Automatic Text Simplification Tools for Deaf and Hard of Hearing Adults: Benefits of Lexical Simplification and Providing Users with Autonomy

Oliver Alonzo, Matthew Seita, Abraham Glasser, Matt Huenerfauth
Rochester Institute of Technology (RIT) Rochester, NY
oa7652@rit.edu, mss4296@rit.edu, atg2036@rit.edu, matt.huenerfauth@rit.edu

ABSTRACT
Automatic Text Simplification (ATS), which replaces text with simpler equivalents, is rapidly improving. While some research has examined ATS reading-assistance tools, little has examined preferences of adults who are deaf or hard-of-hearing (DHH), and none empirically evaluated lexical simplification technology (replacement of individual words) with these users. Prior research has revealed that U.S. DHH adults have lower reading literacy on average than their hearing peers, with unique characteristics to their literacy profile. We investigate whether DHH adults perceive a benefit from lexical simplification applied automatically or when users are provided with greater autonomy, with on-demand control and visibility as to which words are replaced. Formative interviews guided the design of an experimental study, in which DHH participants read English texts in their original form and with lexical simplification applied automatically or on-demand. Participants indicated that they perceived a benefit form lexical simplification, and they preferred a system with on-demand simplification.

Author Keywords
Lexical Simplification, Reading Assistance, People who are Deaf or Hard of Hearing, Autonomy.

CSS Concepts
• **Human-centered computing~Empirical studies in accessibility.**

INTRODUCTION
We investigate Automatic Text Simplification (ATS) technologies to benefit adults who are Deaf or Hard of Hearing (DHH). Nearly 1 in 6 adults in the U.S. is DHH [7], and prior literacy research has established that there is a wide range of reading skills among DHH adults: Some studies found over 17% of deaf adults have "low literacy" [18], and standardized educational testing has revealed that the average reading level among U.S. deaf adults finishing secondary school is significantly below the average of their hearing peers [45]. Research has found that there are lower educational outcomes among DHH individuals, e.g. in 2017, 34% of hearing individuals in the U.S. had completed a bachelor's degree, compared to only 18.8% of deaf individuals [22]. Reading literacy is also important in an employment context, e.g. for learning new skills mid-career [41]. Research has revealed a 22% lower rate of employment among DHH adults, and DHH adults' salaries are only 64%, in comparison to their hearing peers [47].

With this context, we investigate applications that provide ATS for complex texts, in particular to support reading texts online. ATS can scan a webpage, find complex words or phrases, and replace them to make the text easier to understand. Prior work had evaluated syntactic simplification with DHH users (simplifying the structure of sentences) [27]. However, prior work has not evaluated DHH users' perception of automatic reading-assistance tools based on *lexical* simplification (replacing a difficult word with a simpler synonym), nor of *hybrid* systems that combine both syntactic and lexical approaches. As a first step, this study investigates whether DHH users perceive a benefit from lexical simplification alone.

Prior reading-assistance work has investigated using ATS with various groups, e.g. people with aphasia [15], people with dyslexia [36, 37], and second-language learners [3]. However, most of these studies have focused on improving the quality of the underlying ATS technology, with less focus on the human-computer interaction (HCI) aspects, e.g. how users actually interact with the software itself, what modifications they can make to the software, how much decision-making they have to do, and how much control they have over what the software automatically does for them.

Having more ATS reading assistance may not necessarily be better: as with any assistive technology that uses artificial intelligence, there is risk that it may make errors or that users may experience a loss of agency, if they only see a filtered or modified view of the original text. Various prior work in HCI has investigated the issue of autonomy or related concepts [10, 21]. In the context of ATS, we define **autonomy** as how much control a user has over what words are simplified or visibility of what words have already been simplified. We speculate that providing DHH users with a desired level of autonomy may result in an improved user experience and motivation to use the software again.

CHI 2020, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04...$15.00
https://doi.org/10.1145/3313831.3376563

After conducting a preliminary study to inform our design of a set of prototypes, we conducted an experimental study with DHH participants who read four articles using various prototypes, responded to subjective questions about each prototype, and shared feedback. Participants reported that they perceived a benefit from lexical simplification tools, and they preferred prototype designs that provided them with greater autonomy. Participants also discussed tradeoffs between their highest-rated prototypes in this study.

The main contributions of this paper are empirical:

• Our experimental study with DHH participants yielded quantitative data (e.g. Likert-scale questions regarding their experience using the text simplification software) and qualitative data (open-ended feedback questions). Our analysis revealed that DHH users perceived a benefit from automatic reading-assistance tools that provide lexical simplification; prior research had not investigated lexical simplification tools among DHH users.

• We found that DHH users preferred prototypes that provided greater levels of autonomy for ATS-assisted reading. We also found that DHH users could articulate the benefits of autonomy in the interface and comment on the trade-offs among designs in this space. These findings provide a foundation for future HCI research and design of reading-assistance technologies for DHH users.

RELATED WORK

As context for our work, this section describes prior work on reading assistance technologies, including systems that simplify text or modify its appearance for readability. Prior HCI research on the role of autonomy is also discussed.

Reading Assistance and Text Simplification

Prior research has found that over 17% of deaf adults have "low literacy" [18]. Other work has measured fourth-grade reading levels among DHH high school graduates in the United States [45] and sixth-grade reading levels among DHH university students [1, 34]. One study found that over 30% of deaf high school graduates in the United States were "functionally illiterate" [30]. This is certainly not to say that all DHH adults have weak reading skills, as *many DHH people are excellent readers*. Rather, there is a great reading-skill diversity among DHH adults.

Researchers have investigated using text simplification technology to support reading online text for people with aphasia [15], people with dyslexia [36, 37], low-literacy readers [48], second language learners [3], and people who are DHH [24, 27]. Few studies measured any difference in objective measures (e.g. comprehension questions) [27], but some measured a difference in users' subjective perception of whether they benefitted from the technology [36, 40].

