MMSE media= 23,63, dt=4,50) o EA (23,8%, MMSE media=20,22, dt=5,87). En este estudio, a los criterios de exclusión generales, se añadió el de la presencia de enfermedades psiquiátricas. Se excluyeron aquellas personas que obtuvieron una puntuación de >6 en el test Goldberg y que, por tanto, presentaban una depresión o ansiedad severa. Esto se debe a que estudios recientes indican que los síntomas neuropsiquiátricos pueden influir en el habla de los adultos mayores (König et al., 2021), en particular, la depresión y la ansiedad tienen sus propias características acústicas y prosódicas que pueden distorsionar los resultados (Lopez-Otero & Docio-Fernandez, 2021; Morales & Levitan, 2016).

Se dividió a los participantes en tres grupos en función de su puntuación en el MMSE, utilizando como puntos de corte los obtenidos en la revisión de Arevalo-Rodriguez et al., (2015). Así, los grupos fueron: el 1º rango, personas con menos de 23 puntos (deterioro cognitivo moderado-grave); el 2º rango, personas entre 23 y 27 puntos (deterioro leve); y el 3º rango, personas con más de 27 puntos (sin deterioro cognitivo). Las características de cada rango se recogen en la Tabla 4.

En el estudio participaron un mayor número de mujeres que de hombres (70,5% frente a 29,5%), pero no se encontraron diferencias en su distribución en los rangos del MMSE.

Para explorar el efecto de la edad, se clasificó a los participantes en cuatro intervalos: menos de 70 años, entre 70 y 79, entre 80 y 84, y más de 84 años. El criterio para establecer los rangos fue el de obtener grupos homogéneos en cuanto al número de participantes en cada uno de ellos.

Un ANOVA mostró diferencias de edad entre los grupos de estado cognitivo ($F_{2, 397} = 24,084$, $p < .001$), siendo las personas del rango 1º <23 significativamente mayores que las del rango 3º 23-27 ($p < .001$). Las personas en el rango 2º 23-27 eran también significativamente mayores que las del rango 3º >27 ($p < 0,001$). En otro ANOVA se encontró diferencias significativas entre los rangos en cuanto a los años de escolaridad ($F_{2, 385} = 11,388$, $p < 0,001$). Estas se encontraban entre los rangos 1º <23 y 2º 23-27 ($p < .05$) y 3º >27 ($p = .001$) de las puntuaciones en el MMSE.

Tabla 4. Características de la muestra.

	N	Edad	Años escolaridad	MMSE
		Media (dt)	Media (dt)	Media (dt)
1° rango MMSE <23	97	81,37 (8,60)	7,66 (3,59)	17,68 (4,37)
2° rango MMSE 23-27	115	78,80 (8,04)	9,00 (3,98)	25,23 (1,44)
3° rango MMSE >27	188	74,29 (8,93)	10,03 (4,05)	29,12 (,96)
Total	400	77,31 (9,09)	9,16 (4,03)	25,23 (5,16)
Hombres	118	76,60 (9,31)		
Mujeres	282	77,60 (8,99)		

Procedimiento

Los participantes completaron una tarea de lectura que consistía en leer en voz alta el primer párrafo de El Ingenioso hidalgo Don Quijote de la Mancha.

Se han seleccionado 14 parámetros, que debido a la revisión que habíamos realizado considerábamos que podrían estar relacionados con el deterioro cognitivo, y que también han sido señalados como marcadores de la edad (Dehqan et al., 2012; Linville, 2004; Mazzetto de Menezes et al., 2014). Estos fueron: Duración total del tiempo de emisión de habla, número de pausas en el habla, velocidad de habla medido por el número de fonemas por unidad de tiempo, duración media de las sílabas, desviación típica de la duración de las sílabas, nPVI, intensidad media de habla, LTAS, LTAS 50-1k (LTAS entre frecuencias de 50 y 1000), F0, desviación típica de F1, asimetría, HNR y jitter local. Según nuestra hipótesis, estas variables podrían afectar a determinados parámetros del habla tanto durante el envejecimiento no patológico como el patológico.

Análisis estadístico

Para el análisis se utilizó un ANOVA con un diseño factorial 2x3x4 que incluía el género, el estado cognitivo y la edad. Este ANOVA se realizó sobre todas las variables hipotetizadas. A continuación, desglosamos los resultados mediante comparaciones post hoc aplicando la corrección de Bonferroni.

Resultados

Los resultados de todos los análisis de varianza de los parámetros se encuentran en la Tabla 5.

Podemos identificar tres tendencias diferentes dentro de los resultados: parámetros del habla que varían dependiendo del estado cognitivo del hablante; parámetros del habla que dependen de la edad del hablante, y parámetros que difieren según el género. Por ello, hemos organizado los resultados de esta forma.

Parámetros del habla que cambian en función del estado cognitivo:

Este grupo de parámetros del habla incluye variables que mostraron diferencias significativas entre los participantes debido al efecto principal del rango de puntuación en el MMSE. No hubo efectos de interacción con la edad, ni efectos principales debidos a la edad. Sin embargo, en algunos casos hubo interacciones con el género.

Tabla 5. Influencia del género, el estado cognitivo (MMSE) y las categorías de edad en los parámetros del habla implicados en la predicción de la EA.

Parámetros	Género	Interacción MMSE-edad	MMSE	Edad
	F _{1,376}	F _{6,376}	F _{2,376}	F _{3,376}
Duración total	0.142	0.620	28.140***	1.244
Número de pausas	0.153	0.439	28.477***	0.959
Velocidad de habla	10.845*	0.353	16.835***	1.725
Duración media de sílabas	0.254	2.271*	3.886*	1.646
Dt de duración de sílabas	2.058	0.936	10.917***	0.315
nPVI	0.622	0.356	6.408***	1.973
Intensidad media	1.759	0.560	8.202**	0.624
LTAS	0.005	0.581	5.103**	2.616
LTAS 50-1K	1.638	0.869	6.217**	1.134
F1 dt	11.507**	2.130*	7.253**	.562
F0	170.203***	0.689	2.654	0.992
Asimetría	5.649*	1.116	0.932	0.582
HNR	27.830***	1.011	1.137	2.877*
Jitter (Local)	27.740***	0.968	0.805	3.427*

* p < .05; ** p < .01; *** p < .001

Dentro de este primer grupo, se encuentran parámetros temporales como la duración y el número de pausas. Se encontraron diferencias significativas en la duración total entre los participantes con MMSE <23 (70seg.) y el resto de las categorías (MMSE = 23-27, dif: -24seg., p <,001; y MMSE >27, dif: -30seg., p <,001). Véase la tabla 6, para un resumen de las pruebas post hoc en parámetros con efectos de estado cognitivo o edad.

Tabla 6. Comparaciones por pares de parámetros con un efecto principal de la edad o el estado cognitivo. * es significativo a $p < 0,05$ y ** es significativo a $p < 0,001$.

	Duración	Número de pausas	Velocidad de habla	Duración media de la sílaba	Dt de la duración de la sílaba	nPVI	Intensidad media	LTAS	LTAS 50-1k	F1 dt	HNR	Jitter local
<70 vs 70-79	-7,997	-5,032	0,245	0,006	0,001	-2,301	0,587	0,551	0,815	-5,576	0,974	-0,407
<70 vs 80-84	-9,087	-3,567	0,138	0,007	-0,001	-2,827	0,229	-0,542	0,266	-12,4	1,279	-0,41
<70 vs >85	-6,883	-2,407	0,008	0,011	0,000	-2,472	0,26	-0,578	0,487	10,858	1,711*	-0,493*
70-79 vs 80-84	-1,09	1,465	-0,106	0,001	-0,003	-0,527	-0,358	-1,093	-0,549	-6,824	0,305	-0,002
70-79 vs >85	1,114	2,625	-0,237	0,005	-0,001	-0,171	-0,326	-1,129	-0,328	16,434	0,736	-0,086
80-84 vs >85	2,204	1,160	-0,131	0,005	0,001	0,355	0,032	-0,036	0,221	23,258	0,431	-0,083
<23 vs 23-27	23,590**	14,876**	-0,496**	0,008	0,007	2,374	-1,089*	-1,156	-0,979	54,372*	0,371	0,042
<23 vs >27	29,988**	19,28**	-0,685**	0,012*	0,013**	3,533*	-1,558**	-1,502*	-1,426*	69,948*	0,704	-0,088
23-27 vs >27	6,398	4,304	-0,189	0,004	0,006*	1,16	-0,469	-0,347	-0,448		0,333	-0,13

En cuanto al número de pausas, se observaron diferencias entre los participantes con MMSE <23 (39 pausas) y el resto de las categorías (MMSE = 23-27, dif: -15, $p < ,001$; y >27, dif: -19, $p < ,001$). Así, un estado cognitivo de <23, correspondiente a un deterioro cognitivo moderado, implica un aumento significativo de la duración, así como del número de pausas en el habla (véase la Figura 19 como ejemplo representativo de una representación visual de este grupo de parámetros).

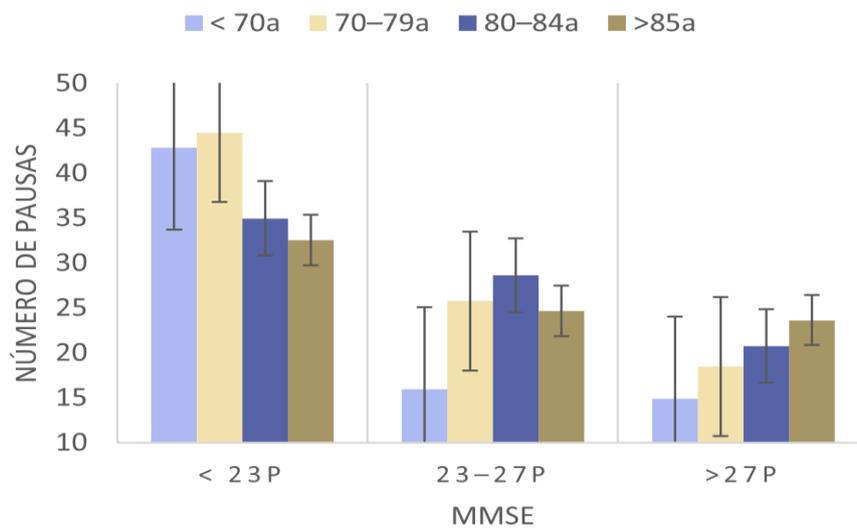


Figura 19. Evolución del número de pausas según el estado cognitivo medido por el MMSE y los rangos de edad.

Los resultados fueron similares para los parámetros de fluidez y ritmo, como la velocidad del habla, la duración media de las sílabas, la desviación típica de las sílabas y el nPVI. En la velocidad del habla (ver Figura 20), medida en fonemas por segundo, hubo un efecto del género, siendo las mujeres (3,42) más rápidas que los hombres (3,11) (dif= ,306, $p < ,01$). Sin embargo, no hay interacción del género con el MMSE ni con la edad.

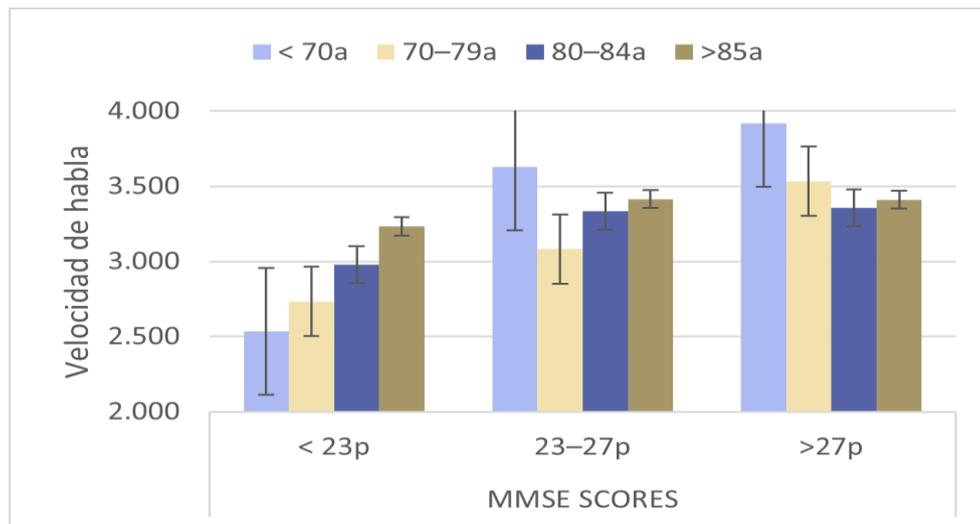


Figura 20. Evolución de la velocidad de habla según el estado cognitivo medido por el MMSE y los rangos de edad.

Fue significativo el efecto principal del MMSE que mostró diferencias entre los participantes con MMSE <23 (2,87) y el resto (MMSE = 23-27, dif: -,496, $p < ,001$; y >27, dif: -,685, $p < ,001$). Así, un estado cognitivo de <23 implica una disminución significativa de la velocidad de habla. Se encontró un efecto de interacción entre el MMSE y la edad ($F_{6,376} = 2,385$, $p < ,05$). Las diferencias en el MMSE se encuentran en los menores de 70 años y en los MMSE>27.

También encontramos cambios en varios parámetros del ritmo silábico. La duración media de la sílaba mostró un efecto del estado cognitivo. Se encontraron diferencias entre los participantes con MMSE<23 (,20ms) y aquellos con MMSE>27 (dif: ,02, $p < ,05$). Por tanto, la duración media de las sílabas es mayor en las personas con deterioro cognitivo moderado. Se encontró un efecto de interacción significativo entre el estado cognitivo y la edad ($F_{6,376} = 2,467$, $p < ,05$). Las diferencias entre rangos de MMSE aparecen sólo en los participantes menores de 70 años. En este grupo de edad hay diferencias entre MMSE<23 (,23s) y tanto MMSE 23-27 (dif = ,04s, $p < ,01$) como MMSE>27 (dif = ,04s, $p < ,001$).

En este parámetro encontramos también un efecto de interacción entre el estado cognitivo, la Edad y el Género ($F_{6,376} = 2,271$, $p < ,05$). En los hombres sólo hay diferencias en la duración de la sílaba ($F_{3,376} = 2,891$) en $MMSE < 23$, mientras que en las mujeres hay diferencias en $MMSE < 23$ ($F_{3,376} = 2,704$, $p < ,05$) y $MMSE 23-27$ ($F_{3,376} = 2,683$, $p < ,05$). La figura 21 muestra la distribución de los parámetros de la sílaba en diferentes rangos de estado cognitivo para hombres y mujeres.

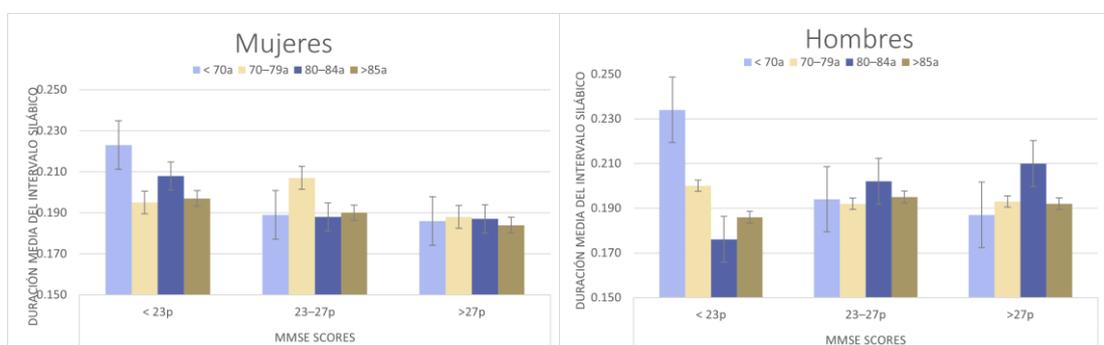


Figura 21. Evolución de la duración media de la sílaba en hombres y mujeres según el estado cognitivo medido por el MMSE y los rangos de edad.

La desviación típica de la duración de las sílabas es un parámetro que analiza la variabilidad media de las distancias entre sílabas. Se encontraron diferencias entre los participantes con $MMSE < 23$ (,11ms) y $MMSE > 27$ (dif: ,01 $p < ,001$). También hubo diferencias entre los participantes con $MMSE = 23-27$ (,012) y $MMSE > 27$ (dif: ,006 $p < ,05$). Por lo tanto, el ritmo de producción de las sílabas muestra una mayor variabilidad cuando el estado cognitivo es típico del deterioro cognitivo leve y moderado.

El nPVI es un parámetro que analiza la variabilidad de las distancias entre sílabas. Un valor alto de nPVI corresponde a una mayor variabilidad rítmica. En este parámetro, se observaron diferencias entre los participantes con MMSE<23 (58,128ms) y aquellos con MMSE>27 (dif: 3,53, p <0,001). Esto significa que la duración media de las sílabas tiene un patrón rítmico significativamente más irregular cuando el estado cognitivo corresponde a un deterioro cognitivo. La figura 22 muestra la distribución del nPVI en diferentes rangos de estado cognitivo. Este patrón es similar al de otros parámetros de alteración del ritmo silábico, como la duración media de las sílabas, y la desviación típica de la duración de las sílabas.

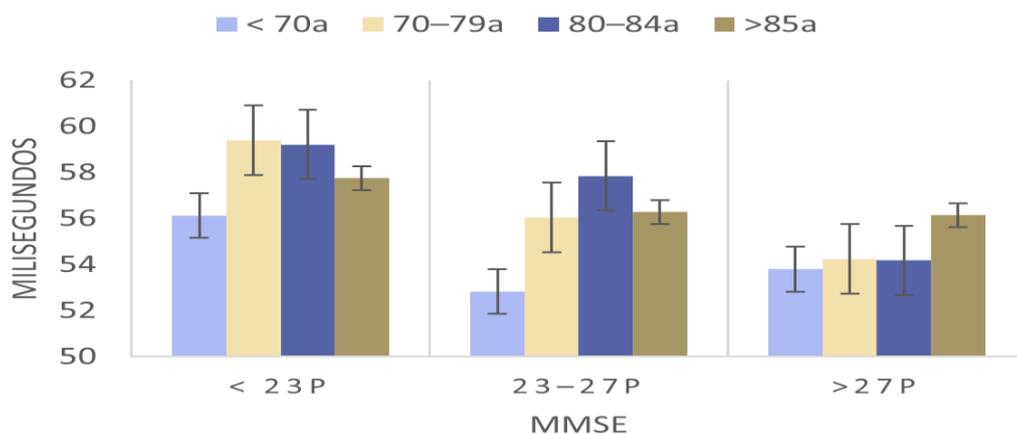


Figura 22. Evolución del nPVI según el estado cognitivo medido por el MMSE y los rangos de edad.

Los parámetros de amplitud y calidad de la voz, como la amplitud media, el LTAS y el LTAS 50-1K, también mostraron diferencias significativas. El parámetro de amplitud media refleja la energía que el hablante da al enunciado. Se encontraron diferencias entre los participantes con MMSE<23 (73,48dB) y los de las otras categorías, MMSE 23-27 (dif: -1,089, p < 0,05) y MMSE >27 (dif: -1,558, p < 0,001). Por lo tanto, la intensidad media de la voz es menor cuando el estado cognitivo se corresponde con el deterioro cognitivo. Véase el ejemplo de la Figura 23.

El LTAS es un parámetro de calidad de la voz que correlaciona el ajuste en tiempo de las frecuencias de los formantes con su intensidad adecuada. Este parámetro mostró diferencias significativas entre los participantes con MMSE<23 (30,674) y MMSE>27 (dif: -1,502, $p < 0,01$). Por tanto, la energía espectral es menor cuando el estado cognitivo es de deterioro cognitivo. El parámetro LTAS 50-1K es una medida de LTAS centrada en el espectro de los primeros 1000Hz (de 50 a 1000Hz). Las diferencias se produjeron entre los participantes con MMSE<23 (43,12dB) y MMSE>27 (dif: -1,426, $p < ,01$). En este sentido, la energía del espectro en el rango de 1000 Hertz es menor cuando el estado cognitivo corresponde con un deterioro cognitivo.

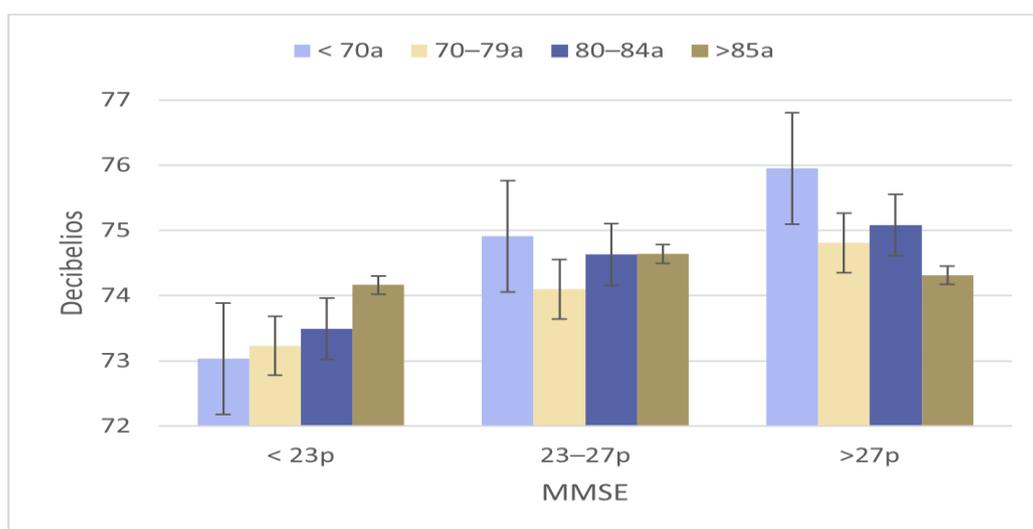


Figura 23. Evolución de la intensidad media según el estado cognitivo medido por el MMSE y los rangos de edad.

Por último, la desviación estándar del Formante 1 (F1sd) depende del estado cognitivo, pero también del sexo y la edad. Por un lado, se encontraron diferencias significativas entre los participantes con MMSE<23 (440hz), y tanto MMSE 23-27 (dif: 54,32hz; $p < ,05$) como MMSE >27 (dif: 69,948hz, $p < ,01$). No se encontró interacción entre las variables de estado cognitivo y edad. Sin embargo, encontramos diferencias debidas al género entre hombres (423hz) y mujeres (374hz, $p < ,01$). Se encontró una interacción cuando se comparan las tres variables juntas: edad, género y estado cognitivo. Las diferencias se encontraron en las mujeres con un estado cognitivo entre 23 y 27 puntos

en el MMSE, y menores de 70 años con respecto a otros grupos. También hubo diferencias entre los hombres y las mujeres menores de 79 años y con un MMSE <27.

Parámetros del habla que cambian en función de la edad:

En esta sección se analizan los parámetros del habla que mostraron diferencias significativas entre los participantes debido al efecto principal de la edad, que son la relación entre armónicos y ruido (HNR) y el jitter (local). No hubo efectos de interacción de la edad con el estado cognitivo y el género, aunque, en ambos casos, se encontró un efecto principal del género.

