

Transferring Web Accessibility through Localization and Internationalization Standards

Jesús Torres-del-Rey

Department of Translation and Interpreting
Faculty of Translation and Documentation
University of Salamanca
Cod.eX Research Group
jtorres@usal.es

Lucía Morado Vázquez

Department of Translation Technology
Faculty of Translation and Interpreting
University of Geneva
Cod.eX Research Group
lucia.morado@unige.ch

This is the **author's accepted manuscript version** of “Transferring Web Accessibility through Localization and Internationalization Standards”, *The Journal of Internationalization and Localization*, Vol. 6:1, pp. 1-24, DOI: <https://doi.org/10.1075/jial.19002.tor>. The document is under copy right and the publisher should be contacted for permission to re-use or reprint the material in any form.

Abstract

In recent years, translation and localization studies have started to include accessibility, and web accessibility in particular, as one of the key aspects to take into consideration when adapting a web product to another language and culture (web localization). This paper provides a comprehensive insight of the connections between these two fields and concepts and, above all, it discusses the possibility of transferring accessibility throughout the localization process. In particular, we analyze how the use of current localization and internationalization data exchange standards is connected to this notion of transfer, and how those standards may be capable of transferring accessibility qualities or information, or supporting localizers in their task. Finally, we present an analytical and pragmatic approach to explore this transferability challenge, which includes the study of the techniques proposed by the W3C to help web developers fulfil a set of success criteria that are included in the latest Web Content Accessibility Guidelines (WCAG) 2.1, in relation to their possible integration in localization and internationalization standards.

Keywords

Web accessibility, localization, internationalization, standards, WCAG guidelines, XLIFF, ITS

1 Introduction

Since the first decade of the 21st Century, web accessibility has attracted more and more interest in localization, particularly after the publication of the Web Content Accessibility Guidelines (WCAG) 2.0 as a W3C Recommendation in December 2008, which later became an ISO standard (ISO/IEC 40500:2012). For example, prominent members of SIDAR, who produced a Spanish translation candidate of the WCAG 2.0 guidelines (Fundación SIDAR 2009), published a seminal article in 2010, suggesting which of the 12 guidelines would be the most relevant for translation and localization (Gutiérrez y Restrepo and Martínez Normand 2010). On 5 June, 2018, the new WCAG 2.1 were officially published as a W3C Recommendation, including an additional guideline (13 in all), among other novelties.

In the web translation and localization field, accessibility has traditionally been associated with key issues such as internationalization and localization challenges (Ó Broin 2004), cultural and multimedia aspects (Tercedor 2010), and localization quality (Rodríguez Vázquez and Torres-del-Rey 2014), and it has even been assimilated with interlinguistic (Ó Broin 2004) or intercultural (Torres-del-Rey and Rodríguez Vázquez 2013) translation. In general, as is usually the case with localization, accessibility has been regarded as something that straddles both functionality-related and textual-pragmatic issues (Jiménez Crespo 2013, 127, figure 5.2), and which goes hand in hand with usability aspects (ibid., 135, figure 6.1).

A few years ago, members of the Cod.eX Research Group (University of Salamanca and University of Geneva), which later gave rise to the ALMA Research Team with members from other universities, started approaching web accessibility as an integral part of the teaching and practice of web localization (Torres-del-Rey and Rodríguez Vázquez 2016; Torres-del-Rey et al 2018).¹ The present paper springs from these teaching and research interests, and introduces a new aspect to the relationship between localization and web accessibility: the connection with localization data exchange standards.

In the following sections, we will present the theoretical foundations as well as a replicable methodology that will allow us to study not only the aforementioned connection, but also the possibility of enriching current processes, standards and tools to integrate accessibility more efficiently in the web localizer's task. The current paper presents preliminary insights from an exploratory analysis of a selected set of accessibility issues in relation to localization standards and the localization process. Forthcoming projects and papers will include in-depth analyses, tests and results.

2 Background and Motivation

2.1 Content and Affordances in Web Accessibility and Localization

Web accessibility can generally be understood as making sure that all users — functionally diverse users, but, in general, all kinds of users — can access, understand and use the interactive content available on websites and in other digital products, as well as participate in its development. "Content" is meant here not only as textual information but also as other forms of multimodal signs or non-verbal language (such as color, auditory signs, layout, use of white space, images, table organization, timing and tempo of events like alerts or tooltips), and also, crucially, action and interaction provided through the ever-evolving modes of web technology.²

Interactive content and the affordances of any objects and artefacts on websites also need to be meaningfully communicated in order to be understood and successfully used, through a variety of access modes.³ Take, for instance, a dropdown menu: unless clearly

¹ ALMA in the context of our research project stands for "Approaching Localisation by Means of Accessibility".

² See Jiménez Crespo (2013, 47), where he widens "the definition of text" to include "all signals used in a communication situation", which "incorporates visual, sound, interactive or typographic aspects".

³ "The concept of affordances originates from ecological psychology (...) to denote action possibilities provided to the actor by the environment" (Kaptelinin 2014: v) and by the objects, according to users' capabilities. "A chair affords ('is for') support and, therefore, affords sitting. Most chairs can also be carried by a single person (they afford lifting), but some can only be lifted by a strong person or by a team of people. If young or relatively weak people cannot lift a chair, then for these people, the chair does not have that affordance, it does not afford lifting" (Norman 2013: 11). Norman's use of the term (ibid.) in the field of human-computer interaction (HCI) has been mistaken as only meaning *perceived* action possibilities.

indicated through styles or other visual, auditory or textual techniques, it may not be clear whether the parent element has a submenu or not, or whether the parent element is linked to a specific web page, or is just a mechanism to drop down the submenu. Also, a typical "X" close button is meaningless for visually impaired people, without any further textual information. Last but not least, whether certain fields are required or not in an input form, and what the appropriate format of those fields should be, together with any perceivable and understandable information that may prevent or remedy errors in the form, are typical usability issues, which can become extremely challenging for people with disabilities. Bad communication of possible interaction or affordances can result in despair and annoyance, or in users quitting the website altogether for good.

In localization, web content needs to be adapted to a specific language and culture, which is also by definition a way of "accessibilizing" content: making it available for users who do not have access to the source content because they do not understand the language and/or the culture from which the product has originated (Ó Broin 2004). If the source content has been designed with accessibility in mind, the localized content should, at the very least, maintain that quality and adapt it to the target locale, culture, language and environment.