While for second language learners, translation of text into another language might be beneficial, translating English text on websites into American Sign Language (ASL) would not benefit all DHH individuals in the U.S., since there is great diversity in the levels of ASL proficiency [20]. In addition, even for DHH users who are ASL signers, providing ASL translations of English text on-demand is currently infeasible: Requesting human sign-language interpreters to translate texts online into ASL would be very resource-intensive, and the state of the art for automatic ASL machine translation is not yet sufficient for this task, as discussed in [9]. Furthermore, research has found that providing text-simplification not only has language-learning benefits, but also when non-native readers were provided with both translation and simplification technology, they preferred simplification [17].

There are three main approaches to modifying text for readability or understandability: *syntactic* simplification, in which the structure of sentences or phrases are modified to reduce grammar complexity; *lexical* simplification, which involves identifying a complex word and replacing it with a simpler synonym [43]; and hybrid approached that combine both technologies. Table 1 shows an example of a sentence simplified using both approaches, from [48].

Original	The senators rejected the proposal from the Information Technology Division of the Senate to change all computers in the plenary room, alleging that public opinion would not receive the disbursement willingly.
Syntactic	The senators alleged that public opinion would not receive the disbursement willingly. Then, the senators rejected the proposal from the Information Technology Division of the Senate to change all computers in the plenary room.
Lexical	The senators rejected the proposal from the Information Technology Division of the Senate to change all computers in the plenary room, alleging that public opinion would not receive the **expense** willingly.

Table 1. Example from [48] of syntactic and lexical simplification (for lexical, the replaced word is shown in bold).

The three different approaches identified above may benefit user-groups to a different degree, depending on their literacy profile [40]. Prior work has revealed benefits from providing syntactic simplification alone for DHH adults [27], which aligns with research on word acquisition suggesting that syntactic structure may be a key source of reading difficulty for DHH readers [13]. However, more recent education research has suggested that, in addition to syntax, vocabulary knowledge also factors into the diversity in literacy levels among people who are DHH [12, 28]. Furthermore, research into the reading strategies used by DHH readers has identified unfamiliar vocabulary as a key source of difficulty in expository text [4]. This suggests that these readers might also benefit from lexical simplification, especially for expository text. Notably, prior studies with lexical simplification among users with dyslexia [36] did not reveal significant effects on comprehension scores, but participants reported perceiving the text significantly easier to read and understand under certain interface treatments. However, it remains unknown whether automatic lexical simplification would be beneficial for DHH readers or whether they would be interested in – and perceive a benefit from – a system that provides this form of simplification *alone* (this second question is investigated specifically in our study).

User Interface for Text Simplification

Aside from transforming the text itself, there has also been research on reading-assistance tools that modify the layout or visual appearance of a page of text, including research on the user-interface design of such systems, e.g. inserting whitespace between sentences in a paragraph [52], highlighting portions of paragraphs [26], annotating text [11], or changing font size and line spacing [38]. This research indicates how visual design affects text readability.

Prior reading-assistance research with lexical simplification has used various user-interface designs, varying in the level of control users are provided and visibility as to what words have been replaced. However, most prior user studies on lexical simplification for reading assistance have focused their investigation on the quality and effect of different simplification methods, e.g. [33, 36, 40], rather than on the user-interface design. In one such study, Rello *et. al* investigated the use of different simplification systems to assist people with dyslexia, and one of their tools provided simplifications on-demand. Participants felt that texts were easier to understand when using the version that provided simplifications on-demand, as compared to the original text and a version that automatically simplified text before the user saw it [37]. Thus, the authors indicated a need for further research to investigate the **user interface** of such systems. However, to the best of our knowledge, the effect of providing on-demand choice in the user-interface design of lexical simplification has not been further investigated.

User Autonomy

Prior HCI research has investigated the effect of providing users with autonomy, with the concept sometimes referred to by different terminology, e.g. agency, independence, or self-determination [10, 19, 21]. This research has found that fine-tuning the levels of users' autonomy can increase motivation and their desire to use software [19]. However, there are tradeoffs: For instance, increased autonomy in an interactive system may also increase the complexity or cognitive load, and user autonomy may need to be limited in order to provide greater security or privacy [21]. A workshop at *CHI 2014* explored the topic of autonomy, which the organizers defined as relating to a sense of agency or self-regulation [10]. The workshop identified four spheres of autonomy in HCI: (1) within a software design, which refers to the sense of control the user has when using software, and has been shown to impact motivation and engagement [19]; (2) the creation of assistive technologies to increase autonomy in daily life; (3) enabling users to design their own technologies; and (4) helping users develop a sense of autonomy through technology. While simplification as reading assistance may fall under *assistive technologies to assist in daily life*, our work focuses on investigating the effect of user autonomy when using text simplification *within the software design*.

In the accessibility field, Rello *et al.* [37] had speculated that providing greater control with *on-demand* assistance explained some of the benefit of simplification for users with dyslexia in their study. ATS is an artificial-intelligence (AI) technology, and in a recent paper on the ethics of AI-based assistive technologies for DHH users, authors have advocated for enabling DHH users to make their own *choices* about when to deploy AI-based systems [25]. This sentiment intersects with recently published guidelines for human-AI interaction [2], which recommend that systems enable efficient invocation of AI services *when needed*. Other accessibility researchers have advocated moving away from paternalistic approaches in accessibility technology design, and instead towards approaches that provide users with more choice or that support "*inter*-dependence" [23]. Given prior work on the benefits of providing users with on-demand AI-powered assistive tools, we investigate whether DHH users of ATS reading-assistance tools prefer designs that provide them with greater control over which words are simplified or greater visibility of what words have already been modified.

RESEARCH QUESTIONS AND METHODS

We investigate two research questions, in the context of a reading-assistance software system for DHH users, capable of providing lexical simplification of words on webpages:

RQ1: Does providing a lexical-simplification reading-assistance tool affect DHH users' perception of their ease in reading a text or perception of how well they understood it?

RQ2: Are DHH users' subjective preferences about lexical-simplification tools affected by whether they are provided with greater autonomy (i.e. in requesting simplifications on-demand or seeing what words have been replaced)?

We first conducted a preliminary study with 12 DHH participants to evaluate a wide variety of prototype designs for the user interface of a reading-assistance system with lexical simplification. Our goal was to select a manageable subset of conditions for our second study. Next, we conducted an experimental study in which 25 DHH participants compared working prototype systems. That final experiment included a baseline system that did not provide any lexical simplification (in support of RQ1 above) and a system that automatically applied lexical simplifications before the users saw the text (for RQ2).