El HNR mide la proporción de energía de los componentes armónicos de la voz (sonido periódico) con respecto al ruido que contiene (aperiódico). Se hallaron diferencias debidas al género, siendo la amplitud de los armónicos menor en los varones (10,988dB frente a 12,994dB). Hubo diferencias debidas a la edad, siendo la amplitud de los armónicos menor en los >85 años que en los <70 años (11,274dB frente a 12,982dB, respectivamente). No se encontró ningún efecto de interacción entre las variables de edad y género.

El jitter (local) mide el porcentaje de variación de la frecuencia media entre dos períodos de voz consecutivos o pulsos de las cuerdas vocales. De nuevo, encontramos diferencias por género, siendo la variabilidad mayor en los hombres (2,936% frente a 2,400%). Hubo diferencias por edad, siendo el Jitter local menor ($p < .05$) en >85 años que en <70 años (2,340% frente a 2,834%, respectivamente). Sin embargo, no se encontró un efecto de interacción. Así, a mayor edad, mayor es el porcentaje de variabilidad en las frecuencias.

Parámetros que cambian en función del sexo:

Estos parámetros del habla no mostraron efectos principales del MMSE ni de la edad; tampoco encontramos efectos de interacción entre las variables. Son parámetros basados en la frecuencia y dependen únicamente del sexo del hablante, por lo que están estrechamente relacionados con aspectos biológicos. Es el caso de la F0 media, ya que los hombres mayores (138Hz) mantienen un tono medio más bajo que las mujeres (180Hz); y de la asimetría espectral, donde el centro de gravedad de la frecuencia es más positivo hacia las frecuencias bajas en los hombres (10,199Hz) que en las mujeres (8,389Hz).

Discusión

En este primer estudio, hemos explorado una amplia gama de parámetros del habla y tres factores que pueden determinar cómo cambian en los adultos mayores: la edad, el estado cognitivo y el género. Muchos de los parámetros parecen estar determinados por el estado cognitivo del hablante medido por el MMSE. Por lo tanto, la evolución neurodegenerativa relacionada con la edad hacia la demencia afectará directamente a los procesos de producción del habla relacionados con estas variables. El habla no depende de un único proceso, sino que requiere de una amplia gama de factores. Cualquier alteración relacionada con la edad en dichos procesos afectará, en consecuencia, a las propiedades y a la producción del habla. Teniendo esto en cuenta, en este estudio se ha analizado qué variables del habla son de carácter cognitivo y están, por tanto, determinadas por el estado mental del adulto mayor, y cuáles de ellas dependen de otros factores biológicos como la edad o el sexo. Entender cómo producen el habla los adultos mayores y cómo puede verse afectada por el estado cognitivo, la edad y el género, es muy

relevante tanto para monitorizar los cambios del habla a lo largo del proceso de envejecimiento, como para establecer su uso como marcadores clínicos para el diagnóstico de enfermedades neurodegenerativas.

En primer lugar, hemos identificado varios parámetros del habla cuya expresión varía en función de las puntuaciones obtenidas por los mayores en el MMSE. Así, cuanto mayor es el deterioro cognitivo en los adultos mayores medido por esta prueba, mayor es el tiempo de fonación, se producen más pausas en el habla, se reduce la velocidad del habla y mayor es la duración media y la variabilidad de las sílabas. Presumiblemente, estas alteraciones reflejarían dificultades en la búsqueda léxica y de las funciones ejecutivas.

Durante la producción del habla, las pausas vacías son más frecuentes y largas cerca de los límites sintácticos de las frases o cuando se pronuncian frases basadas en palabras de baja frecuencia (Lee et al., 2011). Así, la producción del habla está condicionada por la presión de tener que planificar sintácticamente las frases mientras se producen dos o tres sílabas por segundo, lo que en una persona afectada por un proceso neurodegenerativo podría traducirse en una alta carga cognitiva.

Otro factor que puede determinar la importancia de los aspectos silábicos en la producción del habla relacionada con la edad es el hecho de que el español es una lengua de ritmo silábico (Arvaniti & Rodríguez, 2013), lo que significa que cada sílaba tarda aproximadamente el mismo tiempo en ser pronunciada independientemente de que esté acentuada o no. Por lo tanto, el ritmo en la lengua española es monótono, y el habla no patológica se define por la regularidad en la producción de las sílabas. Nuestros datos muestran que dicha regularidad está alterada en el deterioro cognitivo moderado.

Del mismo modo, la intensidad y los parámetros de calidad de la voz como la amplitud media, la LTAS y la LTAS 50-1K, también parecen depender del estado cognitivo del mayor. Hemos encontrado una disminución significativa de la intensidad media asociada al deterioro cognitivo. Sugerimos que esto podría ser una explicación factible a la disminución de la amplitud relacionada con la edad observada en otros estudios (Mazzetto de Menezes et al., 2014). Hay mucha literatura que sugiere un debilitamiento de la voz con la edad, pero se sabe poco sobre la razón de ello (Lortie et al., 2015). Se ha propuesto que la disminución de la intensidad podría deberse a un esfuerzo muscular y fonatorio compensatorio (Huber & Spruill, 2008). El LTAS se relacionaría con el esfuerzo vocal (Núñez Batalla et al., 2004) implicado en los ajustes de cavidades supraglóticas, y se relaciona significativamente con la gravedad de una disfonía (Lowell et al., 2011). Así, valores más altos de este parámetro se relacionan con una voz áspera y ronca, con pobreza de elementos armónicos en las frecuencias altas; mientras que valores bajos, asociados a la demencia, muestran un bajo esfuerzo vocal con inadecuación armónica en las frecuencias bajas.

Son de destacar las similitudes en la evolución de estos parámetros en función del estado cognitivo entre los grupos. La duración total, las pausas, la velocidad del habla, la amplitud y la desviación del F1 no muestran diferencias entre los rangos 23-27 y <27. nPVI, LTAS y LTAS 50-1K sólo muestran diferencias entre <23 y <27. Estos parámetros evolucionarían lentamente y muestran cambios muy leves antes de volverse significativos en la fase de deterioro cognitivo moderado. Sin embargo, la duración media la desviación típica de las sílabas sí diferencian a las personas mayores con un deterioro leve de aquellas con un envejecimiento cognitivo sano.

Por lo tanto, el estado cognitivo global de los adultos mayores parece estar estrechamente relacionado con la producción del habla, que se muestra como una medida

conductual útil de la capacidad de procesamiento. Las dificultades para acceder al léxico y a las estructuras silábicas, impulsadas por los deterioros cognitivos generales, ralentizarían la velocidad del habla y afectarían al ritmo y a la prosodia (Schötz, 2005; Zellner-Keller, 2007). Queda por verificar qué procesos específicos que cambian en las primeras etapas de la demencia son los que se pueden identificar a través del habla, lo que será importante para desarrollar procedimientos de diagnóstico más eficaces.

Los parámetros del habla modulados por la edad como factor biológico son rasgos de alteración vocal, como el HNR y el Jitter local, relacionados con la pérdida de calidad del habla (Mazzetto de Menezes et al., 2014). Confirmaríamos las ideas de la literatura, de modo que el HNR es menor en los hombres y también disminuye con la edad, provocando la sensación de una voz áspera. Tanto en los hombres como en las mujeres, hemos encontrado un HNR con valores típicamente asociados a una voz deteriorada. El jitter local fue mayor en los hombres que en las mujeres y, como otros estudios (Dehqan et al., 2012), hemos observado que aumenta con la edad. Estas alteraciones suelen atribuirse a déficits en el control neurológico de los músculos, así como a cambios mecánicos debidos a la atrofia de las cuerdas vocales.

Por último, algunos de los parámetros del habla que según otros estudios se han relacionado con la EA (Fraser et al., 2019; Themistocleous et al., 2018), parecen depender únicamente del género. Este es el caso de algunos parámetros de frecuencias y análisis espectral, como la F0 y la Asimetría. Los hombres mostraron un tono más bajo y una tendencia más pronunciada hacia las frecuencias bajas que las mujeres. Este resultado no coincide con la mayoría de los estudios (Dehqan et al., 2012; Hughes & Rhodes, 2010), que sostienen que existe una disminución de la F0 relacionada con la edad en las mujeres, mientras que en los hombres aumentaría. Habitualmente se considera que la F0 está determinada por el género y la edad debido a su dependencia directa de la masa y la

longitud de las cuerdas vocales. Sin embargo, tampoco hemos encontrado cambios significativos en este sentido. En cuanto a la asimetría, también está determinada por el deterioro físico de los ligamentos y el cartílago, y por una curvatura más pronunciada de la laringe en los márgenes, lo que provoca un cambio en los parámetros de frecuencia después de los 55 años. Sin embargo, según nuestros datos, no encontramos ningún cambio en este parámetro debido al envejecimiento ni en hombres ni en mujeres.

Hemos comprobado que los parámetros temporales, de amplitud y de calidad de voz pueden servir como medida objetiva de los procesos cognitivos del lenguaje. La cuestión que queda por responder es si los cambios en el lenguaje expresivo reflejan deficiencias en los procesos puramente lingüísticos (es decir, en el acceso léxico-semántico), que acaban afectando al estado cognitivo general de la persona; si esos cambios son el resultado de deficiencias generales en el procesamiento cognitivo (es decir, en los procesos ejecutivos), o si tales cambios surgen como una interacción entre ambos tipos de deficiencias (Le Rhun et al., 2005). Esta cuestión refleja la discusión clásica sobre si el lenguaje actúa como un sistema cognitivo independiente o si, por el contrario, el estado cognitivo -incluyendo todos sus procesos subyacentes- es una herramienta general de procesamiento de la información, de la que el lenguaje es uno de los componentes cognitivos (Code, 2018; Davis, 2011). En base a la relación entre los parámetros y el estado general, sugerimos que estos miden los procesos ejecutivos implicados en la producción del lenguaje. En este sentido, pensamos que los resultados apoyarían modelos teóricos del envejecimiento que defienden que con la edad se da una reducción de los recursos de procesamiento, tanto en el envejecimiento exitoso como en el patológico, dado que encontramos una alteración de los parámetros acorde al estado mental general, sin necesidad de medir procesos o patologías concretas. Así, la explicación implicaría que los recursos atencionales que apoyan el procesamiento

cognitivo disminuirían con la edad, lo que conduce a déficits en los procesos de control ejecutivo (Moscovitch & Winocur, 1995; West, 1996) y en último término, a la producción de lenguaje.

Es necesario tener en cuenta una limitación en este estudio. Se han encontrado diferencias en la escolaridad entre los grupos que afectan principalmente al grupo con mayor deterioro cognitivo. Las personas con MMSE<23 tienen menos años de escolaridad y además son significativamente mayores que las de los otros dos grupos. Aunque parece que la edad no juega un papel relevante en los parámetros relacionados con el MMSE, la escolaridad sigue siendo un factor a explorar.

Estas variables y algunas otras, como la codificación o la capacidad de aprendizaje, podrían contribuir a la capacidad de expresión del habla, y sería necesario explorar su papel en mayor medida. Esperamos que la indagación en los parámetros del habla y de los diferentes factores que afectan a su rendimiento contribuya a explicar y optimizar las herramientas de cribado basadas en el análisis automático del habla.

La principal novedad de este estudio es que vamos más allá de la clásica clasificación binaria de grupos como demencia vs. no demencia. Hemos identificado un conjunto de parámetros del habla que son capaces de cuantificar la gravedad del deterioro cognitivo, medido por el MMSE. Con ello, confirmamos nuestra hipótesis inicial según la cual el habla de los mayores estaría determinada por sus recursos cognitivos. Esto nos lleva a sugerir que, más allá de la relación con el estado mental general puede haber también una relación entre los parámetros del habla y determinados procesos cognitivos específicos.

Estudio III

Los cambios en el ritmo del habla diferencian personas con deterioro cognitivo leve no degenerativo y personas con demencia preclínica

En los estudios anteriores hemos explorado factores que influyen en los parámetros del habla, y hemos visto algunos correlatos cognitivos que contribuyen a explicar los cambios en algunos parámetros debidos, especialmente al deterioro cognitivo. En este nuevo estudio, nos planteamos analizar el funcionamiento de estos parámetros del habla con el fin de elaborar diferenciales perfiles cognitivos, de modo que nos permita distinguir entre el perfil relacionado con la EA frente a otros de deterioro cognitivo no relacionados con la demencia.

El deterioro cognitivo en las personas mayores tiende a ser infradiagnosticado dado el coste económico y personal que supone la realización de pruebas neuropsicológicas o de biomarcadores. De hecho, sus alteraciones suelen atribuirse a procesos comunes de deterioro asociados a la edad longeva, cuando una parte de las personas que lo sufren están padeciendo en realidad una fase preclínica de la EA. Se calcula que entre el 15 y el 20% de las personas con DCL desarrollan posteriormente la EA (Jessen et al., 2010). Detectar estos casos con un diagnóstico precoz es importante porque permitirá a las familias planificar su futuro, implicar a los pacientes en el control de los factores de riesgo, proporcionar los cuidados adecuados, prever nuevos síntomas que pueden afectar a la vida de las personas y organizar un tratamiento más eficaz.

Un objetivo primordial de la investigación en la actualidad en relación con la demencia es perfeccionar el cribado precoz de la EA (Laske et al., 2015), distinguiendo

entre las personas mayores con los déficits típicos del DCL, pero que no tienen EA, de las que sí están desarrollando la enfermedad. Para ello, nos centramos en otros procesos cognitivos que tienen más probabilidades de revelar déficits preclínicos, como cambios en la producción del lenguaje, especialmente las alteraciones en el acceso léxico y semántico (Yi-Hsiu, 2017). Weiner et al., (2008) han encontrado una estrecha correlación entre la gravedad de la demencia y diferentes medidas lingüísticas; sin embargo, los resultados de los estudios con pacientes en las etapas leves de la EA (Forbes-McKay et al., 2013; Fraser, Meltzer, et al., 2016) no pueden generalizarse a las etapas más tempranas de la enfermedad, cuando las estrategias compensatorias pueden ocultar las alteraciones en el lenguaje funcional (Mueller et al., 2018).

Algunos estudios específicos con personas con DCL sí que han identificado déficits sutiles en el lenguaje (Szatloczki et al., 2015) en tareas de denominación de imágenes y pruebas semánticas (Garrard et al., 2014), cambios en la frecuencia de palabras utilizadas para utilizar palabras más comunes (Baldas et al., 2011), cambios en el ritmo en el lenguaje espontáneo (Laske et al., 2015), déficits en tareas de fluidez verbal (Manouilidou et al., 2016), y parafasias fonémicas debidas a la dificultad para producir los sonidos del habla en el orden correcto (Wutzler et al., 2013). Por tanto, podría ser posible identificar marcadores del habla que distinguen entre los individuos con DCL que posteriormente desarrollarán EA y los que no.

El objetivo de este estudio es investigar una serie de aspectos del habla y comprobar si estos parámetros muestran diferencias estadísticas entre individuos con DCL debido a un trastorno no degenerativo (nodMCI) e individuos con DCL con síntomas preclínicos de demencia y que probablemente desarrollarán EA (preEA). Consideramos que los cambios en estos parámetros del control del habla pueden ser clínicamente útiles, reflejando las dificultades de individuos con demencia temprana.

Estudios previos de nuestro grupo, han obtenido una función de cribado discriminante para el diagnóstico de la EA basada en el Análisis de Voz para el Diagnóstico de la Enfermedad de Alzheimer (VAD-EA, Martínez-Sánchez et al., 2018), que califica la probabilidad de que el habla de un individuo corresponda a alguien con un envejecimiento no patológico (o CS) o a alguien con EA. Este algoritmo clasifica correctamente el 92,4% de los casos (logra una clasificación de normalidad en el 97% de los casos). La función discriminante del DAV-EA y su conversión a valores Z con los que se asigna la probabilidad de padecer EA se puede observar en la tabla 10, de modo que el algoritmo contiene las siguientes variables: la edad; la media de la amplitud mínima (Amplit Min); la media del valor de la diferencia de amplitud máxima (Amplit Difference Max Mean), la ubicación de la frecuencia media dentro del intervalo de frecuencias (asimetría); la desviación típica del primer formante (sdF1); el ancho de banda del tercer formante (B3); la trayectoria del tono dentro de los núcleos silábicos, dividida por la duración (TrajIntra); el nPVI; la desviación típica de la ratio entre armónicos y ruido (HNRdt), y la ratio entre armónicos y ruido extraída de los parámetros de calidad de voz AVQI (AVQI HNR).

Tabla 10. Características del algoritmo del DAV-EA.

Parámetros	Coefficientes estandarizados	Puntuaciones Z media (dt)
Amplit Difference Max Mean	0,622	11,376 (2,212)
Amplit Min	0,702	6,071 (81,981)
Edad	0,461	77,690 (7,190)
sdF1	0,376	374,110 (115,190)
nPVI	0,505	57,178 (6,535)
HNRdt	-0,437	5,899 (0,817)
AVQI HNR	-0,478	0,164 (0,07)
Asimetría	-0,386	9,575 (4,329)
TrajIntra	0,289	4,798 (1,656)
B3	0,171	563,613 (107,740)

Utilizaremos una muestra de individuos con DCL que aún no han desarrollado los síntomas clínicos de una enfermedad neurodegenerativa. Creemos que, en esta muestra, podríamos encontrar nuevas y diferentes variables que muestren déficits al inicio de la enfermedad, es decir, parámetros que diferencien a los que tienen la enfermedad de los que no. Por ello, nos propusimos analizar otros parámetros que, en teoría, podrían ser más sensibles en las fases preclínicas de la enfermedad.

Los parámetros de habla evaluados serán:

- 1) Relacionados con la duración: duración, tiempo de fonación, velocidad de articulación y de habla, número de pausas.
- 2) Relacionados con el ritmo del habla: la duración media de las sílabas, la desviación estándar de la duración de las sílabas, el coeficiente de variación de la duración de las sílabas y el nPVI.
- 3) Parámetros derivados del análisis del espectro: asimetría, centro de gravedad y su desviación típica, LTAS, su media y desviación típica y el espectro medio entre las siguientes bandas de frecuencias 50 Hz-1000Hz, 1000Hz- 2000Hz, y 2000Hz-4000Hz.
- 4) Energía: intensidad media en dB y desviación típica.
- 5) Frecuencia fundamental (F0) y perturbación de los pulsos glotales: F0, Jitter local, shimmer local, el porcentaje de roturas de voz, HNR y AVQI.

Método

Participantes

En este estudio se utilizaron 86 participantes que fueron categorizados como mayores con DCL. Se les grabó leyendo el primer párrafo del Quijote. Se pueden consultar las características de la muestra en la tabla 11. Se excluyeron los participantes con depresión o ansiedad (puntuación de >6 en el Test de Goldberg).

Tabla 11. Características de la muestra.

	Media (desviación típica)
Edad	79,36 (9,58)
Años de escolaridad	8,71 (0,2)
MMSE	23,41 (4,59)
MIS	4,84 (1,98)
BNT	19 (10,59)
Isaacs	31,7 (6,92)

A continuación, utilizando el algoritmo del VAD-EA se clasificó a los participantes en dos categorías. El algoritmo proporciona el grado de probabilidad de pertenecer a los grupos CN y EA. En la población con DCL, por tanto, esta probabilidad puede entenderse como la de desarrollar EA. Proponemos así, que esos dos grupos se corresponden con un DCL no degenerativo (nodDCL) y un grupo de DCL con EA subyacente o preclínico (preEA). Estos son los grupos que se compararán en este estudio.

Este cribado arrojó un grupo de 73 personas clasificadas como nodDCL (20 hombres y 53 mujeres; el 85% del total, con una edad media de 78,73 años y 8,78 años de escolaridad), y otro grupo de personas con una probable preEA formado por 13 mayores (cuatro hombres y nueve mujeres; 15% del total, con una edad media de 82,92

años y 8,31 años de escolaridad). Los dos grupos (véase la tabla 12) eran equivalentes en cuanto a la edad ($p = 0,28$), años de escolaridad ($p = 0,29$) y sexo (prueba $\chi^2 = 0,62$, $p = 0,803$). También fueron equivalentes en cuanto al estado cognitivo evaluado con el MMSE ($p = 0,283$), en las pruebas de fluidez verbal semántica evaluadas con el Test de Isaacs ($p = 0,733$), en las pruebas de fluidez verbal fonológica ($p = 0,765$), en el valor global de la escala MIS ($p = 0,181$) y en el Test de Denominación de Boston ($p = 0,15$). Tampoco encontramos diferencias en el estado emocional medido por la escala de ansiedad y depresión de Goldberg ($p = 0,788$). Las únicas diferencias que encontramos entre ambos grupos correspondieron a los valores de recuerdo libre del MIS de Buschke ($P = 0,04$), de modo que el grupo preAD mostró menor recuerdo libre de palabras.

Tabla 12. Comparación de los perfiles cognitivos de los grupos nodDCL y preEA.

	nodDCL (n = 73)	preEA (n = 13)	U de Mann-Whitney (prueba estandarizada)
Edad	78.73 (10.07)	82.92 (5.02)	385 (-1.080)
Hombres / mujeres	20 (27%) / 53 (73%)	4 (31%) / 9 (69%)	
Años de escolaridad	8.78 (4.13)	8.31 (4.71)	558 (1.048)
MMSE	23.25 (4.63)	24.31 (4.37)	380 (-1.073)
MIS (total)	4.96 (2.04)	4.11 (1.36)	315 (1.339)
MIS (recuerdo libre)	1.22 (1.21)	0.33 (0.50)	346 (2.019) *
MIS (recuerdo clave)	1.83 (1.18)	2.44 (1.01)	282 (0.812)
BNT	11.02 (4.61)	7.50 (5.89)	81 (-1.455)
Test de Isaacs	31.52 (7.14)	32.69 (5.75)	434 (-0.341)
Fluidez Fonológica (p)	6.53 (3.03)	6.62 (3.30)	437 (-0.299)
Test de Goldberg	4.73 (4.45)	4 (3.11)	456 (0.269)

Análisis estadístico

Dada la diferencia de tamaño entre los grupos, se utilizó la prueba U de Mann-Whitney con el objetivo de comparar una serie de parámetros del habla entre los mayores con nodDCL y los que se espera que desarrollen EA (preEA). Se comprobó también si había diferencias por género, en los casos en los que fue una variable relevante, se reporta el resultado de hombres y mujeres.

Resultados

La tabla 13 muestra las diferencias entre los grupos nodDCL y preEA en cada parámetro. Los casos en los que la diferencia entre los dos grupos es significativa están marcados con asteriscos. El análisis revela diferencias entre los grupos preEA y nodDCL en varios parámetros.

En la fluidez del habla y en lo que respecta a los parámetros de duración, encontramos diferencias significativas entre los dos grupos en el tiempo de lectura ($p = 0,005$) y en el tiempo de fonación ($p = 0,002$), siendo la duración mayor para el grupo preAD.