2.2 Accessibility Guidelines

The WCAG explain "(...) how to make Web content more accessible to people with disabilities" (2018a). They have been developed by the WAI (Web Accessibility Initiative) of the W3C, and provide a set of 13 accessibility guidelines in its 2.1 version, which are organized around four principles: Perceivable, Operable, Understandable and Robust. They are also complemented by 78 success criteria, categorized in three different levels of accessibility conformance (A, AA, AAA; being AAA the highest level) (WAI 2018c).

In order to help developers achieve the success criteria of each of the 13 guidelines, the WAI provides a set of informative techniques that can be publicly consulted.⁴ The techniques are divided into three categories: sufficient techniques ("reliable ways to meet the success criteria"), advisory techniques ("suggested ways to improve accessibility") and failures ("things that cause accessibility barriers and fail specific success criteria") (W3C 2016c). Furthermore, techniques can be general (applied "to all technologies") or technology-specific (ibid.): the initial letters in the name of the techniques refer to this distinction: e.g. G1 would refer to a general technique whereas H1 would refer to an HTML-related technique. The list of techniques can be enlarged, as the W3C encourages people to submit new ones (ibid.).

It should be understood that a failure to apply the documented sufficient or advisory techniques does not automatically mean that the success criteria associated with them will not be achieved. Different techniques can be used to make the same content accessible, in an alternative or complementary way. Moreover, there might be other means of making the content accessible that are not included in the list of techniques, which is always highlighted for each described technique. As new techniques are developed, those whose job it is to analyze the content in order to adjust or transform it (testers, localizers, etc.) need to be especially aware of their associated accessibility principles and guidelines.

However, in order to avoid that, Norman has recently started to use the term "signifier" for this latter meaning.

⁴ The official URL for version 2.1 is <https://www.w3.org/WAI/WCAG21/Techniques/#introduction>.

This will allow localizers to interpret the success or failure of those new techniques while transforming them linguistically and culturally into a new, more accessible product. Lastly, it should be noted that however useful and informative techniques might be, they are not intended to be a means to determine accessibility conformance; only the success criteria (to which the techniques are associated) can be used to that effect (W3C 2016c). In Figure 1, we include a graphic designed by the WAI (2018b) that condenses the informative documents that they have published in four different boxes. Our interest in the research lies in the box labelled "Techniques for WCAG", which includes instructions for developers.

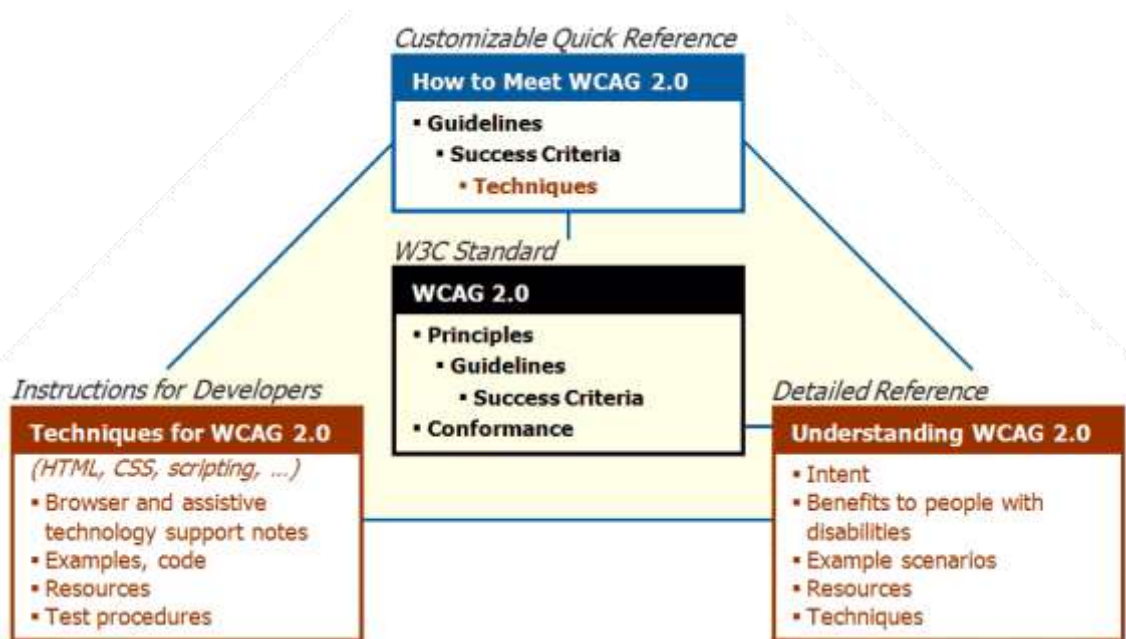


Figure 1. Different documents produced by the WAI Initiative. Source of the image: <https://www.w3.org/WAI/standards-guidelines/wcag/docs/>.

2.3 Multilingual Information Exchange Standards

Information exchange standards are conventions that aid tools and agents in preparing and exchanging content through the different processes they undertake. Specific standards have been developed to facilitate translation and localization tasks. In our current study, we decided to focus on two of them: ITS (Internationalization Tag Set) and XLIFF (XML Localisation Interchange File Format). These were selected because they can be directly used in the preparation of original material (ITS), or during its transformation through the localization process (XLIFF).

ITS 2.0 (the latest version of this standard) defines a series of categories (such as: translatability; terminology; allowed number of characters; locale-specific information; relation to external resources; etc.) that are important when annotating original content to facilitate human translation and other tasks in the multiple "multilingual content production phases". The comprehensive content globalization model targeted by ITS includes:

- Internationalization; Pre-production (e.g. related to marking terminology);
- Automated content enrichment (e.g. automatic hyperlinking for entities);
- Extraction/filtering of translation-relevant content; Segmentation; Leveraging

(e.g. of existing translation-related assets such as translation memories); Machine Translation (e.g. geared towards a specific domain); Quality assessment or control of source language or target language content; Generation of translation kits (e.g. packages based on XLIFF); Post-production; Publishing. (W3C 2013)

XLIFF 2.1 is a format that allows for the interchange of localizable data. Its implementation in the localization process is based on the extraction-merge mechanism: localizable content (mainly textual information) is extracted from the original content and placed in this intermediate format, which is later manipulated (translated and/or adapted), usually in translation or localization management systems, or CAT (Computer-Assisted Translation) tools; finally, the modified content is merged back with the original format to recreate the target content (Torres-del-Rey and Morado Vázquez 2015). The structure of the latest versions of the standard (XLIFF 2.0 and 2.1) is divided into two main components: the core and the modules. The core contains the basic information that is needed to create a minimal XLIFF document that can store extracted localizable data. The modules, on the other hand, are optional and allow for the inclusion of complementary information on the content and localization process, such as Translation matches, Glossary, or Size and Length Restriction. In the current version of the standard (2.1), there is also a specialized ITS module that regulates how to define ITS categories by means of XLIFF inline annotations, attributes and elements.