PROTOTYPE

To investigate RQ2, we needed a set of prototype variations for reading-assistance software that would enable us to explore the concept of user autonomy. Based on prior work on lexical simplification, we identified two principal design dimensions that appeared relevant to this issue: how the user initiates a request for a simplification (henceforth referred to as *user initiative*) and whether the system provides some visible indication that a simplification has occurred (henceforth referred to as *change visibility*). The context of our prototype is that it would provide reading assistance while the user is visiting web pages, and it would therefore

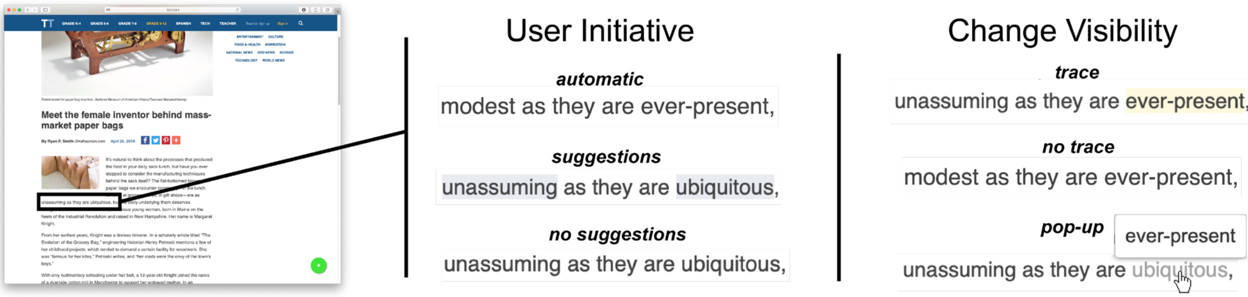

Figure 1: An article used for the video demonstrations in our preliminary study, with a zoomed-in view of what some text looked like under different design variations for *user initiative* and *change visibility*, which were displayed to users in this study.

consist of some type of web browser enhancement or plug-in, an approach recommended in [48].

Within the *user initiative* dimension, we identified five levels based on an analysis of prototypes in previous research on text simplification, with each level varying as to the degree of autonomy it provided to the user:

1. As a baseline, we provided users with a non-interactive prototype, in which the system automatically replaced all the words with simpler versions available before the user even saw the text [27, 37]. As illustrated in Figure 1, we refer to this level as ***automatic***.
2. Providing some additional *user initiative*, the ***replace all*** version allowed the user to click on a button to simultaneously replace all words on the page for which simpler versions were available.
3. ATS software may not be able to provide a simpler replacement for every word in a text, or it might determine that the existing text is already simple enough for the reader. Researchers have therefore sometimes used text decoration, such as highlighting or text color differences, to suggest to the reader which words in the text the user may wish to simplify [6, 31, 36, 37]. We refer to this level as ***suggestions***.
4. Given that this text decoration would be added by the system, we developed a variation of the *suggestions* level that allowed the user more autonomy by providing a button to activate or deactivate this decoration. In this way, the user would be able to experience the web page in its original intended appearance without any visual decoration, and the system would not be suggesting that they might want to simplify particular words. We refer to this level as ***toggle suggestions.***
5. Researchers in prior work have also created reading-assistance tools in which the user was allowed to request a simplification on any word they felt was complex, without providing the user with any text decorations [37] to suggest that they may want to simplify particular words. We refer to this level as ***no suggestions.***

In addition to design differences that influence user initiative over the simplification process (listed above), we identified a relatively orthogonal design dimension in which prototypes vary in the degree of *change visibility*, that is, whether the system indicates to users retrospectively which words have been modified. Along this dimension, we identified four levels of increasing autonomy, with the perspective that designs likely provide more autonomy if they clearly indicate to the user which words have been changed, making it easy to see what the original word had been, and minimize blocking the onscreen text:

1. Researchers in prior text simplification work have investigated providing simplified versions of text in a separate region to the side of the window displaying the text [40, 44], thereby leaving the original word in the text visibly unchanged. We refer to this level as ***sidebar***.
2. Researchers have also used pop-ups (small rectangles that appear floating above the text near the word that has been hovered-over) to provide the reader with a simpler word while still allowing the user to see the original text [31, 36, 37]. We refer to this level as ***pop-up***.
3. While the two options above would actually leave the original word in place in the text, in this next design variation, the word is replaced with a simpler synonym, but some text decoration, e.g. a background highlight color, is used to convey visually that a change has happened [6]. We refer to this level as ***trace.***
4. Lastly, we include a version, as a baseline, in which the word is replaced without leaving behind any visible indication that it had been replaced [8, 27, 37]. We refer to this level as ***no trace.***

Simplifications

While the state of the art in ATS has been increasing in accuracy in recent years, for both the identification of complex words and the selection of appropriate synonyms (given the specific meaning of the word in context) [29], ATS is still not perfect, with many potential places for errors, which could be detrimental to reading rather than helpful [42]. Thus, to ensure that the quality was kept constant across texts, we used a **Wizard-of-Oz** approach to create the simplifications for our prototype by using human-produced simplifications aided by Par4Sim [51] (a semantic writing tool that automatically identifies complex words, highlights them and suggests potential simplifications) and SimplePPDB++ [29] (a resource containing lexical paraphrase rules with readability scores). A team of three (2 Deaf researchers and 1 hearing researcher whose native language is not English) used the following protocol to obtain lexical simplifications for stimuli:

1. The researchers input the text into Par4Sim to begin identifying complex words. While they used the complex words suggestions from the system, they also read through the text in case there were any complex words the system had missed. The first-hand perspective of the researchers as members of the Deaf community or as someone who had learned English as a second language was valuable in their identification of potentially difficult-to-read words in the text.
2. After identifying all complex words, they would look at the suggestions from Par4Sim.
3. If no Par4Sim suggestions were considered simpler or fitting to the context, then they would query the SimplePPDB++ to consider other suggestions.
4. If nothing above provided satisfactory simplifications, then they would perform a Google Search for a synonym, until a simpler synonym was found.