No encontramos diferencias ni en la velocidad de articulación ($p = 0,592$) ni de habla ($p = 0,191$). Sin embargo, sí que observamos que los participantes del grupo preEA producen más pausas ($p = 0,021$). En cuanto a los parámetros del ritmo silábico, encontramos diferencias significativas entre los dos grupos en la desviación típica de la duración de las sílabas ($p = 0,006$), siendo más variables los del grupo preEA, y también en el nPVI ($p = 0,002$), pero no en la duración media de las sílabas ($p = 0,481$) ni en el coeficiente de variación de la duración de las sílabas ($p = 0,130$).

En cuanto a los parámetros de frecuencia fundamental, no hay diferencias entre los dos grupos ($p = 0,634$). Ya hemos visto que este parámetro puede estar determinado por las diferencias de género, cuestión que también se refleja en este estudio confirmando que las mujeres utilizan frecuencias más altas que los hombres ($F_{2,81} = 11,109$, $p < 0,01$). Sin embargo, no observamos diferencias entre los dos grupos en función del género.

En los parámetros del análisis espectral, los valores de asimetría ($p = 0,04$) mostraron una concentración de energía en las frecuencias bajas en los individuos con alta probabilidad de desarrollar demencia, y en las altas en los que no parecen estar desarrollándola. No hay diferencias entre los dos grupos en los valores de las frecuencias del centro de gravedad ($p = 0,06$), pero sí en el grado de difusión en el espectro respecto a su centro de gravedad, siendo mayor entre las personas con riesgo de demencia (CdG_dt , $p = 0,02$). Tampoco hay diferencias en los parámetros que componen el AVQI,

En LTAS encontramos diferencias significativas entre los dos grupos, con una menor intensidad en el espectro global de frecuencias en los individuos del grupo preEA ($p = 0,001$), siendo su desviación típica también menor ($p = 0,001$). En estos parámetros también aparecen diferencias entre ambos grupos en su rango de bajas frecuencias, entre 50 y 1000 Hz ($p = 0,008$). Esas diferencias aparecen en las mujeres ($Ltas\ media$, $p = 0,001$; $Ltas\ dt$, $p = 0,003$; $ltas\ entre\ 50\ y\ 1000\ Hz$, $p = 0,009$), pero no en los hombres ($Ltas\ media$, $p = 0,13$; $Ltas_dt$, $p = 0,21$; $ltas\ entre\ 50\ y\ 1000\ Hz$, $p = 0,43$).

En cuanto a los parámetros de intensidad, el grupo preEA muestra menor amplitud media ($p = 0,002$). Si tenemos en cuenta el género, hay diferencias significativas entre las mujeres ($p = 0,001$), pero no entre los hombres. No encontramos diferencias en la desviación típica de la amplitud ($p = 0,98$), ni en el rango entre la amplitud máxima y la media ($p = 0,24$) o la amplitud mínima ($p = 0,36$), parámetros relevantes en el algoritmo del VAD-AD.

Por último, en a los parámetros de perturbación, no encontramos diferencias entre los dos grupos en la ratio armónicos-ruido ($p = 0,281$) o en los valores de disfonía evaluados por el AVQI ($p = 0,551$). Tampoco hay diferencias en los valores de jitter local ($p = 0,519$), shimmer local ($p = 0,199$) ni roturas de voz ($p = 0,088$).

Tabla 13. Comparación de los parámetros del habla de los grupos nodDCL y preEA.

Parámetros	nodDCL media (dt)	preEA media (dt)	U de Mann-Whitney (prueba estandarizada)
Parámetros de duración			
Duración	47,48 (22,20)	61,74 (20,95)	243 (-2,791)**
Tiempo de fonación	32,35 (10,39)	39,44 (8,75)	213 (-3,153)**
Velocidad de habla	3,35 (0,80)	3,07 (0,92)	583 (1,308)
Número de pausas	25,79 (16,34)	34,23 (14,64)	283 (-2,310)
Parámetros de fluidez y ritmo			
nPVI	54,27 (5,42)	59,97 (5,76)	215 (-3,128)**
Duración sílaba	0,19 (0,03)	0,20 (0,02)	416 (-0,705)
Dt Duración sílaba	0,10 (0,02)	0,11 (0,02)	246 (-2,755)**
VARCO duración sílaba	0,51 (0,07)	0,55 (0,09)	349 (-1,513)
Velocidad de articulación	4,63 (0,59)	4,54 (0,60)	519 (0,537)
Parámetros de análisis espectral			
Asimetría	11,33 (4,14)	8,62 (4,38)	645 (2,056)*
Centro de gravedad	446 (98)	568 (227)	321 (-1,851)
Dt Centro de gravedad	547 (191)	753 (318)	276 (-2,393)*
LTAS media	32,46 (2,31)	30,20 (2,00)	763 (3,478)**
LTAS media (hombres)	32,41 (2,64)	30,47 (2,33)	30 (-0,878)
LTAS media (mujeres)	32,48 (2,20)	30,07 (1,97)	401 (3,247)**
LTAS dt	40,32 (2,10)	38,16 (2,82)	738 (3,177)**
LTAS dt (hombres)	40,21 (2,31)	38,46 (3,08)	57 (1,317)
LTAS dt (mujeres)	40,37 (2,03)	38,03 (2,88)	388 (2,987)**
LTAS 50-1000Hz	44,87 (2,03)	43,26 (2,37)	694 (2,646)**
LTAS 50-1000Hz (hombres)	46,24 (2,12)	44,76 (3,27)	51 (0,852)
LTAS 50-1000Hz (mujeres)	46,40 (2,06)	43,98 (4,12)	369 (2,608)**
AVQI	6,00 (1,18)	6,17 (0,95)	425 (-0,597)

Parámetros de intensidad				
Intensidad media		75,36 (1,77)	73,63 (2,00)	735 (3,141)**
Intensidad media	media	74,94 (0,41)	73,94 (0,91)	55 (1,162)
(hombres)				
Intensidad media (mujeres)		75,98 (0,41)	73,51 (0,60)	390 (3,027)**
Intensidad dt		17,28 (10,54)	14,77 (2,94)	551 (0,922)
Intensidad dt (hombres)		19,24 (14,31)	12,76 (4,26)	51 (0,852)
Intensidad dt (mujeres)		16,55 (8,78)	15,66 (1,79)	198 (-0,809)
Intensidad mínima		-9,08 (13,53)	25,01 (12,44)	551 (0,922)
Intensidad mínima	mínima	23,19 (129,34)	27,59 (14,61)	46 (0,465)
(hombres)				
Intensidad mínima	mínima	-3,76 (107,83)	23,85 (12,14)	285 (0,929)
(mujeres)				
Intensidad dif, Max-min		95,55 (13,40)	61,76 (12,26)	377 (-1,175)
Intensidad dif, Max-min		109,22 (129,34)	58,79 (14,07)	34 (-0,465)
(hombres)				
Intensidad dif, Max-min		90,37 (107,68)	63,07 (12,03)	181 (-1,149)
(mujeres)				
Parámetros de F0 y perturbación				
F0		168,61 (28,90)	164,82 (20,21)	514 (0,476)
F0 (hombres)		138,86 (25,17)	138,84 (16,31)	41 (0,077)
F0 (mujeres)		172,12 (16,34)	170,09 (13,73)	250 (0,230)
Jitter local		2,42 (1,07)	2,35 (0,66)	427 (-0,567)
Shimmer local		1,12 (0,27)	1,19 (0,22)	394 (-0,965)
% roturas de voz		38,93 (11,01)	43,93 (13,24)	333 (-1,706)
HNR		12,96 (3,50)	11,91 (3,18)	570 (1,151)
* $P < 0,05$; ** $P < 0,01$				

Discusión

Este estudio se ha centrado en determinar hasta qué punto las medidas del ritmo del habla pueden utilizarse como parámetros para predecir el futuro desarrollo de la demencia en individuos con DCL. De hecho, pone de manifiesto las dificultades que entraña el uso de las pruebas neuropsicológicas para el diagnóstico presintomático de la EA, ya que las pruebas de cribado no muestran diferencias significativas en las distintas escalas entre los individuos con un fonotipo de habla cercano a la EA (preEA) y los que

no tienen esas características de habla (nodDCL). En ambos grupos, cuando los mayores solicitan una evaluación cognitiva, obtienen una media de 23/24 puntos en el MMSE, lo que les sitúa en el grupo de DCL. Sólo la puntuación del test MIS para el recuerdo libre mostró diferencias entre los dos grupos.

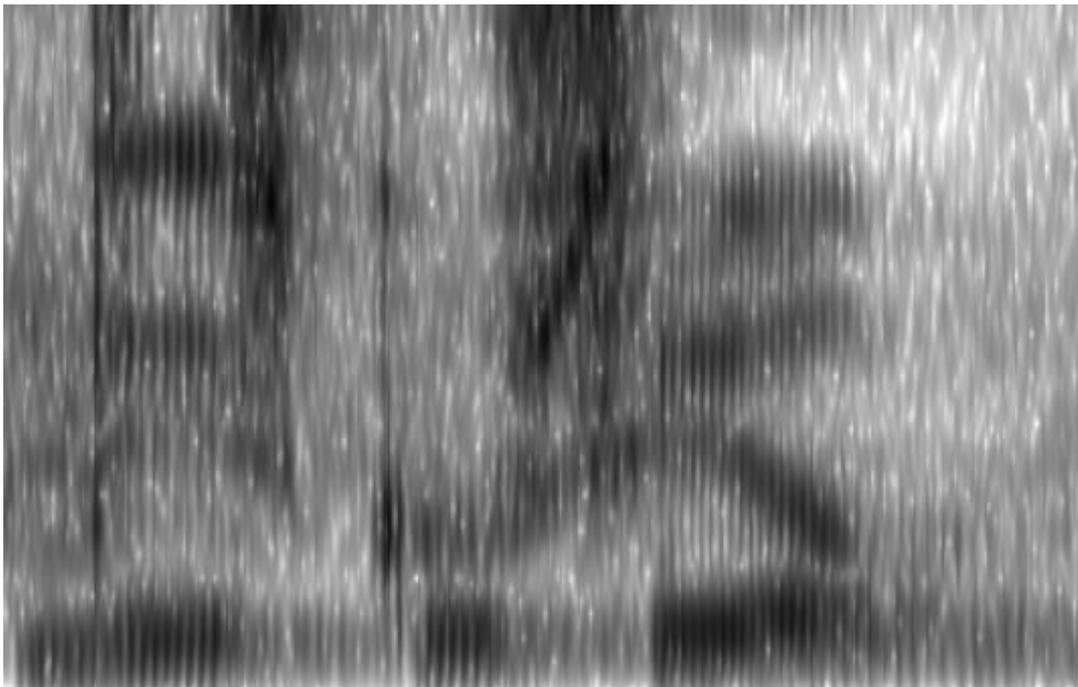
En cuanto a los 12 parámetros utilizados en el estudio que han demostrado ser significativos para distinguir entre aquellas personas con DCL que pueden o no desarrollar EA, sólo dos de los parámetros del VAD-EA mostraron diferencias entre ambos grupos: la asimetría y el nPVI. Esto sugiere una diferencia entre las alteraciones del lenguaje que caracterizan a la EA en comparación con los parámetros alterados en las fases preclínicas del trastorno, lo que pone de manifiesto que hay una evolución del habla en el desarrollo de la enfermedad.

Estos hallazgos tienen implicaciones para el uso de medidas acústicas a la hora de caracterizar el habla en las primeras etapas de la demencia. La presencia de ciertos rasgos rítmicos específicos en la ejecución del habla diferencia a estos individuos, revelando problemas sutiles que distinguen a los individuos con DCL que tienen más o menos probabilidades de desarrollar EA. Estos rasgos no parecen estar relacionados con el tono y las medidas acústicas o fonológicas. En la mayoría de los casos, son parámetros de duración, ritmo e intensidad. El grupo preEA se caracteriza por producir más pausas, un volumen de habla bajo y cambios en la fluidez, una entonación monótona y un menor ritmo de habla.

Por último, el conjunto general de parámetros que hemos obtenido relacionados con la probabilidad de desarrollar demencia parece confirmar la relación entre la demencia preclínica y los parámetros de duración, del análisis espectral y de ritmo. La obtención de un perfil de "alta probabilidad de demencia", junto con la existencia de un déficit cognitivo constatado en las pruebas neuropsicológicas, nos sugiere que podríamos

estar un indicador clave del desarrollo de una demencia, que debe ser confirmada mediante un biomarcador. La confirmación de los hallazgos de este estudio pasa por verificar la evolución de los individuos con DCL que han participado en este estudio, requiriendo su seguimiento longitudinal para seguir su evolución futura y realizar una comparación intrasujeto de sus muestras de habla.

6. Discusión



6. Discusión

El objetivo fundamental de esta tesis era el de verificar la utilidad del análisis del habla en la medición de procesos cognitivos alterados en los mayores con demencia y comprobar su efectividad para el desarrollo de herramientas diagnósticas. En este apartado pondremos en común las evidencias aportadas por los estudios en función de ese objetivo. Los resultados de estas tesis apoyan la idea de que el habla ofrece una información relevante sobre muchas de las operaciones psicológicas que realiza un individuo. Basándonos únicamente en nuestros estudios podemos afirmar que el habla puede servir como medida general de procesamiento a través del estado mental, de monitorización de sus cambios, en la estimación de la edad, en la medición de procesos lingüísticos de búsqueda en el léxico y la memoria semántica, del estado de la función ejecutiva, la identificación de un deterioro cognitivo y el pronóstico de su evolución, y la detección de estados afectivos patológicos.

Los primeros estudios de esta investigación plantean el habla como una medida de las funciones cognitivas, concretamente de los mayores, más allá de las consecuencias del envejecimiento a un nivel fisiológico. Así, planteamos que algunos parámetros del habla relacionados en otros estudios con el envejecimiento (Lortie et al., 2015; Mazzetto de Menezes et al., 2014) se deberían al estado cognitivo de los mayores. Esto revela la importancia del control de los factores de edad, género y estado cognitivo en el estudio del habla de los mayores. El resultado no es sorprendente, puesto que numerosos estudios habían puesto de manifiesto antes que esos parámetros del habla reflejaban estados de deterioro cognitivo (Hoffmann et al., 2010; Meilán et al., 2012), sí es más sorprendente que no haya una interacción entre la edad y el estado cognitivo, lo que sugiere que, según nuestros datos, la relación entre esos parámetros del habla y la edad podría ser, en realidad, un artefacto debido a un control deficiente. En otro estudio que realiza una

aproximación similar a la nuestra, comparando el habla en distintos tramos de edad, Oliveira Santos et al., (2021) encuentran diferencias significativas únicamente entre los participantes más jóvenes, de 30 años, y los mayores de 60. Es decir, las diferencias halladas en el habla de los mayores en otros estudios podrían surgir únicamente cuando se comparan con muestras de jóvenes. Con ello queda patente la importancia de que la evaluación del habla, como cualquier otra, se realice en comparación al grupo de referencia apropiado.

Más allá de un marcador del estado cognitivo general, nuestra investigación sigue la estela de trabajos que plantean que los cambios en el habla estarían determinados por funciones concretas (De Looze et al., 2018; Meilan et al., 2018). La explicación cognitiva de los parámetros de habla se ha hecho habitualmente de manera teórica, deduciendo que se debían a numerosos procesos con mayor o menor interrelación, entre los que encontrábamos la alteración del acceso al léxico, la planificación fonológica y sintáctica, mecanismos de compensación, etc. La evidencia en este sentido es contradictoria y los mismos parámetros han mostrado ser explicados por procesos diferentes, o, como sugerimos aquí, depender incluso de las tareas que se realicen. Por ello, hemos intentado explotar los conocidos déficits asociados a la demencia tipo Alzheimer. En el *estudio II* hemos hipotetizado que el habla en estas personas sería un marcador de los procesos de acceso léxico-semántico y la función ejecutiva, que no sólo son procesos habituales en la explicación de la producción de habla, sino que además sus déficits son específicamente característicos de la EA. Estos resultados tienen dos implicaciones relevantes. Por un lado, nos permiten afirmar que habrá parámetros del habla que son mejor explicados por déficits en un proceso específico y podríamos plantear incluso, que pueden ser utilizados para medir ese proceso concreto. Por otro, pone de manifiesto que los estudios sobre el habla en la EA (o cualquier otra alteración), y específicamente los de detección, deberán

implicar tareas que tengan en cuenta sus alteraciones cognitivas críticas. En este sentido son interesantes los estudios que exploran el habla espontánea durante tareas de memoria episódica (Pistono et al., 2019). También es lo que nos ha llevado a nosotros a plantear una tarea como la de fluidez verbal basada en el juego del parchís. Un reto que afrontar a este respecto será el de seleccionar una tarea adecuada para continuar los estudios acerca de la depresión. Entre los antecedentes descritos, la propuesta de König et al., (2019) consiste en solicitar al participante que relate un evento positivo y uno negativo que le haya ocurrido. Esta tarea contiene un elemento importante como es implicar un recuerdo afectivo, sin embargo, al recaer en parte sobre la memoria episódica, podría introducir error en la muestra con EA. A pesar de ello, es muy probable que esta propuesta marque el camino de estudios próximos.

Explorados los procesos alterados en la producción del habla de los mayores con demencia, propusimos una siguiente fase aplicada, es decir, el uso del análisis del habla en la detección de estados patológicos. El *estudio I* indicó que algunos de los parámetros del habla son de utilidad para determinar el estado cognitivo de una persona. De hecho, aunque la mayoría de los parámetros sólo mostraron diferencias en los estados más avanzados, cuando la demencia es más evidente, otros fueron más sensibles a cambios intermedios, por lo que podrían utilizarse como marcadores de un deterioro cognitivo leve. En esta fase de la investigación podríamos diferenciar dos partes. La primera se corresponde con el estudio III en el que buscamos posibles perfiles que se correspondan con un DCL debido a la enfermedad de Alzheimer. La segunda incluiría los estudios IV y V en los que planteamos posibles nuevas combinaciones de parámetros del habla que complementen en el futuro el algoritmo del DAV-EA (Martínez-Sánchez et al., 2018).

El *estudio III* plantea que dentro de las personas con DCL hay una parte cuya habla se asemeja al de las personas con EA. Aunque cognitivamente son prácticamente

indistinguibles, diferenciándose únicamente en el recuerdo libre del MIS (codificación y recuperación de palabras), muestran además características del habla diferentes. Lo que sugerimos es que a partir de estas características podríamos detectar a las personas con DCL debido a EA y que, por tanto, continuarán su deterioro. Sin embargo, este trabajo queda incompleto en este punto, ya que requiere el seguimiento longitudinal de estas personas, de modo que se pueda confirmar el diagnóstico supuesto a través del que se ha construido el estudio. Este estudio entraría en el cuerpo de trabajos que aventuran que es posible detectar a través del habla perfiles concretos de personas con DCL y una evolución probable a EA (Beltrami et al., 2018; Pastoriza-Domínguez et al., 2022) o diferenciar aquellos que continuarán su deterioro de aquellos que no (O'Malley et al., 2020). Esta es sin duda una de las líneas de futuro más interesantes, porque más allá de la identificación del estado actual del individuo, conduce a la predicción de su desarrollo. Sin embargo, no hay por el momento estudios que confirmen el deterioro de los participantes. Un estudio de Luz et al., (2021) sí que ha realizado un seguimiento longitudinal de dos años, pero sus resultados de predicción del deterioro son poco alentadores, con únicamente un 61% de precisión. Este resultado no es extraño poniéndolo en perspectiva, si el habla detecta adecuadamente entre el 70 y el 86% de los casos de DCL y el objetivo que planteamos es detectar subgrupos dentro de estos, el reto es considerable.

Los dos estudios de la siguiente fase son más cercanos a la investigación mayoritaria sobre el habla y deterioro cognitivo. Aquí hemos tratado de partir de las investigaciones previas del GIR en Neurofisiología, Cognición y Conducta que llevaron al desarrollo del DAV-EA, marcando como objetivo el desarrollo de nuevos algoritmos de detección. Hemos tratado de manipular el habla utilizando como pista la tarea de los colores del parchís. No parece que esta tarea, por sí sola, arroje unos resultados ventajosos

y se queda en una precisión mediocre. Sin embargo, la combinación de esta tarea con la de lectura nos da, en el *estudio V*, uno de los mejores resultados de la literatura en una clasificación multiclase. Creemos que clasificar correctamente los tres grupos de CS, DCL y EA, convierte esta combinación de parámetros en la opción más razonable para ser usada en el ámbito clínico, dado que atiende a más necesidades que los anteriores y, además, sin que se vea mermada su sensibilidad a la etapa de DCL, es decir, la más temprana y con más probabilidades de no tener un diagnóstico.

Hemos de recordar que en nuestra revisión se hallaron estudios que afirmaban que los efectos del deterioro cognitivo sobre el habla son pobres o incluso podría ser artefactos estadísticos (Balagopalan et al., 2021; Lindsay et al., 2021). De hecho, esos estudios indican que el análisis del habla no produce parámetros discriminadores. Nuestros resultados no coinciden con esta postura y se alinean con los de una mayoría de trabajos publicados que comparten resultados positivos. Más allá del posible sesgo de publicación, creemos que las coincidencias de resultados con los de otros grupos son una prueba de la solidez de los hallazgos. Ya se habían observado distorsiones en los formantes (Themistocleous et al., 2018), en la F0 (Horley et al., 2010), alteraciones en el AVQI, el PVI y el trajintra (Martínez-Sánchez et al., 2018), la duración de las sílabas (Themistocleous et al., 2020) y la pendiente espectral (Thomas et al., 2020). Aportamos en este estudio, además, las bandas de energía, que hasta donde sabemos, no se había usado antes en el análisis del habla en deterioro cognitivo, aunque sí otras medidas de análisis espectral. En cualquier caso, esta afirmación podría considerarse no más que evidencia incidental, y se hará necesario realizar meta-análisis que confirmen la robustez de los hallazgos. Por el momento no se ha publicado ninguno, a pesar de que abundan las revisiones sistemáticas, y pensamos que será extremadamente complicado que se hagan

debido a las dificultades que indicamos anteriormente como la falta de claridad al informar de los parámetros implicados en los algoritmos.

Con la propuesta de unos parámetros del habla que permitieran el cribado de distintos niveles de deterioro del estado cognitivo se habría completado el objetivo de partida propuesto en el inicio de esta tesis. Sin embargo, como mostramos en la introducción, el habla se ha usado para detectar muchos otros estados. Por ello identificamos una necesidad distinta a la que podría atender el análisis del habla y planteamos una nueva línea que pretende complementar los hallazgos anteriores. Es así como surge la propuesta de una línea sobre el diagnóstico diferencial de la depresión y la demencia. En este terreno, como hemos visto, son muy pocos los estudios publicados. No hay, por el momento, ninguna propuesta sólida para el cribado de deterioro cognitivo y depresión. Pensamos que la aproximación que exponemos aquí puede ser una vía con futuro. Los estudios previos se han basado en muestras con diagnósticos de envejecimiento sano o deterioro cognitivo, y dentro de ellos, han tratado de detectar a los que mostraban síntomas de depresión. Nuestra propuesta incluye mayores de todo tipo de estado cognitivo, lo que pensamos que nos puede dar parámetros más específicos de la depresión, y aplicables a más situaciones. Este hecho situaría con ventaja nuestra propuesta, dado que su relevancia clínica sería mayor. Sin embargo, los resultados obtenidos aquí están sujetos a importantes limitaciones, no suponen más que un primer paso y será necesario refinar los protocolos y las pruebas. Además, esta propuesta tampoco resuelve el problema del diagnóstico diferencial, que será, en todo caso, un objetivo futuro.