Our study brings both worlds together: on the one hand, we have web accessibility standards that help create and evaluate accessible content; and, on the other, we have internationalization and localization standards that provide a method to define and unify how we produce and exchange content in a localization-friendly way. In the following sections, we will present and analyze our research question: how can those two worlds work together in web localization processes to help localizers preserve or enhance accessibility in their work?⁵

3 Theoretical Approach

3.1 Language- and culture-bound access barriers and pathways

Accessibility of a web product can be achieved through various information techniques at several levels, from more general to more specific:

- overall *structural and interaction design*, anticipating various access needs and preferences;
- semantic and programmatic *coding* of content, layout and aesthetics, according to previously identified access needs, preferences, and the ensuing structural and interaction design;
- *multimodal meaning production*: language, textual strategies and other semiotic resources (e.g. icons, colors, user interface elements), including interactions and overlaps with the previous two levels.

⁵ We understand localization as a broad activity (along the lines of the "broad definition of translation" in Koby et al. 2014) which includes not only interactive digital content translation, but also technical and cultural adjustments, testing, and multilingual content and project management. In the same vein, in this article, we target all the specialized localizer roles defined by Jiménez-Crespo (2013, 174-175, 179-182) when considering their responsibility towards accessibility: localization experts or translators, localization engineers, managers, and so on.

Each of the above levels is, to a certain extent, dependent on language (e.g. how the structure and interaction affordances are communicated) and culture (e.g. what structures and forms of interaction are typical in a certain community, or how any semiotic resources and design spaces are interpreted according to different expectations and cultural meanings of their components). This is obvious in highly textually- or culturally-dependent features such as text alternatives, page titles, abbreviations, reading level, subtitles, sign language, and so on, but it can also be true for more technical aspects of accessibility.

Take, for instance, the WCAG 2.1 requirement that all content must be navigable by keyboard (guideline 2.1), which seems an exclusively technical matter: the logical sequence or reading order of focusable elements (i.e. by default, interactive elements, plus those with a `tabindex` attribute) is important for users with visual and physical disabilities, who might need to jump from one focusable element to the next as their more natural way of browsing a website. If this is not taken into account, the sequence of text and accessible names of those focusable elements, when standing alone, may become nonsensical in certain languages; furthermore, certain rearrangements may need to be made in right-to-left and bi-directional languages. In another example, the initial bypass blocks ("skip to main content") mechanism (WCAG 2.1 success criterion 2.4.1) may need to be linked to a different target section in different locales, in very dynamic websites with frequent updates. Finally, text and non-text contrast (success criteria 1.4.3, 1.4.6 and 1.4.11) would have to be verified again when colors are changed in localization for cultural reasons.

Content, as defined earlier, needs to be *transferred* and adapted for a new community of users, through a technological medium, and across languages and cultures—in the process known as localization. If content is—at the same time or in synergetic terms—linguistically, culturally, semantically, pragmatically and technically accessible for source locale users, those synergies must be reconstructed in a different locale, with different expectations and experiences among communities of people with disabilities. It is therefore logical to assume that success in localizing accessible content depends on how that content is transferred and, if necessary, adapted and transformed linguistically, culturally and technically in relation to its surrounding context, and linguistic, cultural, semantic, pragmatic and technical features, and according to target users and use environment.

3.2 Transferring accessible content in Localization

The "transportation" or "transfer" metaphor is among the most commonly used in translation studies, whether it is meaning or textual material that gets transferred and transformed in the translation process (Munday 2009, 18-19; Pym 2004, 12-14). In localization, language and text are intertwined and co-occur with other logical, functional and semiotic resources: signifiers of affordances (i.e. action and interaction potential), guidance and feedback on them, input methods, interaction results, symbols, multimedia, layout and presentation of objects and text, to name just a few examples.

Well-designed, internationalized products separate logic (computer code) from user interface (language and other semiotic resources); as well as content and structure from presentation. However, for the user, all these elements need to work and communicate together effectively in order for the product to be successful (useful, accessible, usable and providing a good experience) (Nielsen 2012). This duality—on the one hand, appropriate infra-structural separation of elements, and on the other, successful combination of them on the surface and in use—defines and conditions the work of

localizers. As mentioned earlier, text, language and other culture-bound semiotic resources cannot be efficiently translated and transformed without understanding: 1) how they relate to the overall *structure* and *meaning* of the product and the way the product's features can and will be accessed, used and understood by its potential users through their user agents (browsers and others, including assistive technologies); and 2) how they may *convey*, on their own or coherently combined with other resources, what can be, is being and has been *done* with the product, through diverse user agents.

As a consequence of the above, one of the key discussions when looking at accessibility as "content", from the translation and localization viewpoint, should be whether it is something that can be (easily) transferred. In other words, does accessibility have a concrete form or clearly defining characteristics, and can the forms and characteristics that are culture- and language-bound be "captured" formally, or at least annotated? And if so, how?

This idea is crucial if we want to analyze the way in which web accessibility guidelines can be related to localization and internationalization standards, and, in particular, whether accessibility can actually be transferred throughout the localization process, as well as how it can be achieved via those standards.

If we consider accessibility as *a* quality of the product, we cannot state that accessibility *per se* can be transferred, because qualities are not tangible entities that can be transposed; similarly, "usability" or any other abstract feature of the product cannot be transported. Following this analogy, you cannot transfer "good quality" as a feature, but you can implement a series of best practices that can help to achieve the same (or a better) level of quality in the target product (such as hiring a good translator, understanding what the text will be used for, using correct terminology or a spellchecker). Therefore, what we—as accessibility-aware localizers—can actually transfer or exploit is the information (coding elements, attributes, text, relations) features that can make a product or content accessible: for example, including appropriate alternative texts to images or form labels can make the website content more accessible to non-sighted users, both in the source and the target locale.