PRELIMINARY STUDY

The factorial combination of both design dimensions, with their 4 or 5 levels, would yield 20 prototype variations, which was too many for comparison in an experimental study. So, it was essential for us to conduct a preliminary study to reduce this number. Thus, the main goal of our preliminary study was to identify potential design variations that may be preferred by people who are DHH for software to provide lexical simplification on-demand.

Materials

Reading Stimuli. To show our participants various designs for an ATS system, we needed to select a webpage with text that could be simplified. Using an article from the Smithsonian's website TweenTribune[1], which provides news articles at various levels of reading complexity, we created non-interactive high-visual-fidelity prototype videos, demonstrating each level of our two design dimensions discussed above. These videos were created using Safari in an iMac 27-inch desktop. Each video lasted approximately 10 seconds and showed a user requesting 3 simplifications on the first paragraph of the text. Given that this preliminary study was formative in nature, we did not attempt to display 20 videos to users, with all combinatorial possibilities, which may have been overwhelming. We instead provided a set of video prototypes for each design dimension. Thus, videos for the *user initiative* design dimension showed replacement of words when simplification was triggered, without any visual trace of which words had been replaced. The videos illustrating various levels of *change visibility* showed a word being clicked to trigger a replacement, without any visible decoration on the word beforehand to suggest replacing it.

Additional Demographic Questions. In addition to standard demographic questions, e.g. age, we also asked participants about how they identify themselves, e.g. Deaf, deaf, hard-of-hearing, etc. In addition, because there is great diversity among DHH individuals in the U.S., especially in regard to reading literacy skills, we also gathered additional demographic information about our participants:

Specifically, we also asked our participants to respond to an instrument that measured their English literacy skills and to respond to a psychology questionnaire for measuring an individual's inclination toward autonomy, since this was a focus of our work. Our intention was to provide future researchers with sufficient information about our participants, to facilitate any future replication of our study.

- **English Reading Literacy:** Following in the methodology of [5], we administered the sentence comprehension subtest of the Wide Range Achievement Test 4th Edition (WRAT4) [50] to measure literacy because it is brief, does not use audio stimuli, and has been previously validated with users who are DHH [33].
- **Autonomy:** After considering some relatively longer psychological instruments for measuring an individual's inclination towards autonomy, e.g. [14], we chose to administer the relatively brief Index of Autonomous Functioning (IAF) [49], consisting of 15 Likert items.

Data Collection Procedure

This formative study used an interview-based methodology with a set of video prototypes, illustrating various levels of design dimensions that were discussed with the user. After the participant signed a consent form and answered demographic questions, the interviewer explained the basic concept of ATS tools and the scenario in which the user may be reading a web page that contains unfamiliar words. Next, the interviewer discussed each design dimension and showed the participant all the videos for each. After each video, participants were asked to discuss the advantages and disadvantages of each; the goal of this was to encourage the participant to carefully consider each option. After all the videos for each dimension were shown, the participant was presented with a screen with small pictures of each levels simultaneously, and the interviewer asked participants to rate **how likely they were to use each version** on a 6-point scale (using an even number to avoid a middle point during this preliminary study). At the end of the study, participants were asked to provide feedback on the various visual text decorations that had appeared in the prototypes, to determine whether further adjustment of the designs would be needed. Finally, participants completed the IAF and the sentence comprehension subtest of the WRAT test was administered.

[1] https://www.tweentribune.com/article/teen/meet-female-inventor-behind-mass-market-paper-bags/

Recruitment and Participants

Deaf or Hard-of-Hearing participants were recruited through email and social media; participants met with a research assistant for a 70-minute study in a private office and received $40 cash compensation. This study was approved by our university's institutional review board for the protection of human subjects. The studies were conducted in English or American Sign Language (ASL), at the participant's preference. We recruited a total of 12 participants. Participants' self-identified gender included 7 male and 5 females, mean age of 24 (SD = 1.5). While 8 participants identified as culturally Deaf [32], 3 identified as Hard-of-hearing and 1 as deaf. Participants' standardized average WRAT score was 84.1 (SD= 14.03); the average WRAT score for U.S. adults is 100 [35]. The participants' average IAF score was 3.8 out of 5 (SD = 0.32).

Results and Discussion

The goal of this formative study had been to help us to narrow our set of prototypes to be included in our final experimental study. We knew in advance that, in our final experimental study, we wanted to include two "end points" of the design space: (1) a baseline condition in which users see a webpage without any text-simplification assistance provided and (2) a version in which the complex words are automatically replaced before the user sees the text without any visual indication of which words have been changed (thereby providing users with the least autonomy). Thus, our goal was to identify 1-2 additional prototypes to include in our final experiment. Notably, this formative study included few participants, and it was **not designed to be sufficiently powered for statistical difference testing**. Our goal was simply to avoid making an arbitrary choice about which prototypes to include our final experiment.

As Figure 2(a) illustrates, for the *user initiative* design dimension, users expressed most interest in the *toggle suggestions* prototype, while the *no suggestions* prototype was a close second. The *replace all* and *automatic* were the least preferred options. DHH users were most interested in designs that did not provide visual suggestions about which words can be simplified or enable users to toggle these visual suggestion indicators on or off.

As for the *change visibility* dimension, in Figure 2(a) the *trace* was first place, with *pop-up* second, *no signifier* in third, and *sidebar* in fourth. Thus, users were interested in designs that conveyed which words had been changed by leaving a visual trace (e.g. some decoration on the text) or by conveying simpler synonyms for a word using a pop-up (without replacing the word in the text itself).