En la línea de cuáles son los parámetros obtenidos en nuestros distintos estudios, observamos cierta variabilidad. Tomando como referencia el *estudio I* en el que mostramos que varios parámetros son indicadores de deterioro, veremos que esos mismos

parámetros apenas se repiten en los siguientes. Los parámetros de duración, desviación típica de la duración de la sílaba, nPVI, intensidad media, LTAS y LTAS 50-1000, también muestran diferencias significativas en el *estudio III*, mostrándose como parámetros asociados a la progresión del deterioro. Sin embargo, en los estudios siguientes, en los que seleccionamos los parámetros con mayor poder discriminativo, únicamente se incluye la duración de la sílaba. Si bien es cierto que en el caso de los estudios *IV* y *V* las tareas de lenguaje oral son diferentes y esto podría estar detrás de la falta de coherencia en la selección de parámetros. Esto nos lleva a pensar de nuevo en la dificultad que supone el establecimiento de parámetros de habla discriminadores, dada la cantidad de factores que influyen en la producción. No será posible generalizar los resultados de unas tareas a otras, por lo que se hará necesario el uso de protocolos muy cerrados si se pretende usar en el diagnóstico de una patología. Entre los estudios *IV* y *V*, se repiten la duración de la sílaba y la desviación típica del ancho de banda del tercer formante, confirmando en parte los resultados hallados. Por otro lado, entre el *estudio V*, sobre deterioro cognitivo, y el *estudio VI*, sobre depresión, coinciden los parámetros de la pendiente del LTAS extraído del AVQI (tilt) y la trayectoria del tono en la sílaba o TrajIntra. Dado que los análisis posteriores han descartado que los resultados se deban a la interacción de la depresión con el deterioro cognitivo, podemos pensar que su inclusión se debe a que, como apunta Ramanarayanan et al., (2022), los parámetros del habla podrían tener poca especificidad. Más que parámetros individuales, será su relación con otros la que podría determinar si son indicadores de una patología u otra.

Limitaciones

Los estudios propuestos tienen algunas limitaciones que deben tenerse en cuenta a la hora de interpretar sus resultados. En primer lugar, aunque la muestra del primer estudio es amplia, al incluir una base de datos basada en el trabajo de años del propio grupo de investigación, los grupos no eran perfectamente homogéneos y se observaron algunas diferencias entre ellos. El *estudio II*, ha incluido una prueba costosa en tiempo lo que ha limitado la cantidad de participantes. Además, observamos una diferencia de edad y nivel educativo que tratamos de tener en cuenta incluyendo las variables en el análisis. Este es en realidad un problema habitual en los estudios que incluyen mayores con demencia y del que los estudios sobre el habla no están libres, los grupos de personas con deterioro tienen más edad, a lo que se añade que están compuestos por más mujeres (Petti et al., 2020). Acorde a nuestra revisión, la mayoría de estos estudios tratan de controlar estos efectos bien mediante emparejamiento o con correcciones en el análisis estadístico. En nuestro caso, hemos adoptado la estrategia del emparejamiento en los estudios *IV* y *V*. Así, creamos grupos emparejados en edad y nivel educativo que además comprobamos que tenían número similar de hombres y mujeres. Este método nos está dando parámetros de habla que muy probablemente se relacionen únicamente con el deterioro, pero es difícil saber si será adecuado en la práctica clínica con sujetos variados también en otras condiciones y con comorbilidades. Será necesaria una validación clínica que permita comprobar la eficacia. Para evitar este problema de aplicabilidad, la muestra del estudio *VI* incluye un amplio rango de perfiles, pero vuelve a mostrar el defecto de que los grupos son diferentes en edad. Para descartar que este y otros factores tuvieran influencia, llevamos a cabo una serie de análisis posteriores que indicaron que es poco probable que estos factores hayan inducido error. Es decir, a lo largo de los estudios hemos intentado

minimizar el impacto de los defectos en la recogida de la muestra, pero evidentemente requerirá su ampliación y replicación si queremos confirmar estos resultados.

Otro factor a tener en cuenta tiene que ver con las tareas para elicitación del habla. En primer lugar, la tarea de lectura implica que el participante debe tener un mínimo de escolaridad. Por ello, para minimizar el impacto de la habilidad de lectura se ha utilizado como criterio de inclusión el de tener al menos seis años de escolaridad. Ante esto podemos señalar que habrá una parte de la población excluida de este tipo de herramientas de cribado (como lo están en realidad con muchas otras), problema que no sufren las tareas de lenguaje espontáneo. Además, un factor que podría afectar y que no se ha tenido en cuenta es la reserva cognitiva, podría no ser tan importante el nivel educativo como el hábito lector del mayor a lo largo de su vida. Esto es relevante porque una alta reserva cognitiva se ha relacionado con procesos que en esta tesis se han señalado como determinantes de la ejecución del habla, como la fluidez fonológica (Gonzalez-Burgos et al., 2020). Otra cuestión relacionada con las tareas es la dependencia de un conocimiento cultural de la tarea del parchís. Señalamos que una mayoría abrumadora de participantes conocían el juego, pero podría criticarse que el mero conocimiento no implica ni conocer los colores que lo forman ni que, conociéndolos, la respuesta deba ser rápida si no hay una alta familiaridad con él. Además, aunque es un juego universalmente conocido, será mucho menos popular en otras culturas y por tanto será complejo adaptar la tarea a distintos contextos.

Una última limitación que queremos señalar tiene que ver con críticas recientes al estudio del habla a través de lectura oral. Se ha señalado que con este fin debería utilizarse textos experimentales equilibrados fonéticamente y por la frecuencia léxica. Que, además, deberían incluir diferentes tipos de oraciones, dado que no es igual la prosodia de una frase declarativa que la de una interrogativa o exclamativa (Del Carmen Pérez-

Sánchez et al., 2021), y además, que no se suele contemplar la prosodia afectiva (Oh et al., 2021). Hemos argumentado que el texto escogido posee algunas de estas características y otras que pensamos que lo hacen ideal para los objetivos propuestos, pero no cumple con todos los criterios señalados.

Sugerencias para futuras investigaciones

A partir de los resultados obtenidos propondremos en este último apartado una serie de recomendaciones, implicaciones y potenciales líneas de investigación que se derivan directa e indirectamente de nuestras aportaciones.

Los resultados de varios de los estudios de esta tesis sugieren que podría haber una influencia de la edad y el género en la producción de algunos de los parámetros de habla identificados en los estudios sobre deterioro cognitivo. Una de las tareas pendientes es la de hacer estudios normativos que permitan comparar a cada individuo con su grupo de referencia. Esto permitiría mejorar la precisión de las herramientas de cribado. Tal tarea requerirá una gran cantidad de participantes y por tanto la movilización de numerosos recursos, sin embargo, es un paso necesario para la estandarización de la evaluación.

En esta tesis hemos hecho hincapié en la importancia de las tareas de lenguaje oral y hemos sugerido que permiten manipular los procesos cognitivos implicados en la producción de habla. Esto abre varias puertas. En primer lugar, a nivel del habla como cribado, será necesario desarrollar protocolos de grabación adecuados. Esto implica buscar cuáles son los más indicados para comprometer los procesos cognitivos clave en la identificación de cada alteración. Aventuramos que serán cada vez menos comunes las tareas utilizadas generalmente como la descripción de láminas, y estarán enfocadas a

maximizar los efectos que tenga sobre el habla el trastorno explorado. En segundo lugar, el estudio de las tareas nos lleva a poner el foco sobre la medición de procesos cognitivos a través del habla. Hemos defendido que tareas diferentes implicarán procesos diferentes. Por el momento esto es una asunción teórica respaldada solo parcialmente, será necesario confirmar tal afirmación a través de más estudios. Además, planteamos la posibilidad de que algunos parámetros sean indicadores independientes, es decir, que dependan exclusivamente de un proceso, lo que los convertiría en marcadores específicos. De nuevo, esto requerirá progresar en la investigación.

Uno de los retos más importantes y fundamentales es el de llevar a cabo seguimientos de los participantes y realizar estudios longitudinales. Aunque como hemos visto los resultados de diversos estudios plantean que las alteraciones del habla son indicadores tempranos, debe confirmarse con evaluaciones posteriores. En el fondo, cuando los estudios identifican alteraciones del habla, el mayor ya muestra otros signos de deterioro, aunque aún no suponga dificultades en su vida. Sin embargo, se sabe que el deterioro comienza muy sutilmente mucho tiempo antes de que sea evidente. Concretamente, se han identificado alteraciones en el lenguaje expresivo hasta nueve años antes de obtener un diagnóstico (Amieva et al., 2005). Por tanto, potencialmente se podrían hallar estos cambios en el habla, convirtiéndose en una herramienta de predicción del deterioro. Además, a través de esta clase de estudios se podría evaluar la evolución progresiva de los cambios en el habla en función de la severidad del deterioro, y, por tanto, mejorar su capacidad de monitorización.

Otra corriente interesante, que hemos intentado dar comienzo aquí, es la del diagnóstico diferencial. Aunque nosotros hemos planteado concretamente el de la depresión en mayores, hay muchas otras patologías susceptibles de identificación a través del habla que afectan a los mayores. Ya no solo las mismas que a cualquier otra edad, en

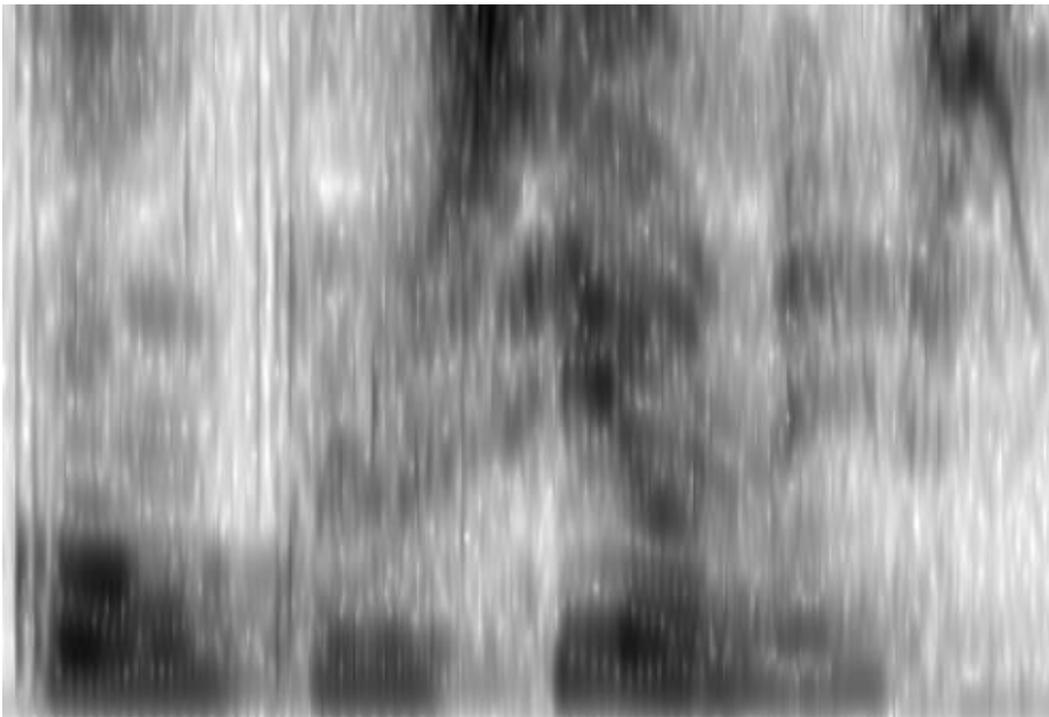
lo que nos remitimos a la variedad de trastornos que mencionamos en los antecedentes del estudio. También otras enfermedades relacionadas con la edad, concretamente otras clases de demencia. Si bien la demencia tipo Alzheimer es la más prevalente y por ello es lógico que los estudios se hayan centrado en ella, el método se puede hacer extensible a otras. Por ejemplo, la demencia vascular y frontotemporal han suscitado muy pocos estudios en este sentido.

Una cuestión sobre la que escasea la evidencia es si los marcadores del habla son independientes de la lengua hablada y, por tanto, generalizables a otros contextos. En la literatura hemos hallado pruebas a favor (Gosztolya et al., 2021) y en contra (Lindsay et al., 2021). Es lógico pensar que habrá parámetros dependientes de la lengua ya mencionados como los derivados del ritmo, la producción de sílabas y fonemas. Pero, por otro lado, podría ser que algunos sean independientes dado que no hay razones para pensar que una pausa utilizada como mecanismo de compensación por un fallo en el acceso léxico-semántico sea dependiente del idioma. Es decir, la producción del habla afectada por la alteración de un proceso cognitivo debería ser afectada de forma similar, aunque aparezcan determinadas especificidades. De nuevo, al ser escasa la información disponible al respecto, esta puede ser una cuestión a explorar.

Acabamos recordando que el análisis de habla es una medida conductual directa y objetiva, y esto es lo que le dota de interés para nosotros. Pero su uso tiene también un interés social. El potencial de estos resultados reside en la posibilidad de realizar estos análisis en dispositivos electrónicos (Martínez-Sánchez et al., 2018), aplicaciones (Hall et al., 2019; Yamada et al., 2021) o a través de llamadas telefónicas (Gregory et al., 2022). Aunque la muestra recogida para este estudio se realizó en el contexto del laboratorio, es fácil trasladar la misma recogida de habla a otros contextos, sin pérdida de calidad, e incluso es posible hacerlo en entornos ruidosos (Rohanian et al., 2021). Además, la

evidencia muestra que el cribado a través del habla es aceptado y valorado positivamente por los usuarios debido a su simplicidad (Arevian et al., 2020; Diaz-Asper et al., 2021). Otras ventajas son la objetividad de la evaluación, la capacidad de llegar a un amplio sector de la población debido al uso generalizado de tecnologías como los smartphones, y la consecuente baja carga para el sistema sanitario (Kourtis et al., 2019). Llegados a este punto, no es difícil imaginar una aplicación de evaluación que contenga varios algoritmos que permitan hacer un perfil de la probabilidad de que la voz de una persona sea indicadora de diferentes patologías, permitiendo, a través de una o unas pocas tareas de habla, hacer un cribado de numerosas alteraciones.

7. Conclusiones



7. Conclusiones

Con los estudios propuestos hemos tratado de responder a los objetivos iniciales, que, de forma general, planteaban la necesidad de explorar los factores que afectan al habla de las personas con Enfermedad de Alzheimer o Deterioro Cognitivo Leve. Pretendíamos, además, buscar la mejora de los métodos de cribado a través del análisis del habla. Así, hemos presentado estudios que indican que la edad, el género y aspectos cognitivos del mayor que afectan a la producción del habla. Hemos acompañado estos resultados con propuestas para emplear la medición de esos factores a través del análisis del habla como herramientas de cribado.

Atendiendo a los objetivos específicos, obtenemos las siguientes conclusiones:

1. *Investigar la causa de los cambios en el habla y sus correlatos cognitivos, observar de qué factores dependen y si pueden ser estimadores de la gravedad del deterioro.*
 - 1.1. *Comprobar la influencia de factores sociodemográficos y cognitivos sobre los cambios en el habla del mayor.*

Conclusión 1. Hemos identificado factores que afectan a los parámetros de habla habitualmente asociados a la edad como son el género y, fundamentalmente, el estado cognitivo. El envejecimiento conlleva cambios en el tracto vocal y la capacidad pulmonar que afectan a la producción del habla, como sería el caso del jitter y el HNR, por lo que en los estudios sobre el habla en procesos neurodegenerativos será necesario controlar este efecto. También hemos hallado diferencias debidas al género, la diferencia en la velocidad de habla ya había sido documentada, pero parece evolucionar en un mismo sentido cuando se presenta un deterioro cognitivo.

Conclusión 2. Otros parámetros identificados parecen no cambiar ni por la edad ni por el deterioro. Estos son la asimetría y la F0, por lo que será necesario explorar en profundidad estos elementos y hallar una explicación al cambio encontrado en muchos estudios.

Conclusión 3. Algunos parámetros cambian a medida que avanza el deterioro cognitivo, y serán buenos estimadores de la gravedad del deterioro. Son parámetros de duración, pausas y del ritmo silábico. Es especialmente destacable el hallazgo de la desviación típica de la duración de la sílaba como único parámetro que distingue los primeros estadios del deterioro.

1.2. Evaluar la influencia de diversos procesos cognitivos sobre los parámetros del habla.

Conclusión 4. Los parámetros explorados pueden servir como una medida directa del rendimiento en tareas asociadas a dominios cognitivos concretos. Hemos hallado parámetros cuya varianza viene explicada principalmente por tareas de fluidez verbal semántica como la duración total, velocidad de habla, duración de las sílabas, asimetría y HNR. Y su afectación, por tanto, tendría que ver con el deterioro del almacenamiento en la memoria semántica y dificultades en el acceso léxico-semántico. Otros parámetros, número de pausas, la dt de la duración de la sílaba, la intensidad media y el LTAS, vienen mejor explicados por la fluidez verbal fonológica y tendrían que ver con la planificación fonológica y los componentes de función ejecutiva de la tarea. Un solo parámetro, el nPVI, ha sido mejor explicado por una tarea de función ejecutiva, y podría tener que ver con la capacidad de planificación sintáctica y motora. En otros, F0, dt de F1 y jitter local, no hemos hallado ningún efecto. Los hallazgos nos permiten suponer que el análisis del habla podría servir para la evaluación de procesos cognitivos específicos, más allá del deterioro general.

2. *Abordar el cribado de DCL y EA mediante diferentes tareas y observando perfiles concretos que permitan mejorar la detección temprana y la utilidad diagnóstica actual del análisis automático del habla.*

2.1. *Explorar la existencia de perfiles de habla de personas con DCL que sean predictores del desarrollo de la demencia.*

Conclusión 5. Dentro de las personas con DCL hemos identificado dos perfiles de habla aplicando una herramienta de cribado de EA. Esto nos lleva a pensar que los identificados por la herramienta podrían ser aquellos cuya causa subyacente es el estado preclínico de tal enfermedad.

Conclusión 6. Hemos identificado nuevos parámetros del habla en los que difieren las personas con DCL no degenerativo y aquellos con una EA con presencia de deterioro cognitivo. Sin embargo, queda por explorar su capacidad discriminativa.

2.2. *Hallar combinaciones de parámetros óptimas para el cribado de DCL y EA.*
Estudios IV y V.

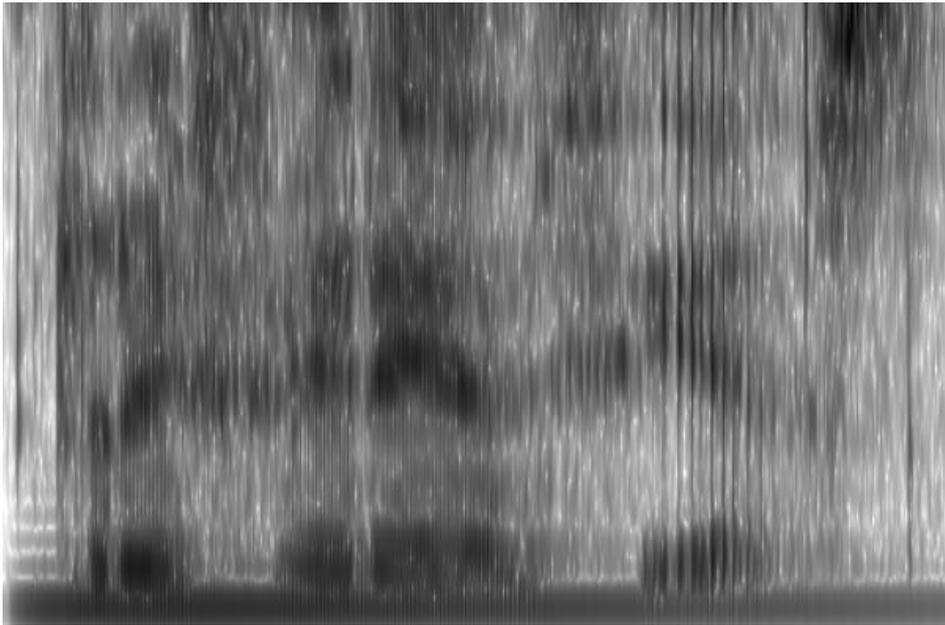
Conclusión 7. Con una tarea controlada de fluidez verbal, logramos distinguir a las personas con DCL de un grupo de mayores sanos con un 75% de precisión. Si bien el resultado es modesto, sugiere que podría ser beneficioso combinar información acústica con otra derivada del rendimiento de la tarea.

Conclusión 8. En la segunda exploración, combinamos esa misma tarea de fluidez con otra de lectura distinguiendo correctamente al 83,3% de los mayores sanos, con DCL o EA. Este es uno de los pocos estudios que ha tenido éxito en una comparación de los tres grupos al mismo tiempo, siendo una prometedora herramienta de cribado capaz de distinguir varios niveles de deterioro.

3. *Hacer extensible el método del análisis del habla a nuevas patologías asociadas al envejecimiento como la depresión en mayores.*

Conclusión 9. Se ha identificado un conjunto de parámetros asociados a los síntomas depresivos de una muestra de mayores independientemente de su estado cognitivo. Utilizando esos parámetros podemos identificar al 75,5% de los mayores con depresión. Esto supone un avance respecto a otros estudios que se limitan a muestras de mayores sanos y abre la puerta al diagnóstico diferencial de la depresión y la demencia.