3.3 Transfer or Support?

Intuitively, then, we might hypothesize:

1. that technical accessibility features (e.g. the guideline to prevent seizures and physical reactions by avoiding quick flashes) would be "*neutrally*" transferred or re-placed (often outside the localizer's view) from one regional version to another—just like, in computer-aided translation tools, structural elements are excluded from the translator's view; or "placeable" in-line items such as numbers, proper names or formatting or anchor tags are included but protected from editing;
2. that it is mainly accessibility-related *textual* and *cultural* elements that need to be transferred *and* translated or localized in the process.
3. Note, however, that again, just as with structural or "placeable" elements, those seemingly neutral elements sometimes need to be localized to accommodate target conventions. See a few examples in section 3.1, or consider the possibility that certain `<abbr>` elements (WCAG 2.1 guideline 3.1.4), or part or all of the short description in `alt` attributes (guideline 1.1.1.), might need to be omitted or expanded, depending on the target language and culture, and on the co-text, to either provide missing knowledge or avoid audible clutter for visually impaired users.

However, *neutrally transferable* (because structural, outside extractable text to be translated) or *re-placeable* features (like in-line formatting) often *need to be taken into account* as accessibility affordances by the localizer, as they provide contextual and pragmatic information that is necessary for an accessible localization, it could be argued that these accessibility features need to be *supported* in the localization process.

On the other hand, accessibility, like usability "is not simply an absolute property of a product, it is the interaction of a product or service with a particular context of use" (Wilson 2007, 47). Accessible or inaccessible content often depends on *holistic* or on *mediated* (i.e. contextual, non-immediately explicit, depending on interaction) relations between functional and meaning-producing elements, the performance of assistive technology, and the knowledge and capabilities of users, as well as conventional expectations. Understanding how those relations work, and *supporting* them implicitly and interweaving them explicitly, if necessary, is key to preserving or enhancing accessibility in the localization process. Ignoring those connections, on the other hand, may diminish accessibility by unwittingly altering or removing certain interrelated elements in the localization process.

What is more, *new* accessibility needs and opportunities may be identified or created in the process of localization, in which case we would not be talking about transferring accessibility from the source. A proper *production context* for the localizer (accessibility-aware localization tools and environment, structure, contextual information and annotations for the accessible content) can also be said to *support* accessibility by allowing the localizer to enhance it—and possibly transfer it to the other locales, including the source.

3.3.1 *Supporting accessibility features: a few examples*

Let us illustrate the previous ideas with a few examples, which we have come across in the course of our exploratory research. What follows is a list of situations in which the accessible content cannot be transferred and translated directly, or is inaccessible. Accessibility can then be *supported* by knowledge and awareness about accessibility by the localizer, first and foremost, but also by the production context (as described before), allowing the localizer to understand the potential interactions, constraints and relations between content items and users, and to fill in any accessibility gaps, if necessary.

- A certain color is used to denote danger, relevance, a required field in a form, or any other cultural-specific meaning. If the instructions or any textual information on the web page refers, by color alone, to that content, visually impaired users will miss out on that connection. Would it not be possible for the accessibility-aware localizer to enhance the overall communication of the web page for all users by adding, in the instructions or the content mentioning it, a reference to some other specific characteristic that connects both pieces of information (such as shape or relative position in a list), even adding a reference to that specific characteristic (for example a warning icon, including its alternative text), or a link between the two?
- A more or less complex table contains interrelated data on product sales, profits, or safety or tax requirements. For accessibility purposes, a caption has been added, header cells are clearly indicated, and relations between specific data cells and headers are introduced. However, when localizing the website, the table may be redesigned to reflect locale-specific categories, interests and existing data, even producing a more relevant table altogether. Here, the localizer would need to recreate (or create) accessible relations that are not directly transferable from the source content. Locale-specific target images provide another clear example of this.

- Structural and semantic HTML5 elements such as `<nav>`, `<header>`, `<footer>`, `<article>`, headings or `<aside>`, as well as other page landmarks, are very important in helping to navigate a website, particularly for visual disabilities. For two textually identical source segments, different translations may often be needed if the functional and situational context changes—for instance a sentence in the middle of a paragraph, and the same sentence as a section heading. In this case, localizers would need to be informed of the structural or semantic item where the translation is taking place.
- Forms (e.g. for data submission) are useful yet complex interactive elements that need to be accessible by clearly disclosing information and relationships between fields and labels, groups of fields, clues or instructions, to prevent or clearly identify errors, and so on. What is more, error validation and feedback are often produced in the background—on the server side, or in script documents other than those containing the form labels and controls. Accessibility is only achieved if those connections are clearly reconstructed linguistically in the target language, as an output of appropriate or inappropriate interaction.
- Textual items in a horizontal navigation menu may vary in *size* in the different language versions. When checking for accessibility, zooming in a smaller viewport (for instance for a mobile phone) to a larger size (which may be necessary for people with low vision), could result in the menu text items overlapping with a background image, thus making it difficult to read those items for a specific language with larger translated menu items.
- WAI-ARIA attributes (W3C 2017) that define the role, current state or properties of a certain web element, and even provide a textual label to be read out by the assistive technology, may not be extracted by a CAT tool, despite the fact that such attributes would be useful for the translator when translating or to provide context.

3.3.2 Provisional conclusions regarding accessibility transfer or support in localization

The above examples and considerations point to several provisional conclusions:

1. Accessibility needs to be interactively probed and proactively recreated and tested, both in the source and localized versions.
2. The agency of localizers should be assessed, particularly when it comes to how the interactive analysis performed by localizers can help pinpoint problems and improve overall quality aspects of the whole product (including the source).⁶
3. Certain structural or programmatic elements only need to be automatically transferred or re-placed in the target product in order for accessibility to be maintained. However, identifying and understanding the rationale behind many technical aspects related to accessibility are useful and even crucial for localization in at least three ways:
 - a. To prevent those more "technical" elements and attributes from being lost through the more technical processes involved in localization, such as format filtering by CAT tools or automatic CMS authoring.
 - b. As functional context or co-text, to help translate and localize the associated texts or semiotic resources.
 - c. As generic (neither directly related nor contextual) semiotic resources: an accessible product is typically more explicit; therefore, it can give more information to localizers on its meanings, affordances, structure, and so on.

⁶ That is, the ability to act (not just to "simply" replace sentences or strings from one language into another) and to make significant decisions in order to achieve a functionally appropriate translation and the intended effects. See, for instance, O'Hagan and Mangiron (sections 2.4, 4.4 and 4.4.2).