While we had found the top options for *user initiative* (*toggle suggestions* and *no suggestions*) and for *change visibility* (*trace* and *pop-up*), our goal had been to identify two prototypes for our later experiment. So, we decided to combine design elements as follows: (a) *toggle suggestions* combined with *trace* and (b) *no suggestions* combined with *pop-up*. Our rationale was that since the *toggle suggestions*

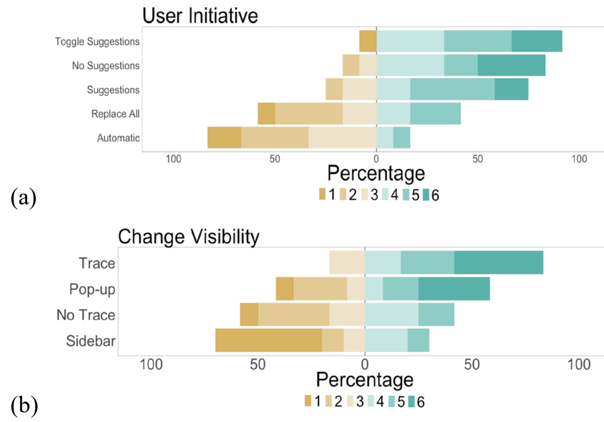

Figure 2. Participants' response to a 1-to-6 scalar question on how likely they were to use a system based on each level of the *user initiative* and *change visibility* design dimensions, from 1 (very unlikely) to 6 (very likely). This small study was insufficiently powered for statistical difference testing, but it was meant to guide prototype selection for the main study. Responses are displayed using a diverging stacked bar graph, as recommended in [39].

option already introduced text decoration into the interface, it was more natural for that to be paired with *trace* (which also used text decoration to indicate modified words).

Based on feedback from DHH participants in this formative study, we selected to use gray background color highlighting to indicate suggestions and yellow background color highlighting to indicate words which had already been replaced. As part of our preliminary study, we also discussed various design options with participants for how various elements of the user interface might appear, and we asked participants to share their thoughts about these alternatives. Participants responded negatively to the concept of using squiggly underlines, e.g. similar to the red underlines used in Microsoft Word to indicate spelling problems, because participants felt like such decoration indicated errors on the page. In addition, traditional underlining or font color changes were not preferred since they may look similar to hyperlinks on a web page.

EXPERIMENTAL STUDY

The goal of our main experimental study was to answer our research questions. RQ1 considered whether DHH users report subjective improvements in the readability or understandability of texts when using automatic reading-assistance tools with lexical simplification, and RQ2 considered whether DHH users' subjective impressions are affected by the degree of autonomy provided by the system. We examined several subjective measures that had been used in prior work on reading assistance tools, which are discussed below. In contrast to our preliminary formative study above, in this subsequent experimental study, users were provided with interactive prototypes. Guided by the results of our preliminary study, we implemented the following four interactive prototypes:

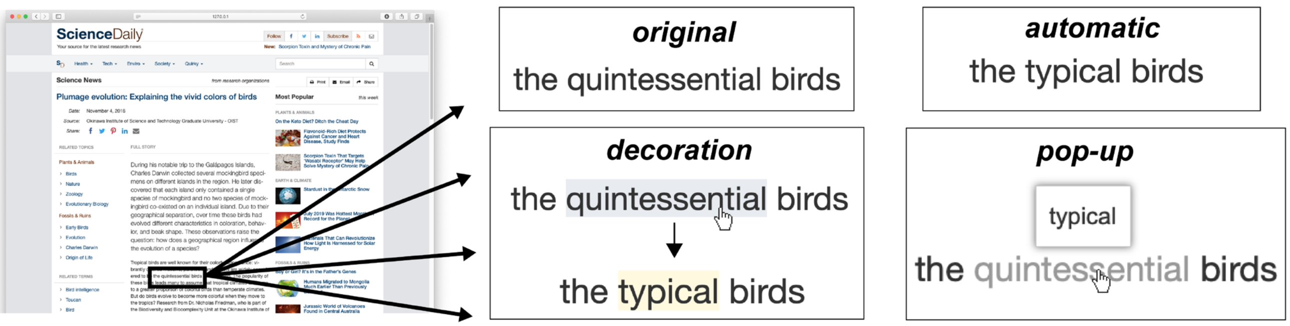

Figure 3: One of the articles used as stimuli in our experimental study, along with zoomed-in views of how text appeared under the four different conditions, e.g. clicking on a word in the *decoration* condition or hovering over a word in the *pop-up* condition.

1. A webpage with all the text in its original form without any simplifications applied nor available was provided as a baseline. We refer to this condition as ***original.***
2. We provided a version in which all complex words were already replaced, without any visible trace, before the user saw the text. In terms of the autonomy provided to the user, this was our least autonomous condition. We refer to this condition as ***automatic.***
3. We combined the *toggle suggestions* level for user initiative with the *trace* level for change visibility, to produce a system that provided decorations to suggest complex words or to show words that had been simplified. Since the system decorated and transformed text upon user request, this was our medium autonomy condition. We refer to this condition as ***decoration.***
4. We combined the *no suggestions* level for user initiative with the *pop-up* level for change visibility, to produce a system that provided a temporary pop-up near any word that can be simplified if the user hovers their pointer over it. This was our most autonomous condition given that the user could select any word and the text was never replaced. We refer to this condition as ***pop-up***.

Figure 3 illustrates each prototype. Readers should note that the name "pop-up" had referred to a level of *change visibility* in our formative study and now as the codename of one of the four prototypes in our final experimental study. Similarly, "automatic" had referred to a level of *user initiative* in our formative study and now as a codename for a prototype in our final experimental study. Given that the prototypes corresponding to each of these labels were essentially identical in both the formative and experimental studies, these seemed to be the most logical codenames for the prototypes, but we do note this identical nomenclature.

Materials

Reading Stimuli. In this study, we selected four texts from the ScienceDaily website, as they were freely available and were suitably complex, with Flesch-Kincaid grade levels ranging from 12.3 to 17.6. All of the articles consisted of research news about scientific studies concerning flying creatures (honey bees, birds and fruit flies). All four texts had a similar length: between 617 and 682 words, with an average of 654. After identifying complex words and their respective simplifications using our Wizard-of-Oz approach as described above, between 40 and 70 complex words were identified in each text. After applying all simplifications, the Flesch-Kincaid grade levels ranged from 11.6 to 16.4. These articles[2] originally contained a summary or abstract at the beginning, which we omitted to prevent our participants from simply reading the summary.