8. Referencias



8. Referencias

- Abur, D., MacPherson, M. K., Shembel, A. C., & Stepp, C. E. (2021). Acoustic Measures of Voice and Physiologic Measures of Autonomic Arousal During Speech as a Function of Cognitive Load in Older Adults. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2020.12.027>
- Ahmed, S., Haigh, A.-M. F., de Jager, C. A., & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, *136*(12), 3727-3737. <https://doi.org/10.1093/brain/awt269>
- Ahmed, S., Jager, C. A. de, Haigh, A.-M. F., & Garrard, P. (2012). Logopenic aphasia in Alzheimer's disease: Clinical variant or clinical feature? *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(11), 1056-1062. <https://doi.org/10.1136/jnnp-2012-302798>
- Aita, S. L., Beach, J. D., Taylor, S. E., Borgogna, N. C., Harrell, M. N., & Hill, B. D. (2019). Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Applied Neuropsychology. Adult*, *26*(5), 441-451. <https://doi.org/10.1080/23279095.2018.1439830>
- Albertini, G., Bonassi, S., Dall'Armi, V., Giachetti, I., Giaquinto, S., & Mignano, M. (2010). Spectral analysis of the voice in Down Syndrome. *Research in Developmental Disabilities*, *31*(5), 995-1001. <https://doi.org/10.1016/j.ridd.2010.04.024>
- Albuquerque, L., Valente, A. R. S., Teixeira, A., Figueiredo, D., Sa-Couto, P., & Oliveira, C. (2021). Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan. *PLOS ONE*, *16*(4), e0248842. <https://doi.org/10.1371/journal.pone.0248842>

- Alghowinem, S., Goecke, R., Wagner, M., Epps, J., Breakspear, M., & Parker, G. (2013). Detecting depression: A comparison between spontaneous and read speech. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 7547-7551. <https://doi.org/10.1109/ICASSP.2013.6639130>
- Al-Hameed, S., Benaissa, M., Christensen, H., Mirheidari, B., Blackburn, D., & Reuber, M. (2019). A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints. *PloS One*, *14*(5), e0217388. <https://doi.org/10.1371/journal.pone.0217388>
- Almeida, J. S., Rebouças Filho, P. P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R., & de Albuquerque, V. H. C. (2019). Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recognition Letters*, *125*, 55-62. <https://doi.org/10.1016/j.patrec.2019.04.005>
- Alonso, J. B., De Leon, J., Alonso, I., & Ferrer, M. A. (2001). Automatic detection of pathologies in the voice by HOS based parameters. *EURASIP Journal on Applied Signal Processing*, *4*, 275-284.
- Alzheimer, A., Stelzmann, R. A., Schnitzlein, H. N., & Murtagh, F. R. (1995). An English translation of Alzheimer's 1907 paper, «Uber eine eigenartige Erkankung der Hirnrinde». *Clinical Anatomy (New York, N.Y.)*, *8*(6), 429-431. <https://doi.org/10.1002/ca.980080612>
- Alzheimer's Disease Association. (2022). *Alzheimer's Disease Facts and Figures*. <https://www.alz.org/alzheimers-dementia/facts-figures>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)* (5th edition). American Psychological Association. <https://doi.org/10.1176/appi.books.9780890425596>

- Amieva, H., Jacqmin-Gadda, H., Orgogozo, J.-M., Le Carret, N., Helmer, C., Letenneur, L., Barberger-Gateau, P., Fabrigoule, C., & Dartigues, J.-F. (2005). The 9 year cognitive decline before dementia of the Alzheimer type: A prospective population-based study. *Brain*, *128*(5), 1093-1101. <https://doi.org/10.1093/brain/awh451>
- Amjad, H., Roth, D. L., Sheehan, O. C., Lyketsos, C. G., Wolff, J. L., & Samus, Q. M. (2018). Underdiagnosis of Dementia: An Observational Study of Patterns in Diagnosis and Awareness in US Older Adults. *Journal of General Internal Medicine*, *33*(7), 1131-1138. <https://doi.org/10.1007/s11606-018-4377-y>
- Arevalo-Rodriguez, I., Smailagic, N., Figuls, M. R. i, Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O. L., Cosp, X. B., & Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews*, *3*. <https://doi.org/10.1002/14651858.CD010783.pub2>
- Arevian, A. C., Bone, D., Malandrakis, N., Martinez, V. R., Wells, K. B., Miklowitz, D. J., & Narayanan, S. (2020). Clinical state tracking in serious mental illness through computational analysis of speech. *PLOS ONE*, *15*(1), e0225695. <https://doi.org/10.1371/journal.pone.0225695>
- Arvaniti, A., & Rodriguez, T. (2013). The role of rhythm class, speaking rate, and F0 in language discrimination. *Laboratory Phonology*, *4*(1), 7-38. <https://doi.org/10.1515/lp-2013-0002>
- Asgari, M., Shafran, I., & Sheeber, L. B. (2014). Inferring clinical depression from speech and spoken utterances. *2014 IEEE international workshop on Machine Learning for Signal Processing (MLSP)*, 1-5.

- Balagopalan, A., Eyre, B., Robin, J., Rudzicz, F., & Novikova, J. (2021). Comparing Pre-trained and Feature-Based Models for Prediction of Alzheimer's Disease Based on Speech. *Frontiers in Aging Neuroscience*, *13*, 635945. <https://doi.org/10.3389/fnagi.2021.635945>
- Baldas, V., Lampiris, C., Capsalis, C., & Koutsouris, D. (2011). Early Diagnosis of Alzheimer's Type Dementia Using Continuous Speech Recognition. En J. C. Lin & K. S. Nikita (Eds.), *Wireless Mobile Communication and Healthcare* (pp. 105-110). Springer. https://doi.org/10.1007/978-3-642-20865-2_14
- Ballard, K. J., Robin, D. A., Woodworth, G., & Zimba, L. D. (2001). Age-related changes in motor control during articulator visuomotor tracking. *Journal of Speech, Language, and Hearing Research: JSLHR*, *44*(4), 763-777. [https://doi.org/10.1044/1092-4388\(2001/060\)](https://doi.org/10.1044/1092-4388(2001/060))
- Balogh, R., Imre, N., Gosztolya, G., Hoffmann, L., Pákáski, M., & Kálmán, J. (2022). The Role of Silence in Verbal Fluency Tasks – A New Approach for the Detection of Mild Cognitive Impairment. *Journal of the International Neuropsychological Society*, 1-13. <https://doi.org/10.1017/S1355617721001454>
- Banks, R. E., Beal, D. S., & Hunter, E. J. (2021). Sports Related Concussion Impacts Speech Rate and Muscle Physiology. *Brain Injury*, *35*(10), 1275-1283. <https://doi.org/10.1080/02699052.2021.1972150>
- Barth, J., Nickel, F., & Kolominsky-Rabas, P. L. (2018). Diagnosis of cognitive decline and dementia in rural areas—A scoping review. *International Journal of Geriatric Psychiatry*, *33*(3), 459-474. <https://doi.org/10.1002/gps.4841>
- Basilakos, A. (2018). Contemporary Approaches to the Management of Post-stroke Apraxia of Speech. *Seminars in Speech and Language*, *39*(1), 25-36. <https://doi.org/10.1055/s-0037-1608853>

- Basilakos, A., Yourganov, G., den, O. D.-B., Fogerty, D., Rorden, C., Feenaughty, L., & Fridriksson, J. (2017). A Multivariate Analytic Approach to the Differential Diagnosis of Apraxia of Speech. *Journal of Speech, Language, and Hearing Research, 60*(12), 3378-3392. https://doi.org/10.1044/2017_JSLHR-S-16-0443
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*(1), 77-100. [https://doi.org/10.1016/0272-7358\(88\)90050-5](https://doi.org/10.1016/0272-7358(88)90050-5)
- Bell, E. E., Chenery, H. J., & Ingram, J. C. L. (2001). Semantic Priming in Alzheimer's Dementia: Evidence for Dissociation of Automatic and Attentional Processes. *Brain and Language, 76*(2), 130-144. <https://doi.org/10.1006/brln.2000.2416>
- Beltrami, D., Gagliardi, G., Rossini Favretti, R., Ghidoni, E., Tamburini, F., & Calzà, L. (2018). Speech Analysis by Natural Language Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline? *Frontiers in Aging Neuroscience, 10*, 369. <https://doi.org/10.3389/fnagi.2018.00369>
- Bennett, S., & Thomas, A. J. (2014). Depression and dementia: Cause, consequence or coincidence? *Maturitas, 79*(2), 184-190. <https://doi.org/10.1016/j.maturitas.2014.05.009>
- Benson, A. D., Slavin, M. J., Tran, T.-T., Petrella, J. R., & Doraiswamy, P. M. (2005). Screening for Early Alzheimer's Disease: Is There Still a Role for the Mini-Mental State Examination? *Primary Care Companion to The Journal of Clinical Psychiatry, 7*(2), 62-69. <https://doi.org/10.4088%2Fpcc.v07n0204>
- Berg, E. A. (1948). A Simple Objective Technique for Measuring Flexibility in Thinking. *The Journal of General Psychology, 39*(1), 15-22. <https://doi.org/10.1080/00221309.1948.9918159>

- Biesbroek, J. M., van Zandvoort, M. J. E., Kappelle, L. J., Velthuis, B. K., Biessels, G. J., & Postma, A. (2016). Shared and distinct anatomical correlates of semantic and phonemic fluency revealed by lesion-symptom mapping in patients with ischemic stroke. *Brain Structure & Function*, *221*(4), 2123-2134.
<https://doi.org/10.1007/s00429-015-1033-8>
- Blair, M., Marczinski, C. A., Davis-Faroque, N., & Kertesz, A. (2007). A longitudinal study of language decline in Alzheimer's disease and frontotemporal dementia. *Journal of the International Neuropsychological Society*, *13*(2), 237-245.
<https://doi.org/10.1017/S1355617707070269>
- Blesa, R., Pujol, M., Aguilar, M., Santacruz, P., Bertran-Serra, I., Hernández, G., Sol, J. M., & Peña-Casanova, J. (2001). Clinical validity of the 'mini-mental state' for Spanish speaking communities. *Neuropsychologia*, *39*(11), 1150-1157.
[https://doi.org/10.1016/S0028-3932\(01\)00055-0](https://doi.org/10.1016/S0028-3932(01)00055-0)
- Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The Association Between Quantitative Measures of Dementia and of Senile Change in the Cerebral Grey Matter of Elderly Subjects. *The British Journal of Psychiatry*, *114*(512), 797-811. <https://doi.org/10.1192/bjp.114.512.797>
- Boer, J. N. de, Voppel, A. E., Brederoo, S. G., Schnack, H. G., Truong, K. P., Wijnen, F. N. K., & Sommer, I. E. C. (2021). Acoustic speech markers for schizophrenia-spectrum disorders: A diagnostic and symptom-recognition tool. *Psychological Medicine*, 1-11. <https://doi.org/10.1017/S0033291721002804>
- Boersma, Paul & Weenink, David (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.14, retrieved 24 May 2022 from <http://www.praat.org/>

- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001).
Disfluency rates in conversation: Effects of age, relationship, topic, role, and
gender. *Language and Speech, 44*(Pt 2), 123-147.
<https://doi.org/10.1177/00238309010440020101>
- Boullosa, R. R., & Pérez Ruíz, S. J. (1984). Long-term average spectrum of Spanish.
Applied Acoustics, 17(3), 201-208. [https://doi.org/10.1016/0003-682X\(84\)90037-9](https://doi.org/10.1016/0003-682X(84)90037-9)
- Braga, D., Madureira, A. M., Coelho, L., & Ajith, R. (2019). Automatic detection of
Parkinson's disease based on acoustic analysis of speech. *Engineering
Applications of Artificial Intelligence, 77*, 148-158.
<https://doi.org/10.1016/j.engappai.2018.09.018>
- Braun, S., Annovazzi, C., Botella, C., Bridler, R., Camussi, E., Delfino, J. P., Mohr, C.,
Moragrega, I., Papagno, C., Pisoni, A., Soler, C., Seifritz, E., & Stassen, H. H.
(2016). Assessing Chronic Stress, Coping Skills, and Mood Disorders through
Speech Analysis: A Self-Assessment 'Voice App' for Laptops, Tablets, and
Smartphones. *Psychopathology, 49*(6), 406-419.
<https://doi.org/10.1159/000450959>
- Bruzzi, C., Salsi, D., Minghetti, D., Negri, M., Casolino, D., & Sessa, M. (2017).
Presbyphonia. *Acta Bio Medica : Atenei Parmensis, 88*(1), 6-10.
<https://doi.org/10.23750/abm.v88i1.5266>
- Buschke, H. (1984). Cued recall in amnesia. *Journal of Clinical Neuropsychology, 6*(4),
433-440. <https://doi.org/10.1080/01688638408401233>
- Buschke, H., Kuslansky, G., Katz, M., Stewart, W. F., Sliwinski, M. J., Eckholdt, H.
M., & Lipton, R. B. (1999). Screening for dementia with the Memory

- Impairment Screen. *Neurology*, 52(2), 231-231.
<https://doi.org/10.1212/WNL.52.2.231>
- Cahoon, C. G. (2012). Depression in Older Adults. *AJN The American Journal of Nursing*, 112(11), 22-30. <https://doi.org/10.1097/01.NAJ.0000422251.65212.4b>
- Cassol, M., Reppold, C. T., Ferrão, Y., Gurgel, L. G., & Almada, C. P. (2010). Análise de características vocais e de aspectos psicológicos em indivíduos com transtorno obsessivo-compulsivo. *Revista da Sociedade Brasileira de Fonoaudiologia*, 15, 491-496. <https://doi.org/10.1590/S1516-80342010000400004>
- Cera, M. L., Ortiz, K. Z., Bertolucci, P. H. F., & Minett, T. (2018). Phonetic and phonological aspects of speech in Alzheimer's disease. *Aphasiology*, 32(1), 88-102. <https://doi.org/10.1080/02687038.2017.1362687>
- Chan, J. C. S., Stout, J. C., & Vogel, A. P. (2019). Speech in prodromal and symptomatic Huntington's disease as a model of measuring onset and progression in dominantly inherited neurodegenerative diseases. *Neuroscience & Biobehavioral Reviews*, 107, 450-460.
<https://doi.org/10.1016/j.neubiorev.2019.08.009>
- Chan, J. Y. C., Kwong, J. S. W., Wong, A., Kwok, T. C. Y., & Tsoi, K. K. F. (2018). Comparison of Computerized and Paper-and-Pencil Memory Tests in Detection of Mild Cognitive Impairment and Dementia: A Systematic Review and Meta-analysis of Diagnostic Studies. *Journal of the American Medical Directors Association*, 19(9), 748-756.e5. <https://doi.org/10.1016/j.jamda.2018.05.010>
- Chertkow, H., Bub, D., Bergman, H., Bruemmer, A., Merling, A., & Rothfleisch, J. (1994). Increased semantic priming in patients with dementia of the Alzheimer's

type. *Journal of Clinical and Experimental Neuropsychology*, 16(4), 608-622.

<https://doi.org/10.1080/01688639408402672>

Chiaramonte, R., & Bonfiglio, M. (2020a). Análisis acústico de la voz en la enfermedad de Parkinson: Revisión sistemática de la discapacidad vocal y metaanálisis de estudios. *Rev. neurol. (Ed. impr.)*, 393-405.

Chiaramonte, R., & Bonfiglio, M. (2020b). Acoustic analysis of voice in bulbar amyotrophic lateral sclerosis: A systematic review and meta-analysis of studies. *Logopedics Phoniatrics Vocology*, 45(4), 151-163.

<https://doi.org/10.1080/14015439.2019.1687748>

Chien, Y.-W., Hong, S.-Y., Cheah, W.-T., Yao, L.-H., Chang, Y.-L., & Fu, L.-C. (2019). An Automatic Assessment System for Alzheimer's Disease Based on Speech Using Feature Sequence Generator and Recurrent Neural Network. *Scientific Reports*, 9(1), 19597. <https://doi.org/10.1038/s41598-019-56020-x>

Chlasta, K., & Wołk, K. (2020). Towards Computer-Based Automated Screening of Dementia Through Spontaneous Speech. *Frontiers in Psychology*, 11, 623237. <https://doi.org/10.3389/fpsyg.2020.623237>

Chupin, M., Gérardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., Benali, H., Garnero, L., Colliot, O., & Alzheimer's Disease Neuroimaging Initiative. (2009). Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus*, 19(6), 579-587. <https://doi.org/10.1002/hipo.20626>

Code, C. (2018). Aphasia is not a disorder of language: Discuss. *Aphasiology*, 32(sup1), 52-53. <https://doi.org/10.1080/02687038.2018.1486390>

Cohen, A. S., Dinzeo, T. J., Donovan, N. J., Brown, C. E., & Morrison, S. C. (2015). Vocal acoustic analysis as a biometric indicator of information processing:

- Implications for neurological and psychiatric disorders. *Psychiatry Research*, 226(1), 235-241. <https://doi.org/10.1016/j.psychres.2014.12.054>
- Cohen, A. S., Najolia, G. M., Kim, Y., & Dinzeo, T. J. (2012). On the boundaries of blunt affect/alogia across severe mental illness: Implications for Research Domain Criteria. *Schizophrenia Research*, 140(1), 41-45. <https://doi.org/10.1016/j.schres.2012.07.001>
- Cohen, A. S., Renshaw, T. L., Mitchell, K. R., & Kim, Y. (2016). A psychometric investigation of “macroscopic” speech measures for clinical and psychological science. *Behavior Research Methods*, 48(2), 475-486. <https://doi.org/10.3758/s13428-015-0584-1>
- Cohen Priva, U. (2017). Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition*, 160, 27-34. <https://doi.org/10.1016/j.cognition.2016.12.002>
- Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, 30(1-2), 1-31. <https://doi.org/10.1080/23273798.2014.963130>
- Conejero, I., Olié, E., Courtet, P., & Calati, R. (2018). Suicide in older adults: Current perspectives. *Clinical Interventions in Aging*, 13, 691-699. <https://doi.org/10.2147/CIA.S130670>
- Corley, M., & Stewart, O. W. (2008). Hesitation Disfluencies in Spontaneous Speech: The Meaning of um. *Language and Linguistics Compass*, 2(4), 589-602. <https://doi.org/10.1111/j.1749-818X.2008.00068.x>
- Cormack, F. K., Barnett, J. H., & Taptiklis, N. (2021). Acoustic features of voice as a measure of cognitive load during performance of serial subtraction in a remote data collection context. *Alzheimer's & Dementia*, 17(S11), e056271. <https://doi.org/10.1002/alz.056271>

- Correia, J. M., Caballero-Gaudes, C., Guediche, S., & Carreiras, M. (2020). Phonatory and articulatory representations of speech production in cortical and subcortical fMRI responses. *Scientific Reports*, *10*(1), 4529. <https://doi.org/10.1038/s41598-020-61435-y>
- Covington, M. A., Lunden, S. L. A., Cristofaro, S. L., Wan, C. R., Bailey, C. T., Broussard, B., Fogarty, R., Johnson, S., Zhang, S., & Compton, M. T. (2012). Phonetic measures of reduced tongue movement correlate with negative symptom severity in hospitalized patients with first-episode schizophrenia-spectrum disorders. *Schizophrenia Research*, *142*(1), 93-95. <https://doi.org/10.1016/j.schres.2012.10.005>
- Croisile, B., Ska, B., Brabant, M.-J., Duchene, A., Lepage, Y., Aimard, G., & Trillet, M. (1996). Comparative Study of Oral and Written Picture Description in Patients with Alzheimer's Disease. *Brain and Language*, *53*(1), 1-19. <https://doi.org/10.1006/brln.1996.0033>
- Croot, K., Hodges, J. R., Xuereb, J., & Patterson, K. (2000). Phonological and Articulatory Impairment in Alzheimer's Disease: A Case Series. *Brain and Language*, *75*(2), 277-309. <https://doi.org/10.1006/brln.2000.2357>
- Cuetos, F., Gonzalez-Nosti, M., & Martínez, C. (2005). The picture-naming task in the analysis of cognitive deterioration in Alzheimer's disease. *Aphasiology*, *19*(6), 545-557. <https://doi.org/10.1080/02687030544000010>
- Cummins, N., Sethu, V., Epps, J., Schnieder, S., & Krajewski, J. (2015). Analysis of acoustic space variability in speech affected by depression. *Speech Communication*, *75*, 27-49. <https://doi.org/10.1016/j.specom.2015.09.003>
- Dahl, K. L., & Stepp, C. E. (2021). Changes in Relative Fundamental Frequency Under Increased Cognitive Load in Individuals With Healthy Voices. *Journal of*

Speech, Language, and Hearing Research, 64(4), 1189-1196.

https://doi.org/10.1044/2021_JSLHR-20-00134

Davis, G. A. (2011). The cognition of Language and Communication. En *Cognition and Acquired Language Disorders* (pp. 1-12). Elsevier.

de la Fuente Garcia, S., Ritchie, C. W., & Luz, S. (2020). Artificial intelligence, speech, and language processing approaches to monitoring Alzheimer's disease: A systematic review. *Journal of Alzheimer's Disease*, 78(4), 1547-1574.

<https://doi.org/10.3233/JAD-200888>

De Looze, C. D., Kelly, F., Crosby, L., Vourdanou, A., Coen, R. F., Walsh, C., Lawlor, B. A., & Reilly, R. B. (2018). Changes in Speech Chunking in Reading Aloud is a Marker of Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease. *Current Alzheimer Research*, 15(9), 828-847.

<https://doi.org/10.2174/1567205015666180404165017>

De Looze, C., Dehsarvi, A., Crosby, L., Vourdanou, A., Coen, R. F., Lawlor, B. A., & Reilly, R. B. (2021). Cognitive and Structural Correlates of Conversational Speech Timing in Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease: Relevance for Early Detection Approaches. *Frontiers in Aging Neuroscience*, 13, 637404.

<https://www.frontiersin.org/article/10.3389/fnagi.2021.637404>

De Roeck, E. E., De Deyn, P. P., Dierckx, E., & Engelborghs, S. (2019). Brief cognitive screening instruments for early detection of Alzheimer's disease: A systematic review. *Alzheimer's Research & Therapy*, 11(1), 21.