4. By understanding the above aspects, the localizer (see our "broad definition" in note 5) would be in a privileged position to test and annotate content for localization and quality-assurance (including accessibility) purposes. Furthermore, accessibility evaluation could be included as a fundamental component of the regular web localization quality assurance process (Rodríguez Vázquez 2016).⁷

As a result of these reflections, we would be inclined to talk about at least four kinds of localization processes in connection with accessibility, which, in one way or another, can be categorized as the *transfer* or *support* of accessibility features (features that contain or afford accessibility):

1. "Neutral" transfer or re-placeability (potentially non-translatable);
2. Transfer of cultural or semiotic material associated with accessibility mechanisms, which must be translated or localized (e.g. textual alternatives to images; page language);
3. Supporting accessibility features for localization by either
 - a. including those features as context; or
 - b. annotating localization-relevant accessibility aspects that are not obvious from the simple transfer of technical, linguistic, cultural or contextual material (e.g. indicating maximum length of sentences for best readability; purpose of certain images; or where and in which circumstances certain strings will appear for the user);
4. Composite or general qualities that require a broader interpretation and particular pragmatic mechanisms (e.g. a certain reading level), including accessibility enhancement.

3.4 Challenging localization context and agency

We anticipate three main objections to the overall approach that we are taking.

First, some may argue that in the previous examples and suggested solutions, there is a high degree of adaptation in localization that might exceed what the localizer is entitled to do.⁸ However, over the last two decades, definitions of localization (mainly from the industry) have tried to press the point that, contrary to "translation proper" (unfairly characterized as word-for-word or sentence-for-sentence content transposition), localization is about functionally adapting content, culture, conventions and even technical aspects so as to make the product look and feel natural in the target locale, and to cater to end user needs according to the intended functions of the product. Localization can certainly share some of the objectives of transcreation (see Morón and Calvo 2017) as regards adaptation to target cultures and users, combined with other more specific technological and digital interactive aspects that define the discipline. In this sense, the adaptation of a product depends on the reception situation (Pym 2004, 1), and accessibility, as discussed before, does not happen in a vacuum: it depends on all the circumstances that surround a particular reception situation.

Secondly, a few examples certainly suggest that the localizer could make the localized content more accessible than the original, ideally, we might add, after validating this decision with the project manager, client or content owner. In fact, rather than narrowly considering our professional as a linguist only concerned with derivatively transforming

⁷ Further information on accessibility evaluation can be found in WAI (2018a).

⁸ This may be the case if the client's localization maturity model is reactive or immature (Ray and DePalma 2017).

source content into equivalent target content, current and future developments may sensibly force us to go back to targeting more comprehensive language service providers, as advocated early on by Gouadec (2002), or linguistic-cultural mediators. After all, it seems unprofessional for a translator or localizer to simply disregard detected issues that would undoubtedly prevent a percentage of target users from accessing, using and understanding the localized product. Finally, it would be useful to look into the possibility that, given the overall communicative user- and experience-oriented nature of accessibility, trained localizers might be the ideal accessibility-curating managers, coordinating technical or graphic designers, and copywriters or "raw" content producers, particularly in organizations that lack specific accessibility experts.

Thirdly, it could be argued that localizers could use their interpreting skills more efficiently and with less difficulty by looking at the rendered content rather than trying to gather contextual information and intentions from annotations, or from a sequence of segments in a specific localization tool or format. However, due to current practices and standardization, content for localization is often presented in isolated segments and various formats, rather than in the final form (or one that can be previewed).⁹ In addition to this, content is often sent for localization while still in the development stage, which means that it is yet to be pieced together and, therefore, does not provide proper visual and functional context.

What is more, accessibility-related context is as crucial for the localizer as it is for users with disabilities, not to mention visually disabled localizers, who would simply be unable to *look* at the rendered content to visually understand the meaningful relations in it.

4 Methodological Approach

As we have discussed, although "accessibility" might not be a transferable quality per se, our initial hypothesis establishes that the information elements, attributes and relations that can render content accessible might indeed be transferable, or can at least be used to support accessibility throughout the localization process. In order to test this from a pragmatic point of view, we decided to analyze the main current accessibility guidelines (WCAG 2.1), in search of specific levels and types of information that would fit with the purposes and structure of the ITS and XLIFF multilingual data exchange standards.

But while the first three top levels of WCAG 2.1 (principles, guidelines and success criteria) are an excellent instrument to guide us through accessibility best practices, they are all intangible and do not have a specific (human or computer) linguistic form that we can use. We needed to base our analysis of how internationalization and localization standards can exchange accessible content on a more tangible, formal or *capturable* level in order to test it. We therefore decided to focus on the suggested *techniques* that describe how to make content accessible in a practical way, with real-case examples and predefined testing mechanisms.

⁹ Despite the fact that, for instance, XLIFF 2.1 does have mechanisms that would allow the content to be previewed, it is, alas, to the best of our knowledge, still unsupported by CAT tools. The birth of initiatives like the Translation API Cases and Classes Initiative (TAPICC) and the publication of its latest derivable: "XLIFF 2 Extraction and Merging Best Practice, Version 1.0" (Filip and Husarčík 2018) are good proof that there is a common interest to improve these aspects.

In this fourth level of detail in WCAG 2.1, we encounter specific (sufficient and advisory) techniques that essentially translate into the use of specific code, text or combination of modes, which can potentially be captured in data exchange formats.

As previously discussed in section 2.2, techniques relate to one or more success criteria, which are "testable" statements that can help us understand whether the localized technique can produce accessible content in the target web page. The wording of the success criteria and the supporting material (explanations, examples, etc.) that accompany success criteria and techniques can also be useful to understanding what can be achieved, and how different users in different cultures who use different languages may need the techniques to be adapted in specific ways.

Let's take technique *H37 -Using alt attributes on img elements* (W3C 2016b) as an example to illustrate what a technique entails. This HTML technique can help accomplish success criterion 1.1.1 and recommends using the value of the attribute `alt` to include a short text alternative to the `` element. Examples and tests to verify whether or not this technique is used are also provided (ibid.).¹⁰ This technique is technology-dependent (it refers to the specific use of an HTML attribute to define the purpose of a web object) and can be directly related to a localization process scenario in which all the values of the `alt` attributes of the source content are extracted to be appropriately translated according to their function.