Subjective Evaluation Questions: To collect subjective feedback from our participants about each of the prototypes, we asked them to respond to three 5-point Likert-type questions for each prototype. These questions were selected based on having been used previously in prior work that had evaluated automatic text simplification systems. Specifically, in our study, participants indicated their agreement with each of the following:

- "This text was easy to read." Several prior studies have used a question like this, e.g. [37, 52];
- "I was able to understand this text well." Prior studies asked participants to indicate their perception of how well they understood the text, e.g. [37, 38, 52].

At the conclusion of the entire study, participants were reminded of the three prototypes they had used that had provided some text simplification, and they were asked to indicate whether they "would be likely to use" each. Our rationale for not asking this question about the baseline "original" condition was that asking the user how likely they would be to use a tool, after simply showing a user a web page without any text-simplification tool provided, would have likely led to confusion among participants.

Comprehension Questions. To encourage participants to pay attention to the texts that were displayed during the study

[2] https://www.sciencedaily.com/releases/2008/09/080909204550.htm
https://www.sciencedaily.com/releases/2019/06/190626160339.htm
https://www.sciencedaily.com/releases/2016/11/161104101848.htm
https://www.sciencedaily.com/releases/2019/08/190822141950.htm

and to attempt to use each prototype to understand the content, we informed participants that there would be a quiz after each text. We developed 3 comprehension questions for each text. These questions were modeled based on [16], using a main idea, main fact and incidental question. We made sure that none of the questions favored a particular version of a word that could be simplified.

Data Collection Procedure

Participants first signed consent forms and responded to demographic questions. Then, they were shown a set of instructions describing the operation of each prototype. While the four texts were presented in the same order, our four conditions were rotated using a balanced Latin-Squares schedule to ensure that neither variations in the complexity of the texts nor the order of presentation of the conditions affected our results. At the end of each text, participants responded to the two subjective questions (listed above) and then answered three comprehension questions about the text. At the conclusion of the study, participants were shown images of the three conditions that involved text simplification (i.e. *automatic*, *decoration,* and *pop-up*) in the same order that each individual participant had seen them. Then, we asked them to rate **how likely they were to use** each one of the conditions using a 5-point Likert-type question. After providing their response for each condition, they were asked open-ended questions about the rationale for their responses. At the end of the experiment, participants were administered the IAF and WRAT tests using an online form, and then informed of the Wizard-of-Oz nature of our system as part of a debriefing session.

Recruitment and Participants

This study was approved by our university's institutional review board. Participants were recruited through email and social media, and the advertisement included our screening criteria: identifying as Deaf or Hard-of-Hearing. Participants met with a research assistant for a 70-minute study in a private office and received $40 cash compensation. The studies were conducted in English or American Sign Language (ASL), at the participant's preference. We recruited 25 participants with mean age of 23.5 (SD = 2.27), including 10 self-identified as male, 14 as female and 1 as non-binary. While 16 participants identified as culturally Deaf [32], 3 identified as Hard-of-hearing and 6 as deaf. The participants' average IAF score was 3.68 out of 5 (SD = 0.36), with an average WRAT score of 83.04 (SD = 12.38).

Results

Figure 4 displays participants responses to the subjective questions in our study that compared all four conditions, including the baseline condition. In this graph, significant pairwise differences are indicated with asterisks as follows: *** if p<0.001, ** if p<0.01, or * if p<0.05. The statistical analysis performed for each question is described below.

After reading each text, we asked participants to indicate how much they agreed with: "This text was easy to read." A Friedman test indicated a significant difference (χ^2 = 7.9682, p = 0.047), yet post-hoc pairwise comparison using Wilcoxon Signed Rank tests with Bonferroni corrections did not reveal any significant pairwise differences.

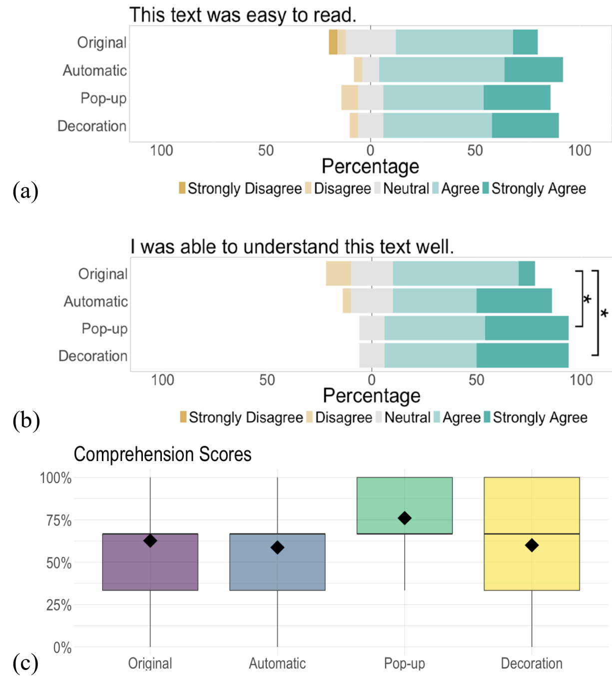

Figure 4. Participants' responses to questions about all four conditions in the experimental study, including subjective Likert-scale responses for (a) *the text was easy to read* and (b) *I was able to understand this text well*, with significant pairwise differences marked with asterisks (* p <0.05). In (c), analysis on objective comprehension questions did not reveal any significant differences between the four conditions.

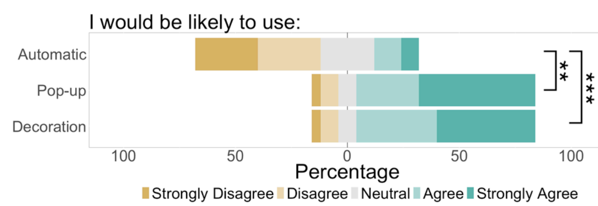

Figure 5. Participants' agreement to a Likert-scale question, presented at the end of the study, as to whether they would be likely to use each of the three simplification conditions, with asterisks marking significant pairwise differences (p<0.01, *** p<0.01).**

We asked participants to indicate how much they agreed with: "I was able to understand this text well" for each condition. A Friedman test indicated a significant difference (χ^2 = 16.261, p-value = 0.001), and post-hoc pairwise comparison using Wilcoxon Signed Rank tests with Bonferroni corrections revealed significant pairwise differences between the following: original / pop-up (p = 0.042) and original / decoration (p = 0.012).