<https://doi.org/10.1186/s13195-019-0474-3>

- Dehqan, A., Scherer, R. C., Dashti, G., Ansari-Moghaddam, A., & Fanaie, S. (2012). The Effects of Aging on Acoustic Parameters of Voice. *Folia Phoniatica et Logopaedica*, 64(6), 265-270. <https://doi.org/10.1159/000343998>
- Del Carmen Pérez-Sánchez, M., González-Nosti, M., Cuetos, F., Martínez, C., & Álvarez-Cañizo, M. (2021). Reading Fluency in Spanish Patients with Alzheimer's Disease. *Current Alzheimer Research*, 18(3), 243-255. <https://doi.org/10.2174/1567205018666210608102012>
- Delavaran, H., Jönsson, A.-C., Lökvist, H., Iwarsson, S., Elmståhl, S., Norrving, B., & Lindgren, A. (2017). Cognitive function in stroke survivors: A 10-year follow-up study. *Acta Neurologica Scandinavica*, 136(3), 187-194. <https://doi.org/10.1111/ane.12709>
- DementiaBank*. (s. f.). Recuperado 3 de mayo de 2022, de <https://dementia.talkbank.org/>
- DeSouza, D. D., Robin, J., Gumus, M., & Yeung, A. (2021). Natural Language Processing as an Emerging Tool to Detect Late-Life Depression. *Frontiers in Psychiatry*, 12, 1525. <https://www.frontiersin.org/article/10.3389/fpsy.2021.719125>
- Docio-Fernandez, L., & Garcia-Mateo, C. (2009). Speech Production. En S. Z. Li & A. Jain (Eds.), *Encyclopedia of Biometrics* (pp. 1290-1295). Springer US. https://doi.org/10.1007/978-0-387-73003-5_199
- Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory Into Practice*, 30(3), 165-175. <https://doi.org/10.1080/00405849109543497>
- Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., DeKosky, S. T., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G. B., Fox, N. C., Galasko, D., Habert, M.-O., Jicha, G. A., Nordberg, A., ... Cummings, J. L. (2014). Advancing research diagnostic

- criteria for Alzheimer's disease: The IWG-2 criteria. *The Lancet. Neurology*, 13(6), 614-629. [https://doi.org/10.1016/S1474-4422\(14\)70090-0](https://doi.org/10.1016/S1474-4422(14)70090-0)
- Duffy, J. R., Hanley, H., Utianski, R., Clark, H., Strand, E., Josephs, K. A., & Whitwell, J. L. (2017). Temporal acoustic measures distinguish primary progressive apraxia of speech from primary progressive aphasia. *Brain and Language*, 168, 84-94. <https://doi.org/10.1016/j.bandl.2017.01.012>
- Espinola, C. W., Gomes, J. C., Pereira, J. M. S., & dos Santos, W. P. (2021). Detection of major depressive disorder using vocal acoustic analysis and machine learning—An exploratory study. *Research on Biomedical Engineering*, 37(1), 53-64. <https://doi.org/10.1007/s42600-020-00100-9>
- Ezzati, A., Katz, M. J., Derby, C. A., Zimmerman, M. E., & Lipton, R. B. (2019). Depressive Symptoms Predict Incident Dementia in a Community Sample of Older Adults: Results From the Einstein Aging Study. *Journal of Geriatric Psychiatry and Neurology*, 32(2), 97-103. <https://doi.org/10.1177/0891988718824036>
- Fagherazzi, G., Fischer, A., Ismael, M., & Despotovic, V. (2021). Voice for Health: The Use of Vocal Biomarkers from Research to Clinical Practice. *Digital Biomarkers*, 5(1), 78-88. <https://doi.org/10.1159/000515346>
- Falcone, M., Yadav, N., Poellabauer, C., & Flynn, P. (2013). Using isolated vowel sounds for classification of Mild Traumatic Brain Injury. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 7577-7581. <https://doi.org/10.1109/ICASSP.2013.6639136>
- Fant, G. (1970). *Acoustic Theory of Speech Production*. Walter de Gruyter.
- Faurholt-Jepsen, M., Busk, J., Frost, M., Vinberg, M., Christensen, E. M., Winther, O., Bardram, J. E., & Kessing, L. V. (2016). Voice analysis as an objective state

- marker in bipolar disorder. *Translational Psychiatry*, 6(7), e856-e856.
<https://doi.org/10.1038/tp.2016.123>
- Ferrand, C. T. (2002). Harmonics-to-Noise Ratio: An Index of Vocal Aging. *Journal of Voice*, 16(4), 480-487. [https://doi.org/10.1016/S0892-1997\(02\)00123-6](https://doi.org/10.1016/S0892-1997(02)00123-6)
- Ferris, S. H., & Farlow, M. (2013). Language impairment in Alzheimer's disease and benefits of acetylcholinesterase inhibitors. *Clinical Interventions in Aging*, 8, 1007-1014. <https://doi.org/10.2147/CIA.S39959>
- Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in Older Adults. *Annual Review of Clinical Psychology*, 5(1), 363-389.
<https://doi.org/10.1146/annurev.clinpsy.032408.153621>
- Flanagan, O., Chan, A., Roop, P., & Sundram, F. (2021). Using Acoustic Speech Patterns From Smartphones to Investigate Mood Disorders: Scoping Review. *JMIR MHealth and UHealth*, 9(9), e24352. <https://doi.org/10.2196/24352>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3), 189-198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Forbes-McKay, K., Shanks, M. F., & Venneri, A. (2013). Profiling spontaneous speech decline in Alzheimer's disease: A longitudinal study. *Acta Neuropsychiatrica*, 25(6), 320-327. <https://doi.org/10.1017/neu.2013.16>
- Franco, C. M., Echeverria, M. P., Campos, R., Campayo, J. G., & Lobo, A. (1993). Escalas de ansiedad y depresión de Goldberg: Una guía de entrevista eficaz para la detección del malestar psíquico. *Atención primaria: Publicación oficial de la Sociedad Española de Familia y Comunitaria*, 12(6), 345-349.

- Fraser, K. C., Lundholm Fors, K., Eckerström, M., Öhman, F., & Kokkinakis, D. (2019). Predicting MCI Status From Multimodal Language Data Using Cascaded Classifiers. *Frontiers in Aging Neuroscience, 11*, 205. <https://doi.org/10.3389/fnagi.2019.00205>
- Fraser, K. C., Meltzer, J. A., & Rudzicz, F. (2016). Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *Journal of Alzheimer's Disease: JAD, 49*(2), 407-422. <https://doi.org/10.3233/JAD-150520>
- Fraser, K. C., Rudzicz, F., & Hirst, G. (2016). Detecting late-life depression in Alzheimer's disease through analysis of speech and language. *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology, 1-11*. <http://dx.doi.org/10.18653/v1/W16-0301>
- Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). "Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis". *Autism Research, 10*(3), 384-407. <https://doi.org/10.1002/aur.1678>
- Galasko, D., Bennett, D., Sano, M., Ernesto, C., Thomas, R., Grundman, M., & Ferris, S. (1997). An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. *Alzheimer Disease and Associated Disorders, 11*(Suppl2), S33-S39. <https://doi.org/10.1097/00002093-199700112-00005>
- Gándara, J. de la. (1997). Manejo de depresión y ansiedad en atención primaria. *Atención primaria: Publicación oficial de la Sociedad Española de Familia y Comunitaria, 20*(7), 389-394.
- Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language, 93*(1), 79-94. <https://doi.org/10.1016/j.bandl.2004.08.003>

- Garrard, P., Rentoumi, V., Gesierich, B., Miller, B., & Gorno-Tempini, M. L. (2014). Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 55, 122-129. <https://doi.org/10.1016/j.cortex.2013.05.008>
- Gayraud, F., Lee, H.-R., & Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics & Phonetics*, 25(3), 198-209. <https://doi.org/10.3109/02699206.2010.521612>
- Goldberg, D., Bridges, K., Duncan-Jones, P., & Grayson, D. (1988). Detecting anxiety and depression in general medical settings. *British Medical Journal*, 297(6653), 897-899. <https://doi.org/10.1136/bmj.297.6653.897>
- Gómez-Vilda, P., Gómez-Rodellar, A., Vicente, J. M. F., Mekyska, J., Palacios-Alonso, D., Rodellar-Biarge, V., Álvarez-Marquina, A., Eliasova, I., Kostalova, M., & Rektorova, I. (2019). Neuromechanical Modelling of Articulatory Movements from Surface Electromyography and Speech Formants. *International Journal of Neural Systems*, 29(02), 1850039. <https://doi.org/10.1142/S0129065718500399>
- Gong, Y., Yatawatte, H., Poellabauer, C., Schneider, S., & Latham, S. (2018). Automatic autism spectrum disorder detection using everyday vocalizations captured by smart devices. *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 465-473. <https://doi.org/10.1145/3233547.3233574>
- Gonzalez-Burgos, L., Barroso, J., & Ferreira, D. (2020). Cognitive reserve and network efficiency as compensatory mechanisms of the effect of aging on phonemic fluency. *Aging*, 12(22), 23351-23378. <https://doi.org/10.18632/aging.202177>

- Gordon, J. K., & Chen, H. (2022). How well does the discrepancy between semantic and letter verbal fluency performance distinguish Alzheimer's dementia from typical aging? *Aging, Neuropsychology, and Cognition*, 0(0), 1-30.
<https://doi.org/10.1080/13825585.2022.2079602>
- Gosztolya, G., Balogh, R., Imre, N., Egas-López, J. V., Hoffmann, I., Vincze, V., Tóth, L., Devanand, D. P., Pákáski, M., & Kálmán, J. (2021). Cross-lingual detection of mild cognitive impairment based on temporal parameters of spontaneous speech. *Computer Speech & Language*, 69, 101215.
<https://doi.org/10.1016/j.csl.2021.101215>
- Gosztolya, G., Vincze, V., Tóth, L., Pákáski, M., Kálmán, J., & Hoffmann, I. (2019). Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech & Language*, 53, 181-197. <https://doi.org/10.1016/j.csl.2018.07.007>
- Grabe, E., & Low, E. L. (2008). Durational variability in speech and the Rhythm Class Hypothesis. En *Durational variability in speech and the Rhythm* (pp. 515-546). De Gruyter Mouton. <https://doi.org/10.1515/9783110197105.2.515>
- Grande, M., Meffert, E., Schoenberger, E., Jung, S., Frauenrath, T., Huber, W., Hussmann, K., Moormann, M., & Heim, S. (2012). From a concept to a word in a syntactically complete sentence: An fMRI study on spontaneous language production in an overt picture description task. *NeuroImage*, 61(3), 702-714.
<https://doi.org/10.1016/j.neuroimage.2012.03.087>
- Granena, G., & Yilmaz, Y. (2019). Phonological short-term memory capacity and L2 oral performance. *Journal of Second Language Studies*, 2(2), 317-335.
<https://doi.org/10.1075/jsls.19005.gra>

- Grünerbl, A., Muaremi, A., Osmani, V., Bahle, G., Oehler, S., Tröster, G., Mayora, O., Haring, C., & Lukowicz, P. (2014). Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE Journal of Biomedical and Health Informatics*, *19*(1), 140-148. <https://doi.org/10.1109/jbhi.2014.2343154>
- Guimarães, M. T., Medeiros, A. G., Almeida, J. S., Falcão y Martin, M., Damaševičius, R., Maskeliūnas, R., Cavalcante Mattos, C. L., & Rebouças Filho, P. P. (2020). An Optimized Approach to Huntington's Disease Detecting via Audio Signals Processing with Dimensionality Reduction. *2020 International Joint Conference on Neural Networks (IJCNN)*, 1-8. <https://doi.org/10.1109/IJCNN48605.2020.9206773>
- Hall, A. O., Shinkawa, K., Kosugi, A., Takase, T., Kobayashi, M., Nishimura, M., Nemoto, M., Watanabe, R., Tsukada, E., Ota, M., Higashi, S., Nemoto, K., Arai, T., & Yamada, Y. (2019). Using Tablet-Based Assessment to Characterize Speech for Individuals with Dementia and Mild Cognitive Impairment: Preliminary Results. *AMIA Summits on Translational Science Proceedings, 2019*, 34-43.
- Hamilton, M. (1960). A Rating Scale for Depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*(1), 56-62. <https://doi.org/10.1136/jnnp.23.1.56>
- Hashim, N. W., Wilkes, M., Salomon, R., Meggs, J., & France, D. J. (2017). Evaluation of Voice Acoustics as Predictors of Clinical Depression Scores. *Journal of Voice*, *31*(2), 256.e1-256.e6. <https://doi.org/10.1016/j.jvoice.2016.06.006>
- Haugrud, N., Crossley, M., & Vrbancic, M. (2011). Clustering and Switching Strategies During Verbal Fluency Performance Differentiate Alzheimer's Disease and

- Healthy Aging. *Journal of the International Neuropsychological Society*, 17(6), 1153-1157. <https://doi.org/10.1017/S1355617711001196>
- Haulcy, R., & Glass, J. (2020). Classifying Alzheimer's Disease Using Audio and Text-Based Representations of Speech. *Frontiers in Psychology*, 11, 624137. <https://doi.org/10.3389/fpsyg.2020.624137>
- Hegde, S., Shetty, S., Rai, S., & Dodderi, T. (2019). A Survey on Machine Learning Approaches for Automatic Detection of Voice Disorders. *Journal of Voice*, 33(6), 947.e11-947.e33. <https://doi.org/10.1016/j.jvoice.2018.07.014>
- Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, 42(9), 1212-1222. <https://doi.org/10.1016/j.neuropsychologia.2004.02.001>
- Hernandez, A., Lee, H., & Chung, M. (2019). Acoustic analysis of fricatives in dysarthric speakers with cerebral palsy. *Phonetics and Speech Sciences*, 11(3), 23-29. <https://doi.org/10.13064/KSSS.2019.11.3.023>
- Higuchi, M., Tokuno, S., Nakamura, M., Shinohara, S., Mitsuyoshi, S., Omiya, Y., Hagiwara, N., Takano, T., Toda, H., Saito, T., Terashi, H., & Mitoma, H. (2018). Classification of Bipolar Disorder, Major Depressive Disorder, and Healthy State Using Voice. *Asian Journal of Pharmaceutical and Clinical Research*, 89-93. <https://doi.org/10.22159/ajpcr.2018.v11s3.30042>
- Hodges, J. R., Patterson, K., Graham, N., & Dawson, K. (1996). Naming and Knowing in Dementia of Alzheimer's Type. *Brain and Language*, 54(2), 302-325. <https://doi.org/10.1006/brln.1996.0077>
- Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge?

- Neuropsychologia*, 30(4), 301-314. [https://doi.org/10.1016/0028-3932\(92\)90104-T](https://doi.org/10.1016/0028-3932(92)90104-T)
- Hoffmann, I., Nemeth, D., Dye, C. D., Pákási, M., Irinyi, T., & Kálmán, J. (2010). Temporal parameters of spontaneous speech in Alzheimer's disease. *International Journal of Speech-Language Pathology*, 12(1), 29-34. <https://doi.org/10.3109/17549500903137256>
- Horley, K., Reid, A., & Burnham, D. (2010). Emotional Prosody Perception and Production in Dementia of the Alzheimer's Type. *Journal of Speech, Language, and Hearing Research*, 53(5), 1132-1146. [https://doi.org/10.1044/1092-4388\(2010/09-0030\)](https://doi.org/10.1044/1092-4388(2010/09-0030))
- Huber, J. E., & Spruill, J. (2008). Age-related changes to speech breathing with increased vocal loudness. *Journal of Speech, Language, and Hearing Research: JSLHR*, 51(3), 651-668. [https://doi.org/10.1044/1092-4388\(2008/047\)](https://doi.org/10.1044/1092-4388(2008/047))
- Hughes, S. M., & Rhodes, B. C. (2010). Making age assessments based on voice: The impact of the reproductive viability of the speaker. *Journal of Social, Evolutionary, and Cultural Psychology*, 4(4), 290. <https://doi.org/10.1037/h0099282>
- Hunter, M. B., Jenkins, N., Dolan, C., Pullen, H., Ritchie, C., & Muniz-Terrera, G. (2021). Reliability of Telephone and Videoconference Methods of Cognitive Assessment in Older Adults with and without Dementia. *Journal of Alzheimer's Disease*, 81(4), 1625-1647. <https://doi.org/10.3233/JAD-210088>
- Huttunen, K. H., Keränen, H. I., Pääkkönen, R. J., Päivikki Eskelinen-Rönkä, R., & Leino, T. K. (2011). Effect of cognitive load on articulation rate and formant frequencies during simulator flights. *The Journal of the Acoustical Society of America*, 129(3), 1580-1593. <https://doi.org/10.1121/1.3543948>

- Isaacs, B., & Akhtar, A. J. (1972). The Set Test: A rapid test of mental function in old people. *Age and Ageing*, *1*(4), 222-226. <https://doi.org/10.1093/ageing/1.4.222>
- Ismail, Z., Rajji, T. K., & Shulman, K. I. (2010). Brief cognitive screening instruments: An update. *International Journal of Geriatric Psychiatry*, *25*(2), 111-120. <https://doi.org/10.1002/gps.2306>
- Ivanova, O., Meilán, J. J. G., Martínez-Sánchez, F., Martínez-Nicolás, I., Llorente, T. E., & González, N. C. (2022). Discriminating speech traits of Alzheimer's disease assessed through a corpus of reading task for Spanish language. *Computer Speech & Language*, *73*, 101341. <https://doi.org/10.1016/j.csl.2021.101341>
- Ivanova, O., Meilán, J. J. G., Martínez-Sánchez, F., Rodríguez, E., & Carro, J. (2017). *Psycholinguistics of dementia: What is known about specific language and speech traits of Alzheimer's disease*. Cambridge Scholars Publishing.
- Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., & Sperling, R. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, *14*(4), 535-562. <https://doi.org/10.1016/j.jalz.2018.02.018>
- Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet. Neurology*, *9*(1), 119-128. [https://doi.org/10.1016/S1474-4422\(09\)70299-6](https://doi.org/10.1016/S1474-4422(09)70299-6)

- Jang, H., Soroski, T., Rizzo, M., Barral, O., Harisinghani, A., Newton-Mason, S., Granby, S., Stutz da Cunha Vasco, T. M., Lewis, C., Tutt, P., Carenini, G., Conati, C., & Field, T. S. (2021). Classification of Alzheimer's Disease Leveraging Multi-task Machine Learning Analysis of Speech and Eye-Movement Data. *Frontiers in Human Neuroscience, 15*, 716670. <https://doi.org/10.3389/fnhum.2021.716670>
- Jansen, W. J., Ossenkoppele, R., Knol, D. L., Tijms, B. M., Scheltens, P., Verhey, F. R. J., Visser, P. J., Amyloid Biomarker Study Group, Aalten, P., Aarsland, D., Alcolea, D., Alexander, M., Almdahl, I. S., Arnold, S. E., Baldeiras, I., Barthel, H., van Berckel, B. N. M., Bibeau, K., Blennow, K., ... Zetterberg, H. (2015). Prevalence of cerebral amyloid pathology in persons without dementia: A meta-analysis. *JAMA, 313*(19), 1924-1938. <https://doi.org/10.1001/jama.2015.4668>
- Jessen, F., Wiese, B., Bachmann, C., Eifflaender-Gorfer, S., Haller, F., Kölsch, H., Luck, T., Mösch, E., van den Bussche, H., Wagner, M., Wollny, A., Zimmermann, T., Pentzek, M., Riedel-Heller, S. G., Romberg, H.-P., Weyerer, S., Kaduszkiewicz, H., Maier, W., Bickel, H., & German Study on Aging, Cognition and Dementia in Primary Care Patients Study Group. (2010). Prediction of dementia by subjective memory impairment: Effects of severity and temporal association with cognitive impairment. *Archives of General Psychiatry, 67*(4), 414-422. <https://doi.org/10.1001/archgenpsychiatry.2010.30>
- Jia, Y., Liang, Y., & Zhu, T. (2019). An analysis of voice quality of Chinese patients with depression. *2019 22nd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, 1-6. <https://doi.org/10.1109/O-COCOSDA46868.2019.9060848>

- Jiang, H., Hu, B., Liu, Z., Yan, L., Wang, T., Liu, F., Kang, H., & Li, X. (2017). Investigation of different speech types and emotions for detecting depression using different classifiers. *Speech Communication, 90*, 39-46.
<https://doi.org/10.1016/j.specom.2017.04.001>
- Jorm, A. F. (2000). Is Depression a Risk Factor for Dementia or Cognitive Decline? *Gerontology, 46*(4), 219-227. <https://doi.org/10.1159/000022163>
- Jorm, A. F., & Jacomb, P. A. (1989). The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Socio-demographic correlates, reliability, validity and some norms. *Psychological Medicine, 19*(4), 1015-1022.
<https://doi.org/10.1017/S0033291700005742>
- Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*.
- Kato, S., Homma, A., & Sakuma, T. (2018). Easy Screening for Mild Alzheimer's Disease and Mild Cognitive Impairment from Elderly Speech. *Current Alzheimer Research, 15*(2), 104-110.
<https://doi.org/10.2174/1567205014666171120144343>
- Kauhanen, M.-L., Korpelainen, J. T., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., Sotaniemi, K. A., & Myllylä, V. V. (2000). Aphasia, Depression, and Non-Verbal Cognitive Impairment in Ischaemic Stroke. *Cerebrovascular Diseases, 10*(6), 455-461. <https://doi.org/10.1159/000016107>
- Kavé, G., Heled, E., Vakil, E., & Agranov, E. (2011). Which verbal fluency measure is most useful in demonstrating executive deficits after traumatic brain injury? *Journal of Clinical and Experimental Neuropsychology, 33*(3), 358-365.
<https://doi.org/10.1080/13803395.2010.518703>
- Kemper, S., LaBarge, E., Ferraro, F. R., Cheung, H., Cheung, H., & Storandt, M. (1993). On the Preservation of Syntax in Alzheimer's Disease: Evidence From

- Written Sentences. *Archives of Neurology*, 50(1), 81-86.
<https://doi.org/10.1001/archneur.1993.00540010075021>
- Kempler, D. (1995). Language changes in dementia of the Alzheimer type. *Dementia and communication*, 1, 98-114.
- Khodabakhsh, A., & Demiroglu, C. (2015). Analysis of Speech-Based Measures for Detecting and Monitoring Alzheimer's Disease. En C. Fernández-Llatas & J. M. García-Gómez (Eds.), *Data Mining in Clinical Medicine* (pp. 159-173). Springer. https://doi.org/10.1007/978-1-4939-1985-7_11
- Khodabakhsh, A., Yesil, F., Guner, E., & Demiroglu, C. (2015). Evaluation of linguistic and prosodic features for detection of Alzheimer's disease in Turkish conversational speech. *EURASIP Journal on Audio, Speech, and Music Processing*, 2015(1), 9. <https://doi.org/10.1186/s13636-015-0052-y>
- Kobayashi, T., & Kato, S. (2011). Depression–dementia medius: Between depression and the manifestation of dementia symptoms. *Psychogeriatrics*, 11(3), 177-182. <https://doi.org/10.1111/j.1479-8301.2011.00363.x>
- König, A., Linz, N., Tröger, J., Wolters, M., Alexandersson, J., & Robert, P. (2018). Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task. *Dementia and Geriatric Cognitive Disorders*, 45(3-4), 198-209. <https://doi.org/10.1159/000487852>
- König, A., Linz, N., Zeghari, R., Klinge, X., Tröger, J., Alexandersson, J., & Robert, P. (2019). Detecting Apathy in Older Adults with Cognitive Disorders Using Automatic Speech Analysis. *Journal of Alzheimer's Disease*, 69(4), 1183-1193. <https://doi.org/10.3233/JAD-181033>
- König, A., Mallick, E., Tröger, J., Linz, N., Zeghari, R., Manera, V., & Robert, P. (2021). Measuring neuropsychiatric symptoms in patients with early cognitive

- decline using speech analysis. *European Psychiatry*, 64(1).
<https://doi.org/10.1192/j.eurpsy.2021.2236>
- König, A., Satt, A., Sorin, A., Hoory, R., Derreumaux, A., David, R., & Robert, P. H. (2018). Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People. *Current Alzheimer Research*, 15(2), 120-129. <https://doi.org/10.2174/1567205014666170829111942>
- König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., Manera, V., Verhey, F., Aalten, P., Robert, P. H., & David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(1), 112-124. <https://doi.org/10.1016/j.dadm.2014.11.012>
- Kopp, B., Lange, F., & Steinke, A. (2021). The Reliability of the Wisconsin Card Sorting Test in Clinical Practice. *Assessment*, 28(1), 248-263.
<https://doi.org/10.1177/1073191119866257>
- Kosztyła-Hojna, B., Moskal, D., & Kuryliszyn-Moskal, A. (2015). Parameters of the assessment of voice quality and clinical manifestation of rheumatoid arthritis. *Advances in Medical Sciences*, 60(2), 321-328.
<https://doi.org/10.1016/j.advms.2015.06.004>
- Kourtis, L. C., Regele, O. B., Wright, J. M., & Jones, G. B. (2019). Digital biomarkers for Alzheimer's disease: The mobile/wearable devices opportunity. *Npj Digital Medicine*, 2(1), 1-9. <https://doi.org/10.1038/s41746-019-0084-2>
- Ladd, D. R. (2008). *Intonational Phonology*. Cambridge University Press.
- Lancu, I., & Olmer, A. (2006). The minimental state examination—An up-to-date review. *Harefuah*, 145(9), 687-690, 701.