In our preliminary study, we selected a limited (small, yet representative) number of success criteria (from three out of the four WCAG 2.1 principles) that, in one way or another, can be directly related to localization for one or several of the following reasons:

1. They specifically relate to language, text or other localizable modes or cultural assets;
2. They refer to essential functionality (e.g. navigation) and affordances in a website;
3. They are important or key to understanding and generally making sense of the content.

The complete list of the 11 success criteria studied is shown in Table 1.

Table 1. Success criteria covered in our preliminary study

Success criterion	Guideline	Principle
1.1.1 Non-text content	1.1 Text Alternative	1. Perceivable
2.4.1 Bypass blocks	2.4 Navigable	2. Operable
2.4.2 Page Titled	2.4 Navigable	2. Operable
2.4.3 Focus order	2.4 Navigable	2. Operable
2.4.4 Link Purpose (In Context)	2.4 Navigable	2. Operable
3.1.1 Language of Page	3.1 Readable	3 Understandable
3.1.2 Language of Parts	3.1 Readable	3 Understandable
3.2.1 On Focus	3.2 Predictable	3 Understandable
3.2.2 On Input	3.2 Predictable	3 Understandable
3.3.1 Error Identification	3.3 Input Assistance	3 Understandable
3.3.2. Labels or Instructions	3.3 Input Assistance	3 Understandable

A protocol and an analytical card were designed to allow us to analyze each technique against possible features and techniques in XLIFF and ITS in a systematic and comparable way. Let us remember that the main objective is to find out how XLIFF and

¹⁰ This success criterion establishes that apart from some exceptions, all the non-text content must have an alternative text (W3C 2018). It is the only success criteria included in guideline 1.1 (Text Alternatives) of the first principle (Perceivable).

ITS are able to transfer or support accessibility through their mechanisms, in a way that might be useful for a localizer, essentially by using the appropriate CAT or localization tools for their task.

The card included the following categories: name of the technique (used as a unique identifier in our analysis), situation (context of use), discussion (a brief description of the technique), possible relation to ITS (how this technique might be related, integrated or mapped to ITS annotation mechanisms), possible relation to XLIFF (how this technique might be related, integrated or mapped to existing XLIFF transfer mechanisms); and, finally, a real example (or its location) is provided. Table 2 provides an example of one of these analytical cards.

Table 2. Example of the analytical card used to compare accessibility techniques and localization standards

Technique (Sufficient)	Situation	Discussion	ITS	XLIFF	Example
G94: Providing short text alternative for non-text content that serves the same purpose and presents the same information as the non-text content.	There are 5 possible situations for the success criteria 1.1.1. This technique applies to “situation A”: “If a short description can serve the same purpose and present the same information as the non-text content”.	This technique is similar to G92, the only difference here is that here it is specified that the text alternative should be short.	If a limitation of characters is associated with a short alt text, this could be annotated using the ITS data category “Allowed Characters”.	Text alternatives should be extracted from the original content and included in XLIFF units. We could foresee the creation of a metadata attribute that includes information on the meaning, function and importance of the alternative text. Besides, a linking mechanism that maps the alternative text to the non-text content might be useful for the translator to have more contextual information of the element that they are translating.	Examples can be found in http://www.tomjewett.com/accessibility/alt-text.html

This collection of cards describing and analyzing each technique of the selected success criteria in comparison with ITS and XLIFF allowed us to have an overall picture of their possible relations and to extract our first preliminary results. An initial overview of these results is provided in the following section.

5 Preliminary Insights: XLIFF and ITS Transfer and Support for Accessibility Content

The following insights are related to the examples and provisional conclusions discussed in sections 3.3.1 and 3.3.2 of this paper, which are also a consequence of the analysis that we have just presented in our methodological approach (section 4). We deemed it appropriate to use those examples and provisional conclusions before the methodological section to advance our arguments. Nonetheless, the above conclusions and examples and the following preliminary insights feed off of each other. What we will be mostly concerned with here is the perceived capabilities and potentials of XLIFF 2.1 and ITS 2.0 to carry out the transfer and support needs identified above. To begin with, we present the positive connection points that we have revealed between accessibility transfer and support needs, and localization standards. Secondly, we discuss the limitations that we have encountered.

5.1 Accessibility transfer and support capabilities

In terms of actual localizable content, XLIFF can contain any extracted content that affords accessibility, whether it is text inside HTML elements or attribute values, as well as the actual structural code (either as a separate skeleton file or code, or as accompanying <group> elements). Therefore, we can state that any information included in the source code can actually be marked and transferred through XLIFF. The tool used to extract this

information should recognize the localizable information (the value of the `alt` attributes, for example) and include it in the XLIFF document that will later be manipulated in a CAT tool by a localizer, although the method for presenting and grouping the structural and textual information in XLIFF is also very important for the end result. For instance, the use of both the `<group>` and the `<unit>` (which can include several related segments) elements needs to be explored. Contextual and structural information included in the XLIFF file also needs to be presented to the localizer through the translation or localization tool.

Sometimes the text itself needs additional contextual information in order for the localizer to understand the accessibilizing function it is meant to serve. Here is where additional mechanisms, categories or modules are needed to help in ITS and XLIFF.

Both ITS and XLIFF have mechanisms to annotate content that is marked for internationalization or transported for localization. For example, the attribute `translate` which is available in both languages, allows the developer to mark the content that should be translated (with the value `yes`) or, on the other hand, protected (`no`). Furthermore, other mechanisms, like the element `<note>` in XLIFF or the Localization Note data category in ITS could include additional information that would then be transferred to the translator or localizer. Of course, once again, this mechanism would need specific Enriching tools (or manual input) in order to be useful, as well as CAT tools to display those notes. Additional annotations that might help make the translator aware of the function and meaning of the element to be localized are therefore crucial to ensure that the target content is as accessible as the original. A new XLIFF accessibility module can also be envisioned to enrich extracted localizable data with specific information on its accessibility function.

ITS is able to provide specific internationalization information, which is also very useful for usability and accessibility, particularly when it can be automatically controlled by a CAT tool: allowed characters, directionality, language information, and, of course, the "translate" category. For example, if we want to introduce a limitation of characters for the value of the `alt` attribute to include short descriptions of images, we could do so through the storage size data category in ITS. Similarly, XLIFF has a "Validation" module that might help verify the omission or the format of certain accessibility features like text length, bad practices for alternative texts (e.g. including "image of"), and so on. Again, for these mechanisms to be successful, the tools used in the internationalization and localization phases need to fully support them.