After the two subjective questions, participants answered three comprehension questions about information from each article, via multiple-choice questions. The average scores for these questions across the four conditions are shown in Figure 4(c). One-Way repeated measures analysis of variance (ANOVA) did not reveal any significant differences between the conditions.

Next, participants were reminded of the 3 simplification conditions (with pictures of each), and they were asked to indicate how much they agreed with the statement "I would be likely to use a system like this." Figure 5 displays participants' responses. A Friedman test indicated a significant difference ($\chi^2 = 18.731$, $p = 8.562e\text{-}05$). Post-hoc pairwise Wilcoxon Signed Rank tests with Bonferroni corrections indicated significant differences between some pairs: *automatic* / *pop-up* ($p = 0.00461$) and *automatic* / *decoration* ($p = 0.00049$). Finally, participants shared feedback comments about all four conditions in the study.

DISCUSSION

RQ1: Perceived Benefit from Lexical Simplification. We did not observe any significant differences across conditions in users' scores on objective comprehension questions, but users indicated stronger agreement with "I was able to understand this text well," for both the *pop-up* and *decoration* conditions, compared to the *original* condition (without any lexical simplification provided). This finding is in line with prior research on lexical-simplification tools for users with dyslexia, in which no difference was found on objective comprehension scores, yet participants reported subjectively that simplification provided on-demand made texts easier to understand [36]. Additional research is needed to determine if there is any objective benefit from this technology, i.e. as measured through comprehension questions or some other method. A challenge is that prior methodological research on evaluation of assistive technologies for users who are DHH with diverse literacy-levels has suggested that some measurement instruments (types of questions used in studies) exhibit bias in regard to DHH users' literacy [5]. Thus, additional methodological research is needed on how to measure the effect of lexical simplification among adults who are DHH at various literacy levels. Assuming that an objective benefit of this technology can be established, then there would be motivation for future computational linguistic research, specifically methods for tailoring the output of automatic simplification tools for these users, which has been important for other groups [31].

RQ2: Providing Simplifications On-Demand. In our results above, the *pop-up* and *decoration* conditions received higher scores on the "likely to use" question, in comparison to the scores for the *automatic* condition. Thus, DHH users believed they would be more likely to use reading-assistance tools that provide simplifications on-demand, suggesting that users preferred greater autonomy in such tools. This finding aligns with speculation by authors in [36] that some of the benefits they observed from simplification tools for users with dyslexia were due to one of their prototypes providing assistance on-demand [36].

When asked about the reasons for their responses in our study, DHH participants mentioned a variety of reasons for why they preferred the on-demand forms of simplification (*pop-up* or *decoration*), including: a desire to have a choice, a desire to see the original text, and a desire to learn new words. When commenting on their preference for the *pop-up* condition, for instance, participant P13 commented:

"I may already know the word and may choose to skip the simpler word conversion but that choice wasn't made for me already. I can choose whether to do [it] or not." - P13

In terms of the desire to see the original text, participant P24, e.g., expressed concerns about missing important terminology for a class if the text is automatically replaced:

"I am not so sure about this method because if I am reading this article for a class, I would prefer to know all the terms that will be brought to class from this article." - P24

In particular, the issue of learning new vocabulary was the most frequently mentioned reason by participants, e.g.:

"I like to know what words were originally used. I have no way of know what words were used when it was automatically simplified, therefore no way to learn." - P21

Our results expand on the HCI literature on autonomy by illustrating that, in the context of lexical simplification, DHH users prefer control in requesting simplifications on-demand and in seeing what portions of text have been transformed. This finding is in line with prior research that had found that increasing users' sense of autonomy within a software environment may increase their engagement with it [19]. Our DHH participants indicated they would be more likely to use tools providing them with greater autonomy.

Our results also have implications for researchers in the field of ATS who want to test the quality of their simplification methods, by highlighting that whether those simplifications are provided on-demand may have an effect on the subjective judgements of their users.

Tradeoffs between On-Demand Prototypes. While DHH users preferred the two on-demand prototypes (*pop-up* or *decoration*), as compared to the other conditions (*original* or *automatic*), we did not observe any significant difference between their preferences for *pop-up* and *decoration*. In open-ended comments, DHH participants did mention some trade-offs between these two conditions, which relate to the different levels of autonomy provided by each. For instance, when discussing the *pop-up* condition, participants mentioned that the pop-ups could be distracting, and some participants disliked having to hover over a word to find out whether the system was capable of providing any simplifications for it. However, other participants indicated that they appreciated that it kept the original text and allowed them to see both the complex word and its simpler synonym at the same time. For instance, participant P9 commented:

"I like [the pop-up] more than the highlighting option, because it allows me to basically know an easier synonym of a difficult word, but allows me to reread the text at the level its at rather than a simplified version since I will know what the difficult word means." - P9

As for the *decoration* condition, participants found it to be less distracting than the *pop-up* condition, and they mentioned that it was helpful to know which words they could click on. Participant P16, for instance said:

"I actually like this method because it allowed me to learn more vocabularies and it was not distract at all like pop up. If there is a word that I am curious to know the meaning and I would click it to see it. Also, unlike pop up, you don't have to move mouse around to find which word that has pop up because highlight is already there and it doesn't bother me when I read the article." - P4

However, some participants commented that clicking may take longer than hovering over words, and some were concerned about not being able to see both the original word and its simpler replacement word simultaneously.