- Lang, B. H. H., Wong, C. K. H., & Ma, E. P. M. (2016). A systematic review and meta-analysis on acoustic voice parameters after uncomplicated thyroidectomy. *The Laryngoscope*, *126*(2), 528-537. <https://doi.org/10.1002/lary.25452>
- Laske, C., Sohrabi, H. R., Frost, S. M., López-de-Ipiña, K., Garrard, P., Buscema, M., Dauwels, J., Soekadar, S. R., Mueller, S., Linnemann, C., Bridenbaugh, S. A., Kanagasingham, Y., Martins, R. N., & O'Bryant, S. E. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *11*(5), 561-578. <https://doi.org/10.1016/j.jalz.2014.06.004>
- Lauriola, M., Mangiacotti, A., D'Onofrio, G., Cascavilla, L., Paris, F., Ciccone, F., Greco, M., Paroni, G., Seripa, D., & Greco, A. (2018). Late-Life Depression versus Amnesic Mild Cognitive Impairment: Alzheimer's Disease Incidence in 4 Years of Follow-Up. *Dementia and Geriatric Cognitive Disorders*, *46*(3-4), 140-153. <https://doi.org/10.1159/000492489>
- Lawrence, B. J., Jayakody, D. M. P., Bennett, R. J., Eikelboom, R. H., Gasson, N., & Friedland, P. L. (2020). Hearing Loss and Depression in Older Adults: A Systematic Review and Meta-analysis. *The Gerontologist*, *60*(3), e137-e154. <https://doi.org/10.1093/geront/gnz009>
- Le, P. N., Ambikairajah, E., Epps, J., Sethu, V., & Choi, E. H. (2011). Investigation of spectral centroid features for cognitive load classification. *Speech Communication*, *53*(4), 540-551. <https://doi.org/10.1016/j.specom.2011.01.005>
- Le Rhun, E., Richard, F., & Pasquier, F. (2005). Natural history of primary progressive aphasia. *Neurology*, *65*(6), 887-891. <https://doi.org/10.1212/01.wnl.0000175982.57472.84>

- Lee, H., Gayraud, F., Hirsch, F., & Barkat-Defradas, M. (2011). Speech Dysfluencies in Normal and Pathological Aging: A Comparison between Alzheimer Patients and Healthy Elderly Subjects. *ICPhS*, 1174-1177.
- Lee, O., & Redford, M. A. (2015). Verbal and spatial working memory load have similarly minimal effects on speech production. *Proceedings of the International Congress of Phonetic Sciences. International Congress of Phonetic Sciences*, 18, 0798.
- Lee, S., Suh, S. W., Kim, T., Kim, K., Lee, K. H., Lee, J. R., Han, G., Hong, J. W., Han, J. W., Lee, K., & Kim, K. W. (2021). Screening major depressive disorder using vocal acoustic features in the elderly by sex. *Journal of Affective Disorders*, 291, 15-23. <https://doi.org/10.1016/j.jad.2021.04.098>
- Lei, P.-W., & Koehly, L. M. (2003). Linear Discriminant Analysis Versus Logistic Regression: A Comparison of Classification Errors in the Two-Group Case. *The Journal of Experimental Education*, 72(1), 25-49. <https://doi.org/10.1080/00220970309600878>
- Li, Y., Li, P., Yang, Q. X., Eslinger, P. J., Sica, C. T., & Karunanayaka, P. (2017). Lexical-Semantic Search Under Different Covert Verbal Fluency Tasks: An fMRI Study. *Frontiers in Behavioral Neuroscience*, 11, 131. <https://doi.org/10.3389/fnbeh.2017.00131>
- Lindsay, H., Tröger, J., & König, A. (2021). Language Impairment in Alzheimer's Disease-Robust and Explainable Evidence for AD-Related Deterioration of Spontaneous Speech Through Multilingual Machine Learning. *Frontiers in Aging Neuroscience*, 13, 642033. <https://doi.org/10.3389/fnagi.2021.642033>
- Linville, S. E. (2004). The aging voice. *The ASHA leader*, 9(19), 12-21.

- Lira, J. O. de, Ortiz, K. Z., Campanha, A. C., Bertolucci, P. H. F., & Minett, T. S. C. (2011). Microlinguistic aspects of the oral narrative in patients with Alzheimer's disease. *International Psychogeriatrics*, 23(3), 404-412. <https://doi.org/10.1017/S1041610210001092>
- López-de-Ipiña, K., Alonso, J. B., Solé-Casals, J., Barroso, N., Henriquez, P., Faundez-Zanuy, M., Travieso, C. M., Ecay-Torres, M., Martínez-Lage, P., & Eguiraun, H. (2013). On Automatic Diagnosis of Alzheimer's Disease Based on Spontaneous Speech Analysis and Emotional Temperature. *Cognitive Computation*, 7(1), 44-55. <https://doi.org/10.1007/s12559-013-9229-9>
- López-de-Ipiña, K., Alonso, J. B., Solé-Casals, J., Barroso, N., Henriquez, P., Faundez-Zanuy, M., Travieso, C. M., Ecay-Torres, M., Martínez-Lage, P., & Eguiraun, H. (2015). On Automatic Diagnosis of Alzheimer's Disease Based on Spontaneous Speech Analysis and Emotional Temperature. *Cognitive Computation*, 7(1), 44-55. <https://doi.org/10.1007/s12559-013-9229-9>
- López-de-Ipiña, K., Alonso, J.-B., Travieso, C. M., Solé-Casals, J., Eguiraun, H., Faundez-Zanuy, M., Ezeiza, A., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., & Martinez de Lizardui, U. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors (Basel, Switzerland)*, 13(5), 6730-6745. <https://doi.org/10.3390/s130506730>
- López-de-Ipiña, K., Alonso-Hernández, J. B., Solé-Casals, J., Travieso-González, C. M., Ezeiza, A., Faúndez-Zanuy, M., Calvo, P. M., & Beitia, B. (2015). Feature selection for automatic analysis of emotional response based on nonlinear speech modeling suitable for diagnosis of Alzheimer's disease. *Neurocomputing*, 150, 392-401. <https://doi.org/10.1016/j.neucom.2014.05.083>

- López-de-Ipiña, K., Martínez-de-Lizarduy, U., Calvo, P. M., Beitia, B., García-Melero, J., Fernández, E., Ecay-Torres, M., Faundez-Zanuy, M., & Sanz, P. (2018). On the analysis of speech and disfluencies for automatic detection of Mild Cognitive Impairment. *Neural Computing and Applications*, 32(20), 15761-15769. <https://doi.org/10.1007/s00521-018-3494-1>
- López-de-Ipiña, K., Martínez-de-Lizarduy, U., Calvo, P. M., Mekyska, J., Beitia, B., Barroso, N., Estanga, A., Tainta, M., & Ecay-Torres, M. (2018). Advances on Automatic Speech Analysis for Early Detection of Alzheimer Disease: A Non-linear Multi-task Approach. *Current Alzheimer Research*, 15(2), 139-148. <https://doi.org/10.2174/1567205014666171120143800>
- López-de-Ipiña, K., Solé-Casals, J., Eguiraun, H., Alonso, J. B., Travieso, C. M., Ezeiza, A., Barroso, N., Ecay-Torres, M., Martínez-Lage, P., & Beitia, B. (2015). Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach. *Computer Speech & Language*, 30(1), 43-60. <https://doi.org/10.1016/j.csl.2014.08.002>
- Lopez-Otero, P., & Docio-Fernandez, L. (2021). Analysis of gender and identity issues in depression detection on de-identified speech. *Computer Speech & Language*, 65, 101118. <https://doi.org/10.1016/j.csl.2020.101118>
- Lortie, C. L., Thibeault, M., Guitton, M. J., & Tremblay, P. (2015). Effects of age on the amplitude, frequency and perceived quality of voice. *Age (Dordrecht, Netherlands)*, 37(6), 117. <https://doi.org/10.1007/s11357-015-9854-1>
- Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, 5(1), 96-116. <https://doi.org/10.1002/lio2.354>

- Lowell, S. Y., Colton, R. H., Kelley, R. T., & Hahn, Y. C. (2011). Spectral- and Cepstral-Based Measures During Continuous Speech: Capacity to Distinguish Dysphonia and Consistency Within a Speaker. *Journal of Voice*, 25(5), e223-e232. <https://doi.org/10.1016/j.jvoice.2010.06.007>
- Luz, S., Haider, F., de la Fuente, S., Fromm, D., & MacWhinney, B. (2020). Alzheimer's Dementia Recognition through Spontaneous Speech: The ADReSS Challenge. *arXiv preprint, arXiv:2004.06833*
- Luz, S., Haider, F., de la Fuente, S., Fromm, D., & MacWhinney, B. (2021). Detecting cognitive decline using speech only: The ADReSSo Challenge. *arXiv preprint, arXiv:2104.09356*
- Luzzi, S., Baldinelli, S., Ranaldi, V., Fiori, C., Plutino, A., Fringuelli, F. M., Silvestrini, M., Baggio, G., & Reverberi, C. (2020). The neural bases of discourse semantic and pragmatic deficits in patients with frontotemporal dementia and Alzheimer's disease. *Cortex*, 128, 174-191. <https://doi.org/10.1016/j.cortex.2020.03.012>
- MacPherson, M. K. (2019). Cognitive Load Affects Speech Motor Performance Differently in Older and Younger Adults. *Journal of Speech, Language, and Hearing Research*, 62(5), 1258-1277. https://doi.org/10.1044/2018_JSLHR-S-17-0222
- MacPherson, M. K., Abur, D., & Stepp, C. E. (2017). Acoustic Measures of Voice and Physiologic Measures of Autonomic Arousal during Speech as a Function of Cognitive Load. *Journal of Voice*, 31(4), 504.e1-504.e9. <https://doi.org/10.1016/j.jvoice.2016.10.021>
- Mahajan, P., & Baths, V. (2021). Acoustic and Language Based Deep Learning Approaches for Alzheimer's Dementia Detection From Spontaneous Speech.

Frontiers in Aging Neuroscience, 13, 623607.

<https://doi.org/10.3389/fnagi.2021.623607>

Manfredi, C., Lebacqz, J., Cantarella, G., Schoentgen, J., Orlandi, S., Bandini, A., & DeJonckere, P. H. (2017). Smartphones Offer New Opportunities in Clinical Voice Research. *Journal of Voice: Official Journal of the Voice Foundation*, 31(1), 111.e1-111.e7. <https://doi.org/10.1016/j.jvoice.2015.12.020>

Manouilidou, C., Dolenc, B., Marvin, T., & Pirtošek, Z. (2016). Processing complex pseudo-words in mild cognitive impairment: The interaction of preserved morphological rule knowledge with compromised cognitive ability. *Clinical Linguistics & Phonetics*, 30(1), 49-67.

<https://doi.org/10.3109/02699206.2015.1102970>

Maor, E., Sara, J. D., Orbelo, D. M., Lerman, L. O., Levanon, Y., & Lerman, A. (2018). Voice Signal Characteristics Are Independently Associated With Coronary Artery Disease. *Mayo Clinic Proceedings*, 93(7), 840-847.

<https://doi.org/10.1016/j.mayocp.2017.12.025>

Marinelli, C. V., Spaccavento, S., Craca, A., Marangolo, P., & Angelelli, P. (2017). Different Cognitive Profiles of Patients with Severe Aphasia. *Behavioural Neurology*, 2017, e3875954. <https://doi.org/10.1155/2017/3875954>

Marmar, C. R., Brown, A. D., Qian, M., Laska, E., Siegel, C., Li, M., Abu-Amara, D., Tsiartas, A., Richey, C., Smith, J., Knoth, B., & Vergyri, D. (2019). Speech-based markers for posttraumatic stress disorder in US veterans. *Depression and Anxiety*, 36(7), 607-616. <https://doi.org/10.1002/da.22890>

Martinc, M., Haider, F., Pollak, S., & Luz, S. (2021). Temporal Integration of Text Transcripts and Acoustic Features for Alzheimer's Diagnosis Based on

- Spontaneous Speech. *Frontiers in Aging Neuroscience*, 13, 642647.
<https://doi.org/10.3389/fnagi.2021.642647>
- Martínez-Nicolás, I., Llorente, T. E., Ivanova, O., Martínez-Sánchez, F., & Meilán, J. J. G. (2022). Many Changes in Speech through Aging Are Actually a Consequence of Cognitive Changes. *International Journal of Environmental Research and Public Health*, 19(4), 2137. <https://doi.org/10.3390/ijerph19042137>
- Martínez-Nicolás, I., Llorente, T. E., Martínez-Sánchez, F., & Meilán, J. J. G. (2021). Ten Years of Research on Automatic Voice and Speech Analysis of People With Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article. *Frontiers in Psychology*, 12, 645.
<https://doi.org/10.3389/fpsyg.2021.620251>
- Martínez-Sánchez, F., García Meilán, J. J., Pérez, E., Carro, J., & Arana, J. M. (2012). Expressive prosodic patterns in individuals with Alzheimer's disease. *Psicothema*, 24(1), 16-21.
- Martínez-Sánchez, F., Meilán, J. J. G., Carro, J., & Ivanova, O. (2018). A Prototype for the Voice Analysis Diagnosis of Alzheimer's Disease. *Journal of Alzheimer's Disease*, 64(2), 473-481. <https://doi.org/10.3233/JAD-180037>
- Martínez-Sánchez, F., Meilán, J. J. G., García-Sevilla, J., Carro, J., & Arana, J. M. (2013). Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects. *Neurologia (English Edition)*, 28(6), 325-331.
<https://doi.org/10.1016/j.nrl.2012.07.012>
- Martínez-Sánchez, F., Meilán, J. J. G., Vera-Ferrandiz, J. A., Carro, J., Pujante-Valverde, I. M., Ivanova, O., & Carcavilla, N. (2017). Speech rhythm alterations in Spanish-speaking individuals with Alzheimer's disease. *Neuropsychology*,

- Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 24(4), 418-434. <https://doi.org/10.1080/13825585.2016.1220487>
- Martínez-Sánchez, F., Muela-Martínez, J. A., Cortés-Soto, P., Meilán, J. J. G., Ferrándiz, J. A. V., Caparrós, A. E., & Valverde, I. M. P. (2015). Can the Acoustic Analysis of Expressive Prosody Discriminate Schizophrenia? *The Spanish Journal of Psychology*, 18. <https://doi.org/10.1017/sjp.2015.85>
- Maryn, Y., & Weenink, D. (2015). Objective Dysphonia Measures in the Program Praat: Smoothed Cepstral Peak Prominence and Acoustic Voice Quality Index. *Journal of Voice*, 29(1), 35-43. <https://doi.org/10.1016/j.jvoice.2014.06.015>
- Mazzetto de Menezes, K. S., Master, S., Guzman, M., Bortnem, C., & Ramos, L. R. (2014). Differences in acoustic and perceptual parameters of the voice between elderly and young women at habitual and high intensity. *Acta Otorrinolaringologica Espanola*, 65(2), 76-84. <https://doi.org/10.1016/j.otorri.2013.07.009>
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 7(3), 263-269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., Carcavilla, N., & Ivanova, O. (2018). Voice Markers of Lexical Access in Mild Cognitive Impairment and

- Alzheimer's Disease. *Current Alzheimer Research*, 15(2), 111-119.
<https://doi.org/10.2174/1567205014666170829112439>
- Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., López, D. E., Millian-Morell, L., & Arana, J. M. (2014). Speech in Alzheimer's Disease: Can Temporal and Acoustic Parameters Discriminate Dementia? *Dementia and Geriatric Cognitive Disorders*, 37(5-6), 327-334. <https://doi.org/10.1159/000356726>
- Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., Sánchez, J. A., & Pérez, E. (2012). Acoustic markers associated with impairment in language processing in Alzheimer's Disease. *The Spanish Journal of Psychology*, 15(2), 487-494.
https://doi.org/10.5209/rev_sjop.2012.v15.n2.38859
- Meilán, J. J. G., Martínez-Sánchez, F., Martínez-Nicolás, I., Llorente, T. E., & Carro, J. (2020). Changes in the Rhythm of Speech Difference between People with Nondegenerative Mild Cognitive Impairment and with Preclinical Dementia. *Behavioural Neurology*, 2020, 4683573. <https://doi.org/10.1155/2020/4683573>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097.
<https://doi.org/10.1371/journal.pmed.1000097>
- Momeni, M., & Rahmani, M. (2021). Speech signal analysis of alzheimer's diseases in farsi using auditory model system. *Cognitive Neurodynamics*, 15(3), 453-461.
<https://doi.org/10.1007/s11571-020-09644-z>
- Moola, S., Munn, Z., Tufanaru, C., Aromataris, E., Sears, K., Sfetcu, R., Currie, M., Qureshi, R., Mattis, P., & Lisy, K. (2017). Chapter 7: Systematic reviews of etiology and risk. *Joanna briggs institute reviewer's manual. The Joanna Briggs Institute*, 5.

- Morales, M. R., & Levitan, R. (2016). Speech vs. text: A comparative analysis of features for depression detection systems. *2016 IEEE Spoken Language Technology Workshop (SLT)*, 136-143.
<https://doi.org/10.1109/SLT.2016.7846256>
- Morris, E., Chalkidou, A., Hammers, A., Peacock, J., Summers, J., & Keevil, S. (2016). Diagnostic accuracy of (18)F amyloid PET tracers for the diagnosis of Alzheimer's disease: A systematic review and meta-analysis. *European Journal of Nuclear Medicine and Molecular Imaging*, *43*(2), 374-385.
<https://doi.org/10.1007/s00259-015-3228-x>
- Moscovitch, M., & Winocur, G. (1995). Frontal lobes, memory, and aging. *Annals of the New York Academy of Sciences*, *769*, 119-150.
<https://doi.org/10.1111/j.1749-6632.1995.tb38135.x>
- Mueller, K. D., Kosciak, R. L., Hermann, B. P., Johnson, S. C., & Turkstra, L. S. (2018). Declines in Connected Language Are Associated with Very Early Mild Cognitive Impairment: Results from the Wisconsin Registry for Alzheimer's Prevention. *Frontiers in Aging Neuroscience*, *9*, 437.
<https://doi.org/10.3389/fnagi.2017.00437>
- Mueller, K. D., Kosciak, R. L., Turkstra, L. S., Riedeman, S. K., LaRue, A., Clark, L. R., Hermann, B., Sager, M. A., & Johnson, S. C. (2016). Connected Language in Late Middle-Aged Adults at Risk for Alzheimer's Disease. *Journal of Alzheimer's Disease*, *54*(4), 1539-1550. <https://doi.org/10.3233/JAD-160252>
- Murtró, P. (2010). *Bases acústicas de la voz*. XVI Congreso Nacional de la Sociedad Médica Española de Foniatría (SOMEF), Pamplona, España.
- Nagumo, R., Zhang, Y., Ogawa, Y., Hosokawa, M., Abe, K., Ukeda, T., Sumi, S., Kurita, S., Nakakubo, S., Lee, S., Doi, T., & Shimada, H. (2020). Automatic

Detection of Cognitive Impairments through Acoustic Analysis of Speech.

Current Alzheimer Research, 17(1), 60-68.

<https://doi.org/10.2174/1567205017666200213094513>

Nakamura, H., Nakanishi, M., Hamanaka, T., Nakaaki, S., & Yoshida, S. (2000).

Semantic Priming in Patients with Alzheimer and Semantic Dementia. *Cortex*,

36(2), 151-162. [https://doi.org/10.1016/S0010-9452\(08\)70521-5](https://doi.org/10.1016/S0010-9452(08)70521-5)

Nasrolahzadeh, M., Mohammadpoori, Z., & Haddadnia, J. (2016). Analysis of mean

square error surface and its corresponding contour plots of spontaneous speech

signals in Alzheimer's disease with adaptive wiener filter. *Computers in Human*

Behavior, 61, 364-371. <https://doi.org/10.1016/j.chb.2016.03.031>

Nasrolahzadeh, M., Mohammadpoory, Z., & Haddadnia, J. (2016). A novel method for

early diagnosis of Alzheimer's disease based on higher-order spectral estimation

of spontaneous speech signals. *Cognitive Neurodynamics*, 10(6), 495-503.

<https://doi.org/10.1007/s11571-016-9406-0>

Nasrolahzadeh, M., Mohammadpoory, Z., & Haddadnia, J. (2018). Higher-order

spectral analysis of spontaneous speech signals in Alzheimer's disease.

Cognitive Neurodynamics, 12(6), 583-596. [https://doi.org/10.1007/s11571-018-](https://doi.org/10.1007/s11571-018-9499-8)

9499-8

Nelson, P. T., Head, E., Schmitt, F. A., Davis, P. R., Neltner, J. H., Jicha, G. A., Abner,

E. L., Smith, C. D., Van Eldik, L. J., Kryscio, R. J., & Scheff, S. W. (2011).

Alzheimer's disease is not "brain aging": Neuropathological, genetic, and

epidemiological human studies. *Acta neuropathologica*, 121(5), 571-587.