5.2 Accessibility transfer and support limitations and opportunities

We have also encountered potential limitations in terms of accessibility transfer and support. First of all, we realized that XLIFF and ITS do not have clear mechanisms to show certain relationships that may be essential for accessibility (and, therefore, for proper localization):

- Cross-references between related elements within the same file that are not contiguous (for instance, between labels and form fields, through the `for` attribute in the `<label>` elements; by means of the `aria-describedby` or `aria-labelledby` attributes; or among links, including to skip navigation). The "Resource Data" module in XLIFF and the "External Resource" category in ITS only seem to connect elements in different files.
- Similarly, semantically-, textually- or structurally-related elements should be linked in one way or another, particularly if not contiguous. Take, for instance, dynamically-

shown error messages for specific form interactions or fields; or the ability to group the sequence of focusable elements and landmarks on a webpage, which are the main access mechanisms for visually impaired users. The `<group>` element could help, although it is also important not to alter other logical groupings. Sub-flows and the "Metadata" module could also offer useful mechanisms to be explored.

- The "Format Style" module also needs to be fully explored to visually and logically display relations, but seems to lack the possibility of processing CSS—including aural—style sheets.

Secondly, during our comparative study, we also discovered limitations in terms of multilingual support. In particular, technique H58 for the success criterion 3.1.2 (Language of Parts) recommends the use of the `lang` attribute to indicate a section of content that is different from the main content of a page. As a matter of principle, when working with XLIFF, only one source language is allowed. However, it is possible to include sections that are in a different source language by making use of the `itsm:lang` attribute;¹¹ therefore, we can confidently state that it is indeed possible to transfer this feature using a combination of XLIFF and ITS. Nevertheless, a quick test run on XLIFF extractors¹² have shown us that this mechanism is not yet implemented and that all the content (main content and content tagged in another language) is just extracted and marked as the general source language¹³. More tests, including a larger and more representative set of XLIFF enrichers, will be conducted to confirm that this mechanism is indeed not being extensively implemented, and if this is the case, we will inform the XLIFF TC and conduct awareness actions to promote its importance and call for support.

6 Conclusions and Future Work

This initial study on the possibility of transferring or supporting accessibility throughout the localization process, and particularly in multilingual data exchange standards, has raised many questions and unveiled many positive interconnections as well as limitations. We soon realized that accessibility is an abstract and intangible notion that depends on each reception situation and cannot be transferred per se. However, by taking a closer look at the guidelines that define how to make content accessible, we found that a certain level of analysis in the Web Content Accessibility Guidelines, namely the documented techniques that can be used to meet the different testable success criteria of the accessibility guidelines, can be very useful, as they provide tangible information and knowledge that can be presented and formally represented to the localizer. We then selected a number of WCAG 2.1 success criteria in an attempt to demonstrate the

¹¹ Please refer to <http://docs.oasis-open.org/xliff/xliff-core/v2.1/os/xliff-core-v2.1-os.html#langAnnotation> in the XLIFF 2.1 specification for more information on this point.

¹² Following the official XLIFF terminology, an "Extractor" is a tool that can carry out "the process of encoding localizable content from a native content or User Interface format as XLIFF payload, so that localizable parts of the content in the source language are available for Translation into the target language along with the necessary context information". Examples of other XLIFF agents (applications or tools) are: "Enricher", "Merger", "Modifier", "Writer".

¹³ This statement is corroborated by the results of the "XLIFF Version 2.1 Support in CAT Tools" report (Morado Vázquez and Filip 2018), which gathered statements of use of the new version 2.1. In this report, none of the tools categorized as extractors claimed to support the `itsm:lang` attribute. However, it should be noted that those results are based on early implementations of the new version, and that the current support situation might have positively evolved since then.

capabilities of two internationalization and localization standards (ITS 2.0 and XLIFF 2.1) to capture, transfer and support accessibility qualities (and potentials).

Our preliminary findings have shown that certain accessibility features, particularly those related to textual content and inter-semiotic purposes, can be captured, marked, transferred and annotated through existing mechanisms in ITS and XLIFF, whereas other qualities related to semantic relations, groupings, and access pathways and navigation probably need to be further explored and developed in order to support the task of localizers. For both the more and the less positive conclusions, however, it is important to understand that accessibility qualities need to be properly identified through ITS annotation and/or the extraction and enriching of the XLIFF mechanism, to be adequately processed by translation and localization tools, and, above all, interpreted and adapted by the localizer. Previous research by Rodríguez Vázquez (2016) on alternative texts points in that direction, but we also need to carry out empirical research to identify what kind of accessibility content transfer or support can lead to the same (or even better) level of accessibility conformance in the target content, and improve the overall quality of source and target products.

Further research in these particular areas will be carried out, and, if judged necessary, the respective technical committees will be informed, so that they can adopt the necessary changes to accommodate these mechanisms in future versions of the standards. Finally, we identified some accessibility features that are yet to be implemented by existing XLIFF extractors, although they can be transferred through existing mechanisms of the exchange data standards analyzed, which is the case when transferring multilingual texts into an XLIFF file.

This initial study that seeks to establish the basis of the relationship between these two areas will serve as the foundation for our future research efforts. In particular, we will work sequentially in the following three phases: first, we will produce a comprehensive analysis of the results of mapping WCAG techniques to ITS and XLIFF. This will allow us to provide a detailed map of the connections between the two areas. Secondly, we will conduct extensive tests with different existing CAT tools to verify whether or not the mechanisms identified in the previous phase are also supported. At this stage, we will further investigate specific areas of connection or ones that have proved challenging (like the aforementioned mechanisms to interrelate content). Finally, we are planning to carry out empirical studies to validate the usefulness of the transferred accessibility features to obtain accessible localized content.

7 Acknowledgements

We would like to express our deepest and sincerest gratitude to Silvia Rodríguez Vázquez and Yves Savourel for reading this article and offering their feedback, advice and vast knowledge.