<https://doi.org/10.1007/s00401-011-0826-y>

- Nielson, K. B., & DeKeyser, R. (2019). Working memory and planning time as predictors of fluency and accuracy. *Journal of Second Language Studies*, 2(2), 281-316. <https://doi.org/10.1075/jsls.19004.bro>
- Nikolaev, A., Ashaie, S., Hallikainen, M., Hänninen, T., Higby, E., Hyun, J., Lehtonen, M., & Soininen, H. (2019). Effects of morphological family on word recognition in normal aging, mild cognitive impairment, and Alzheimer's disease. *Cortex*, 116, 91-103. <https://doi.org/10.1016/j.cortex.2018.10.028>
- Noffs, G., Perera, T., Kolbe, S. C., Shanahan, C. J., Boonstra, F. M. C., Evans, A., Butzkueven, H., van der Walt, A., & Vogel, A. P. (2018). What speech can tell us: A systematic review of dysarthria characteristics in Multiple Sclerosis. *Autoimmunity Reviews*, 17(12), 1202-1209. <https://doi.org/10.1016/j.autrev.2018.06.010>
- Núñez Batalla, F., Corte Santos, P., Señaris González, B., Rodríguez Prado, N., & Suárez Nieto, C. (2004). Evaluación espectral cuantitativa de la hipofunción vocal. *Acta Otorrinolaringológica Española*, 55(7), 327-333. [https://doi.org/10.1016/S0001-6519\(04\)78531-3](https://doi.org/10.1016/S0001-6519(04)78531-3)
- Oh, C., Morris, R. J., & Wang, X. (2021). A Systematic Review of Expressive and Receptive Prosody in People With Dementia. *Journal of Speech, Language, and Hearing Research : JSLHR*, 64(10), 3803-3825. https://doi.org/10.1044/2021_JSLHR-21-00013
- O'Hara, R., Thompson, J. M., Kraemer, H. C., Fenn, C., Taylor, J. L., Ross, L., Yesavage, J. A., Bailey, A. M., & Tinklenberg, J. R. (2002). Which Alzheimer Patients Are at Risk for Rapid Cognitive Decline? *Journal of Geriatric Psychiatry and Neurology*, 15(4), 233-238. <https://doi.org/10.1177/089198870201500409>

- Oliveira Santos, A., Godoy, J., Silverio, K., & Brasolotto, A. (2021). Vocal Changes of Men and Women from Different Age Decades: An Analysis from 30 Years of Age. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2021.06.003>
- O'Malley, R. P. D., Mirheidari, B., Harkness, K., Reuber, M., Venneri, A., Walker, T., Christensen, H., & Blackburn, D. (2020). Fully automated cognitive screening tool based on assessment of speech and language. *Journal of Neurology, Neurosurgery, and Psychiatry*, jnnp-2019-322517. <https://doi.org/10.1136/jnnp-2019-322517>
- OMS. (2022). *11 ed.; CIE-11*. <https://icd.who.int/es>
- Organización Mundial de la Salud. (2005). *EHealth*. http://extranet.who.int/iris/bitstream/10665/20378/1/WHA58_28-en.pdf?ua=1
- Orozco-Arroyave, J. R., Hönl, F., Arias-Londoño, J. D., Vargas-Bonilla, J. F., Daqrouq, K., Skodda, S., Rusz, J., & Nöth, E. (2016). Automatic detection of Parkinson's disease in running speech spoken in three different languages. *The Journal of the Acoustical Society of America*, 139(1), 481-500. <https://doi.org/10.1121/1.4939739>
- Özseven, T., Düğenci, M., Doruk, A., & Kahraman, H. I. (2018). Voice traces of anxiety: Acoustic parameters affected by anxiety disorder. *Archives of Acoustics*, 625-636. <https://doi.org/10.24425/aoa.2018.125156>
- Pakhomov, S. V. S., Marino, S. E., Banks, S., & Bernick, C. (2015). Using Automatic Speech Recognition to Assess Spoken Responses to Cognitive Tests of Semantic Verbal Fluency. *Speech Communication*, 75, 14-26. <https://doi.org/10.1016/j.specom.2015.09.010>

- Parola, A., Simonsen, A., Bliksted, V., & Fusaroli, R. (2020). Voice patterns in schizophrenia: A systematic review and Bayesian meta-analysis. *Schizophrenia Research, 216*, 24-40. <https://doi.org/10.1016/j.schres.2019.11.031>
- Pastoriza-Domínguez, P., Torre, I. G., Diéguez-Vide, F., Gómez-Ruiz, I., Geladó, S., Bello-López, J., Ávila-Rivera, A., Matías-Guiu, J. A., Pytel, V., & Hernández-Fernández, A. (2022). Speech pause distribution as an early marker for Alzheimer's disease. *Speech Communication, 136*, 107-117. <https://doi.org/10.1016/j.specom.2021.11.009>
- Peña-Casanova, J., Blesa, R., Aguilar, M., Gramunt-Fombuena, N., Gómez-Ansón, B., Oliva, R., Molinuevo, J. L., Robles, A., Barquero, M. S., Antúnez, C., Martínez-Parra, C., Frank-García, A., Fernández, M., Alfonso, V., Sol, J. M., & NEURONORMA Study Team. (2009). Spanish Multicenter Normative Studies (NEURONORMA Project): Methods and sample characteristics. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 24*(4), 307-319. <https://doi.org/10.1093/arclin/acp027>
- Peters, J. (2017). Stability of long-term average spectrum measures and cepstral peak prominence in connected speech. *Preface & Acknowledgements, 129*.
- Petti, U., Baker, S., & Korhonen, A. (2020). A systematic literature review of automatic Alzheimer's disease detection from speech and language. *Journal of the American Medical Informatics Association: JAMIA, 27*(11), 1784-1797. <https://doi.org/10.1093/jamia/ocaa174>
- Pinto, T. C. C., Machado, L., Bulgacov, T. M., Rodrigues-Júnior, A. L., Costa, M. L. G., Ximenes, R. C. C., & Sougey, E. B. (2019). Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's

- Disease (AD) in the elderly? *International Psychogeriatrics*, *31*(4), 491-504.
<https://doi.org/10.1017/S1041610218001370>
- Pistono, A., Jucla, M., Barbeau, E. J., Saint-Aubert, L., Lemesle, B., Calvet, B., Köpke, B., Puel, M., & Pariente, J. (2016). Pauses During Autobiographical Discourse Reflect Episodic Memory Processes in Early Alzheimer's Disease. *Journal of Alzheimer's Disease*, *50*(3), 687-698. <https://doi.org/10.3233/JAD-150408>
- Pistono, A., Pariente, J., Bézy, C., Lemesle, B., Le Men, J., & Jucla, M. (2019). What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease. *Neuropsychologia*, *124*, 133-143.
<https://doi.org/10.1016/j.neuropsychologia.2018.12.018>
- Poellabauer, C., Yadav, N., Daudet, L., Schneider, S. L., Busso, C., & Flynn, P. J. (2015). Challenges in Concussion Detection Using Vocal Acoustic Biomarkers. *IEEE Access*, *3*, 1143-1160. <https://doi.org/10.1109/ACCESS.2015.2457392>
- Pohar, M., Blas, M., & Turk, S. (2004). Comparison of Logistic Regression and Linear Discriminant Analysis: A Simulation Study. *Metodoloski Zvezki*, *1*(1), 143-161.
- Price, S. E., Kinsella, G. J., Ong, B., Storey, E., Mullaly, E., Phillips, M., Pangnadasa-Fox, L., & Perre, D. (2012). Semantic verbal fluency strategies in amnesic mild cognitive impairment. *Neuropsychology*, *26*(4), 490-497.
<https://doi.org/10.1037/a0028567>
- Pulido, M. L. B., Hernández, J. B. A., Ballester, M. Á. F., González, C. M. T., Mekyska, J., & Smékal, Z. (2020). Alzheimer's disease and automatic speech analysis: A review. *Expert Systems with Applications*, *150*, 113213.
<https://doi.org/10.1016/j.eswa.2020.113213>
- Qiao, Y., Xie, X.-Y., Lin, G.-Z., Zou, Y., Chen, S.-D., Ren, R.-J., & Wang, G. (2020). Computer-assisted speech analysis in mild cognitive impairment and

- Alzheimer's disease: A pilot study from Shanghai, China. *Journal of Alzheimer's Disease*, 75(1), 211-221. <https://doi.org/10.3233/JAD-191056>
- Ramanarayanan, V., Lammert, A. C., Rowe, H. P., Quatieri, T. F., & Green, J. R. (2022). Speech as a Biomarker: Opportunities, Interpretability, and Challenges. *Perspectives of the ASHA Special Interest Groups*, 7(1), 276-283. https://doi.org/10.1044/2021_PERSP-21-00174
- Reitan, R. M. (1986). *Trail Making Test: Manual for administration and scoring*. Reitan Neuropsychology Laboratory.
- Robin, J., Harrison, J. E., Kaufman, L. D., Rudzicz, F., Simpson, W., & Yancheva, M. (2020). Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations. *Digital Biomarkers*, 4(3), 99-108. <https://doi.org/10.1159/000510820>
- Rochon, E., Leonard, C., & Goral, M. (2018). Speech and language production in Alzheimer's disease. *Aphasiology*, 32(1), 1-3. <https://doi.org/10.1080/02687038.2017.1390206>
- Rogers, S. L., & Friedman, R. B. (2008). The underlying mechanisms of semantic memory loss in Alzheimer's disease and semantic dementia. *Neuropsychologia*, 46(1), 12-21. <https://doi.org/10.1016/j.neuropsychologia.2007.08.010>
- Rousseaux, M., Sève, A., Vallet, M., Pasquier, F., & Mackowiak-Cordoliani, M. A. (2010). An analysis of communication in conversation in patients with dementia. *Neuropsychologia*, 48(13), 3884-3890. <https://doi.org/10.1016/j.neuropsychologia.2010.09.026>
- Rusz, J., Cmejla, R., Ruzickova, H., & Ruzicka, E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early

- untreated Parkinson's disease. *The Journal of the Acoustical Society of America*, 129(1), 350-367. <https://doi.org/10.1121/1.3514381>
- Sakin, B. (2021). Pragmatic Language Disorders Resulting from Semantic Degradation in Patients with Alzheimer's Disease. *Clinical and Experimental Health Sciences*, 11(3), 523-528. <https://doi.org/10.33808/clinexphealthsci.840917>
- Salehi, M., Reisi, M., & Ghasisin, L. (2017). Lexical Retrieval or Semantic Knowledge Which One Causes Naming Errors in Patients with Mild and Moderate Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders Extra*, 7(3), 419-429. <https://doi.org/10.1159/000484137>
- Salhi, L., & Cherif, A. (2006). *A speech processing interface for analysis of pathological voices*. 2nd International Conference on Information Communication Technologies, Damasco.
- Sanchez, M. H., Vergyri, D., Ferrer, L., Richey, C., Garcia, P., Knoth, B., & Jarrold, W. (2011). Using prosodic and spectral features in detecting depression in elderly males. *Twelfth Annual Conference of the International Speech Communication Association*. Florence, Italy.
- Santos, J. F., Brosh, N., Falk, T. H., Zwaigenbaum, L., Bryson, S. E., Roberts, W., Smith, I. M., Szatmari, P., & Brian, J. A. (2013). Very early detection of Autism Spectrum Disorders based on acoustic analysis of pre-verbal vocalizations of 18-month old toddlers. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 7567-7571. <https://doi.org/10.1109/ICASSP.2013.6639134>
- Schötz, S. (2005). Stimulus Duration and Type in Perception of Female and Male Speaker Age. *Proceedings of Interspeech 2005*. INTERSPEECH 2005 -

Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal.

- Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, 5, 772. <https://doi.org/10.3389/fpsyg.2014.00772>
- Shimoda, A., Li, Y., Hayashi, H., & Kondo, N. (2021). Dementia risks identified by vocal features via telephone conversations: A novel machine learning prediction model. *PloS One*, 16(7), e0253988. <https://doi.org/10.1371/journal.pone.0253988>
- Silagi, M. L., Bertolucci, P. H. F., & Ortiz, K. Z. (2015). Naming ability in patients with mild to moderate Alzheimer's disease: What changes occur with the evolution of the disease? *Clinics*, 70, 423-428. [https://doi.org/10.6061/clinics/2015\(06\)07](https://doi.org/10.6061/clinics/2015(06)07)
- Singh, S., Bucks, R. S., & Cuerden, J. M. (2001). Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology*, 15(6), 571-583. <https://doi.org/10.1080/02687040143000041>
- Skoog Waller, S., Eriksson, M., & Sörqvist, P. (2015). Can you hear my age? Influences of speech rate and speech spontaneity on estimation of speaker age. *Frontiers in Psychology*, 6, 978. <https://doi.org/10.3389/fpsyg.2015.00978>
- Smith, M., Dietrich, B. J., Bai, E., & Bockholt, H. J. (2020). Vocal pattern detection of depression among older adults. *International Journal of Mental Health Nursing*, 29(3), 440-449. <https://doi.org/10.1111/inm.12678>
- Solana-Lavalle, G., & Rosas-Romero, R. (2021). Analysis of voice as an assisting tool for detection of Parkinson's disease and its subsequent clinical interpretation. *Biomedical Signal Processing and Control*, 66, 102415. <https://doi.org/10.1016/j.bspc.2021.102415>

- Solé-Casals, J., Munteanu, C., Martín, O. C., Barbé, F., Queipo, C., Amilibia, J., & Durán-Cantolla, J. (2014). Detection of severe obstructive sleep apnea through voice analysis. *Applied Soft Computing*, 23, 346-354.
<https://doi.org/10.1016/j.asoc.2014.06.017>
- Strauss, E., Strauss, P. of P. E., Sherman, N. and A. A. P. D. of P. and C. N. E. M. S., Sherman, E. M. S., Spreen, O., & Spreen, B. P. of P. O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. Oxford University Press.
- Strimbu, K., & Tavel, J. A. (2010). What are Biomarkers? *Current opinion in HIV and AIDS*, 5(6), 463-466. <https://doi.org/10.1097/COH.0b013e32833ed177>
- Surendran, D., & Levow, G.-A. (2008). Can voice quality improve mandarin tone recognition? *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 4177-4180. Las Vegas, USA.
- Suzuki, Y. (2021). Individual differences in memory predict changes in breakdown and repair fluency but not speed fluency: A short-term fluency training intervention study. *Applied Psycholinguistics*, 42(4), 969-995.
<https://doi.org/10.1017/S0142716421000187>
- Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J., & Pakaski, M. (2015). Speaking in Alzheimer's Disease, is That an Early Sign? Importance of Changes in Language Abilities in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, 7, 195. <https://doi.org/10.3389/fnagi.2015.00195>
- Taguchi, T., Tachikawa, H., Nemoto, K., Suzuki, M., Nagano, T., Tachibana, R., Nishimura, M., & Arai, T. (2018). Major depressive disorder discrimination using vocal acoustic features. *Journal of Affective Disorders*, 225, 214-220.
<https://doi.org/10.1016/j.jad.2017.08.038>

- Taler, V., & Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology*, *30*(5), 501-556.
<https://doi.org/10.1080/13803390701550128>
- Tang, F., Chen, J., Dodge, H. H., & Zhou, J. (2021). The Joint Effects of Acoustic and Linguistic Markers for Early Identification of Mild Cognitive Impairment. *Frontiers in Digital Health*, *3*, 702772.
<https://doi.org/10.3389/fdgth.2021.702772>
- Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*, *9*, 1112-1122.
<https://doi.org/10.1016/j.protcy.2013.12.124>
- Thapa, D. K., Visentin, D. C., Kornhaber, R., & Cleary, M. (2020). Prevalence and factors associated with depression, anxiety, and stress symptoms among older adults: A cross-sectional population-based study. *Nursing & Health Sciences*, *22*(4), 1139-1152. <https://doi.org/10.1111/nhs.12783>
- Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2018). Identification of Mild Cognitive Impairment from Speech in Swedish Using Deep Sequential Neural Networks. *Frontiers in Neurology*, *9*, 975.
<https://doi.org/10.3389/fneur.2018.00975>
- Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2020). Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. *PloS One*, *15*(7), e0236009.
<https://doi.org/10.1371/journal.pone.0236009>
- Thomas, J. A., Burkhardt, H. A., Chaudhry, S., Ngo, A. D., Sharma, S., Zhang, L., Au, R., & Hosseini Ghomi, R. (2020). Assessing the utility of language and voice

- biomarkers to predict cognitive impairment in the Framingham Heart Study cognitive aging cohort data. *Journal of Alzheimer's Disease*, 76(3), 905-922.
<https://doi.org/10.3233/JAD-190783>
- Toth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatloczki, G., Banreti, Z., Pakaski, M., & Kalman, J. (2018). A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech. *Current Alzheimer Research*, 15(2), 130-138.
<https://doi.org/10.2174/1567205014666171121114930>
- Tremblay, P., Deschamps, I., & Dick, A. S. (2019). Neuromotor organization of speech production. En Zubicaray G. I. & Schiller, O. (Eds.) *The Oxford Handbook of Neurolinguistics*. Oxford Handbooks.
<https://doi.org/10.1093/oxfordhb/9780190672027.013.15>
- Tremblay, P., Poulin, J., Martel-Sauvageau, V., & Denis, C. (2019). Age-related deficits in speech production: From phonological planning to motor implementation. *Experimental Gerontology*, 126, 110695.
<https://doi.org/10.1016/j.exger.2019.110695>
- Tröger, J., Linz, N., König, A., Robert, P., Andersson, J., Peter, J., & Kray, J. (2019). Exploitation vs. Exploration—Computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease. *Neuropsychologia*, 131, 53-61.
<https://doi.org/10.1016/j.neuropsychologia.2019.05.007>
- Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, and M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society*, 4(2), 137-143.
<https://doi.org/10.1017/S1355617798001374>

- Van Segbroeck, M., Travadi, R., Vaz, C., Kim, J., Black, M. P., Potamianos, A., & Narayanan, S. S. (2014). Classification of cognitive load from speech using an i-vector framework. *Interspeech*, 751-755.
- Vanello, N., Guidi, A., Gentili, C., Werner, S., Bertschy, G., Valenza, G., ... & Scilingo, E. P. (2012). Speech analysis for mood state characterization in bipolar patients. In *2012 annual international conference of the IEEE Engineering in Medicine and Biology Society* (pp. 2104-2107). IEEE.
<https://doi.org/10.1109/EMBC.2012.6346375>
- Vaughan, R. M., Coen, R. F., Kenny, R., & Lawlor, B. A. (2018). Semantic and Phonemic Verbal Fluency Discrepancy in Mild Cognitive Impairment: Potential Predictor of Progression to Alzheimer's Disease. *Journal of the American Geriatrics Society*, 66(4), 755-759. <https://doi.org/10.1111/jgs.15294>
- Vicsi, K., Sztahó, D., & Kiss, G. (2012). Examination of the sensitivity of acoustic-phonetic parameters of speech to depression. *2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*, 511-515. Kosice, Slovakia
- Victor, C., Scambler, S., Bond, J., & Bowling, A. (2000). Being alone in later life: Loneliness, social isolation and living alone. *Reviews in Clinical Gerontology*, 10(4), 407-417. <https://doi.org/10.1017/S0959259800104101>
- Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, 30(4), 1414-1432. <https://doi.org/10.1016/j.neuroimage.2005.11.002>
- Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., Szoëke, C., Macaulay, S. L., Martins, R., Maruff, P., Ames, D., Rowe, C. C.,

- Masters, C. L., & Australian Imaging Biomarkers and Lifestyle (AIBL) Research Group. (2013). Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *The Lancet. Neurology*, *12*(4), 357-367. [https://doi.org/10.1016/S1474-4422\(13\)70044-9](https://doi.org/10.1016/S1474-4422(13)70044-9)
- Vincze, V., Szatlóczki, G., Tóth, L., Gosztolya, G., Pákási, M., Hoffmann, I., & Kálmán, J. (2021). Telltale silence: Temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clinical Linguistics & Phonetics*, *35*(8), 727-742. <https://doi.org/10.1080/02699206.2020.1827043>
- Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, *25*(7-9), 905-945. <https://doi.org/10.1080/01690961003589492>
- Wang, T., Hong, Y., Wang, Q., Su, R., Ng, M. L., Xu, J., Wang, L., & Yan, N. (2021). Identification of Mild Cognitive Impairment Among Chinese Based on Multiple Spoken Tasks. *Journal of Alzheimer's Disease*, *82*(1), 185-204. <https://doi.org/10.3233/JAD-201387>
- Weakley, A., & Schmitter-Edgecombe, M. (2014). Analysis of Verbal Fluency Ability in Alzheimer's Disease: The Role of Clustering, Switching and Semantic Proximities. *Archives of Clinical Neuropsychology*, *29*(3), 256-268. <https://doi.org/10.1093/arclin/acu010>
- Weiner, M. F., Neubecker, K. E., Bret, M. E., & Hynan, L. S. (2008). Language in Alzheimer's disease. *The Journal of Clinical Psychiatry*, *69*(8), 1223-1227. <https://doi.org/10.4088/jcp.v69n0804>

- West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272-292. <https://doi.org/10.1037/0033-2909.120.2.272>
- Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal Fluency: Language or Executive Function Measure? *Applied Neuropsychology. Adult*, *23*(1), 29-34. <https://doi.org/10.1080/23279095.2015.1004574>
- Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M. G., Sterne, J. A. C., Bossuyt, P. M. M., & QUADAS-2 Group. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, *155*(8), 529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., Nordberg, A., Bäckman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., De Leon, M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., ... Petersen, R. c. (2004). Mild cognitive impairment – beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine*, *256*(3), 240-246. <https://doi.org/10.1111/j.1365-2796.2004.01380.x>
- Wutzler, A., Becker, R., Lämmle, G., Haverkamp, W., & Steinhagen-Thiessen, E. (2013). The anticipatory proportion as an indicator of language impairment in early-stage cognitive disorder in the elderly. *Dementia and Geriatric Cognitive Disorders*, *36*(5-6), 300-309. <https://doi.org/10.1159/000350808>
- Xue, C., Karjadi, C., Paschalidis, I. C., Au, R., & Kolachalama, V. B. (2021). Detection of dementia on voice recordings using deep learning: A Framingham Heart

- Study. *Alzheimer's Research & Therapy*, 13(1), 146.
<https://doi.org/10.1186/s13195-021-00888-3>
- Yamada, Y., Shinkawa, K., Kobayashi, M., Caggiano, V., Nemoto, M., Nemoto, K., & Arai, T. (2021). Combining Multimodal Behavioral Data of Gait, Speech, and Drawing for Classification of Alzheimer's Disease and Mild Cognitive Impairment. *Journal of Alzheimer's Disease : JAD*, 84(1), 315-327.
<https://doi.org/10.3233/JAD-210684>
- Yamada, Y., Shinkawa, K., Kobayashi, M., Nishimura, M., Nemoto, M., Tsukada, E., Ota, M., Nemoto, K., & Arai, T. (2021). Tablet-Based Automatic Assessment for Early Detection of Alzheimer's Disease Using Speech Responses to Daily Life Questions. *Frontiers in Digital Health*, 3.
<https://www.frontiersin.org/article/10.3389/fdgth.2021.653904>
- Yap, T. F., Epps, J., Ambikairajah, E., & Choi, E. H. (2011). Voice source features for cognitive load classification. *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5700-5703. Praga, República Checa.
<https://doi.org/10.1109/ICASSP.2011.5947654>
- Yap, T. F., Epps, J., Ambikairajah, E., & Choi, E. H. C. (2015). Voice source under cognitive load: Effects and classification. *Speech Communication*, 72, 74-95.
<https://doi.org/10.1016/j.specom.2015.05.007>
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of psychiatric research*, 17(1), 37-49.
[https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)
- Yeung, A., Iaboni, A., Rochon, E., Lavoie, M., Santiago, C., Yancheva, M., Novikova, J., Xu, M., Robin, J., Kaufman, L. D., & Mostafa, F. (2021). Correlating natural

language processing and automated speech analysis with clinician assessment to quantify speech-language changes in mild cognitive impairment and Alzheimer's dementia. *Alzheimer's Research & Therapy*, 13(1), 109.

<https://doi.org/10.1186/s13195-021-00848-x>

Yi-Hsiu, L. (2017). Language Processing of Seniors with Alzheimer's Disease: From the Perspective of Temporal Parameters. *International Journal of Cognitive and Language Sciences*, 11(7), 425-430.

Zellner-Keller, B. (2007). "Comment est-ce qu'on dit?" Vieillesse et manque de mot en conversation. En *Nouveaux cahiers de linguistique française*. Librairie Droz.

Zhang, L., Duvvuri, R., Chandra, K. K. L., Nguyen, T., & Ghomi, R. H. (2020). Automated voice biomarkers for depression symptoms using an online cross-sectional data collection initiative. *Depression and Anxiety*, 37(7), 657-669.

<https://doi.org/10.1002/da.23020>

Zwirner, P., Murry, T., & Woodson, G. E. (1991). Phonatory function of neurologically impaired patients. *Journal of Communication Disorders*, 24(4), 287-300.

[https://doi.org/10.1016/0021-9924\(91\)90004-3](https://doi.org/10.1016/0021-9924(91)90004-3)