8 References

- Fundación SIDAR. 2009. *Pautas de Accesibilidad para el Contenido Web (WCAG) 2.0*. Spanish Candidate Translation for the Web Content Accessibility Guidelines (WCAG) 2.0 of the W3C. Accessed August 28, 2018. <http://www.sidar.org/traduccion/wcag20/es/>.
- Filip, David, and Ján Husarčík. 2018. *TAPICC. XLIFF 2 Extraction and Merging Best Practice, Version 1.0*. Accessed August 28, 2018. <https://galaglobal.github.io/TAPICC/T1/WG3/wd01/XLIFF-EM-BP-V1.0-wd01.html>.
- Gouadec, Daniel. 2002. "Training Translators: Certainties, Uncertainties, Dilemmas." In *Training the Language Service Providers for the New Millennium*, by Belinda Maia, Johann Haller and Margherita Ullrych, 31-41. Porto: Faculdade de Letras da Universidade do Porto.
- Gutiérrez y Restrepo, Emmanuelle, and Loïc Martínez Normand. 2010. "Localization and Web Accessibility." *Tradumàtica: traducció i tecnologies de la informació i la comunicació* 8: 1-6.
- Jiménez-Crespo, Miguel Ángel. 2013. *Translation and Web Localization*. London/New York: Routledge.
- Kaptelinin, Victor. 2014. *Affordances and Design*. Georgia: Interaction Design.
- Koby, Geoffrey S, Paul Fields, Daryl Hague, Arle Lommel, and Alan Melby. 2014. "Defining Translation Quality." *Tradumàtica: traducció i tecnologies de la informació i la comunicació* 12: 413-420.
- Morado Vázquez, Lucía, and David Filip. 2018. *XLIFF Version 2.1 Support in CAT Tools*. Subcommittee Report, OASIS XLIFF Promotion and Liaison Subcommittee.
- Morón, Marián, and Elisa Calvo. 2017. "Introducing Transcreation Skills in Translator Training Contexts: A Situated Project-based Approach." *The Journal of Specialised Translation* 29: 126-148.
- Munday, Jeremy. 2009. *Issues in Translation Studies*. Oxon: Routledge.
- Nielsen, Jakob. 2012. *Usability 101: Introduction to Usability*. Accessed September 14, 2018. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>.
- Norman, Don. 2013. *The Design of Everyday Things. Revised and Expanded Edition*. New York: Basic Books.
- Ó Broin, Ultan. 2004. "Accessibility Is Just Another Language." *Multilingual Computing and Technology* 15 (3): 17-20.
- OASIS XLIFF TC. 2018. *XLIFF 2.1*. Accessed August 28, 2018. <http://docs.oasisopen.org/xliff/xliff-core/v2.1/os/xliff-core-v2.1-os.html>.
- Pym, Anthony. 2004. *The Moving Text: Localization, Translation, and Distribution*. New York/Amsterdam: John Benjamins.

- Ray, Rebecca, and Donald A DePalma. 2017. *Localization Maturity Model. Release 3.0*. Common Sense Advisory.
- Rodríguez Vázquez, Silvia. 2016. "Assuring Accessibility during Web Localisation: an Empirical Investigation on the Achievement of Appropriate Text Alternatives for Images". PhD dissertation. Universidad de Salamanca/Université de Genève.
- Rodríguez Vázquez, Silvia, and Jesús Torres-del-Rey. 2014. "Fostering Accessibility through Web Localization." *Multilingual Computing and Technology* 25 (7): 32–37.
- Tercedor Sánchez, Maribel. 2010. "Translating Web Multimodalities: Towards Inclusive Web Localization." *Tradumàtica: traducció i tecnologies de la informació i la comunicació* 8: 1-8.
- Torres-del-Rey, Jesús, and Lucía Morado Vázquez. 2015. "XLIFF and the Translator: Why Does it Matter?" *Tradumàtica: traducció i tecnologies de la informació i la comunicació* 13. 584-607.
- Torres-del-Rey, Jesús, and Silvia Rodríguez Vázquez. 2013. "Traducción y accesibilización de discurso público en formato web." In *Traducción, política(s), conflictos: legados y retos para la era del multiculturalismo*, by María Carmen Claramonte, África Vidal, María Rosario and Martín Ruano, 121-14. Granada : Comares.
- Torres-del-Rey, Jesús, and Silvia Rodríguez Vázquez. 2016. "New insights into translation-oriented, technology-intensive localiser education: accessibility as an opportunity." *Proceedings of the 4th International Conference Technological Ecosystems for Enhancing Multiculturality (TEEM '16)*. 971-978. New York: ACM Press.
- Torres-del-Rey, Jesús, Silvia Rodríguez Vázquez, María del Mar Sánchez Ramos, Lucía Morado Vázquez, Manuel Mata Pastor, Luis García Nevares, and Emilio Rodríguez V. de Aldana. 2018. "Educación para la accesibilidad en localización web: el proyecto ALMA." Paper presented at the Seminario Internacional sobre Traducción y Accesibilidad al Patrimonio (SITAP 2018), Granada, Spain, March 2018.
- W3C. 2013. *Internationalization Tag Set (ITS) Version 2.0*. Accessed August 28, 2018. <https://www.w3.org/TR/its20/>.
- W3C. 2016a. *Techniques for WCAG 2.0*. Accessed August 28, 2018. <https://www.w3.org/TR/WCAG20-TECHS/>.
- W3C. 2016b. *H37: Using alt attributes on img elements*. Accessed August 28, 2018. <https://www.w3.org/TR/WCAG-TECHS/H37.html>.
- W3C. 2016c. *Understanding Techniques for WCAG Success Criteria*. Accessed August 28, 2018. <https://www.w3.org/TR/UNDERSTANDINGWCAG20/>.
- W3C. 2017. *Accessible Rich Internet Applications (WAI-ARIA) 1.1*. Accessed August 28, 2018. <https://www.w3.org/TR/wai-aria>.

- W3C. 2018. *Web Content Accessibility Guidelines (WCAG) 2.1*. Accessed August 28, 2018. <https://www.w3.org/TR/WCAG21/>.
- WAI. 2018a. *Evaluating Web Accessibility Overview*. Accessed August 28, 2018. <https://www.w3.org/WAI/test-evaluate/>.
- WAI. 2018b. *WCAG 2 Documents*. Edited by Shawn Lawton Henry. Accessed August 28, 2018. <https://www.w3.org/WAI/standards-guidelines/wcag/docs/>.
- WAI. 2018c. *Web Content Accessibility Guidelines (WCAG) Overview*. Accessed June 28, 2018. <https://www.w3.org/WAI/intro/wcag>.
- Wilson, Chauncey. 2007. "The Problem with Usability Problems: Context is Critical." *Interactions - Societal Interfaces: Solving Problems, Affecting Change* 14 (5): 46-ff.