LIMITATIONS AND FUTURE WORK

There were several limitations in our study, which may limit the generalizability of this work or suggest future research:

- We have only investigated one type of text simplification (lexical), and it is important to clarify that *we are not implying that lexical simplification alone is sufficient for addressing text readability challenges among DHH users.* Instead, we chose to investigate lexical simplification in this study because we had perceived a gap in the literature, i.e. with prior computing research having previously investigating *syntactic* simplification among adults who are DHH [27]. We foresee the need for future research to investigate the preferences among these users for lexical, syntactic, and hybrid/combined simplification tools.
- While we have provided information about the level of difficulty of the texts shown in our study, which were rotated across conditions, these texts did not have identical Flesch-Kincaid scores. To further mitigate this lack of control over the texts' difficulty, we shall use a Greco-Latin square design to rotate conditions in future work.
- Our study did not include a baseline condition that displayed original text with decoration around complex words, without any text simplification provided. Such a baseline could have revealed whether the observed effect on perceived readability might arise from text decoration alone, which could have led to response bias [46], rather than from actually providing lexical simplification.
- This study included one form of objective measure (comprehension questions), but future work can explore other measurements (e.g. reading speed or eye-tracking).
- The responses gathered in our study may not be representative of a group of DHH individuals with a different distribution of ages, genders, identity (Deaf, deaf, or hard of hearing), etc. To enable replication of our work or for readers to better interpret our findings, we have provided these demographic characteristics, as well as WRAT scores (reflecting English reading literacy level) and IAF (reflecting the personality characteristics of the individuals in the study, in regard to autonomy). A further study with a larger group of participants or individuals with a different range of such characteristics may be needed to understand the range of opinions across an even more diverse set of DHH users.
- We used a Wizard-of-Oz method for identifying complex words and providing simplifications; there is a risk that the output provided may differ from an automatic system. Specifically, an automatic system may sometimes provide erroneous word replacements (that change the meaning of the text), and a future study would be needed to investigate DHH participants' evaluation of on-demand simplification systems with some errors in the simplifications provided.
- We used an initial study to select a subset of possible levels of user initiative and change visibility to include in our final study, but that first study was underpowered and used formatively in this work to avoid our arbitrarily selecting prototypes for comparison. In future work, a full factorial experiment with a larger sample size may enable us to determine the best combinations of both user initiative and change visibility. A future study could also focus on how the trade-offs of varying the level of user autonomy may make certain variations more appropriate for specific conditions (e.g. different kinds of webpages or tasks).

In current work, we are investigating properties that affect the complexity of words for DHH readers, as well as how to best evaluate text simplifications with this user group.

CONCLUSION

Our study investigated lexical simplification in automatic reading-assistance tools for English text for DHH users. Our study did not reveal any measurable, objective benefit for the comprehension of texts among DHH readers, but it did reveal that users perceived certain designs incorporating lexical simplification as beneficial and that providing a sense of autonomy influences DHH users' acceptance of such reading-assistance tools. Specifically, having investigated a variety of designs that vary in the degree of autonomy they provide (i.e. control of what words are replaced and visibility of past replacements), we found that DHH adults indicate they are more likely to use systems with greater autonomy. Beyond the implications of this finding for research on reading-assistance tools, this contributes more broadly to HCI research literature on autonomy, i.e. providing evidence of its benefits for these users and task, while identifying some tradeoffs that arise in this context.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under award No. 1822747.

REFERENCES

[1] J. Albertini and C. Mayer. 2011. Using Miscue Analysis to Assess Comprehension in Deaf College Readers. *Journal of Deaf Studies and Deaf Education* 16, 1 (2011), 35–46. DOI:http://dx.doi.org/10.1093/deafed/enq017

[2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 3, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300233

[3] Mahmoud Azab, Chris Hokamp, and Rada Mihalcea. 2015. Using Word Semantics To Assist English as a Second Language Learners. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. Association for Computational Linguistics, Denver, Colorado, 116–120. DOI: http://dx.doi.org/10.3115/v1/N15-3024

[4] A. Banner and Y. Wang. 2011. An Analysis of the Reading Strategies Used by Adult and Student Deaf Readers. *Journal of Deaf Studies and Deaf Education* 16, 1 (2011), 2–23. DOI:http://dx.doi.org/10.1093/deafed/enq027

[5] Larwan Berke, Sushant Kafle, and Matt Huenerfauth. 2018. Methods for Evaluation of Imperfect Captioning Tools by Deaf or Hard- of-Hearing Users at Different Reading Literacy Levels. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 91, 12 pages. DOI: http://dx.doi.org/10.1145/3173574.3173665

[6] Joachim Bingel, Gustavo Paetzold, and Anders Søgaard. 2018. Lexi: A tool for adaptive, personalized text simplification. In *Proceedings of the 27th International Conference on Computational Linguistics*. Association for Computational Linguistics, Santa Fe, New Mexico, USA, 245–258. https://www.aclweb.org/anthology/C18-1021

[7] Debra L Blackwell, Jacqueline W Lucas, and Tainya C Clarke. 2014. *Summary Health Statistics for U.S. Adults: National Health Interview Survey*, 2012. National Center for Health Statistics. https://www.ncbi.nlm.nih.gov/pubmed/24819891

[8] Stefan Bott, Horacio Saggion, and David Figueroa. 2012. A Hybrid System for Spanish Text Simplification. In *Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies (SLPAT '12)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 75–84. http://dl.acm.org/citation.cfm?id=2392855.2392865

[9] Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler and Meredith Morris. 2019. Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. ACM, New York, NY, USA.

[10] Rafael A. Calvo, Dorian Peters, Daniel Johnson, and Yvonne Rogers. 2014. Autonomy in Technology Design. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 37–40. DOI:http://dx.doi.org/10.1145/2559206.2560468

[11] Shiwei Cheng, Zhiqiang Sun, Lingyun Sun, Kirsten Yee, and Anind K. Dey. 2015. Gaze-Based Annotations for Reading Comprehension. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1569–1572. DOI: http://dx.doi.org/10.1145/2702123.2702271

[12] Carol Convertino, Georgianna Borgna, Marc Marschark, and Andreana Durkin. 2014. Word and World Knowledge Among Deaf Learners With and Without Cochlear Implants. *The Journal of Deaf Studies and Deaf Education* 19, 4 (08 2014), 471–483. DOI: http://dx.doi.org/10.1093/deafed/enu024

[13] Beth Davey and Susan King. 1990. Acquisition of word meanings from context by deaf readers. American Annals of the Deaf 135, 3 (1990), 227– 234.

[14] Edward L Deci and Richard M Ryan. 1985. The general causality orientations scale: Self-determination in personality. *Journal of research in personality* 19, 2 (1985), 109–134.

[15] Siobhan Devlin and Gary Unthank. 2006. Helping Aphasic People Process Online Information. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '06)*. ACM, New York, NY, USA, 225–226. DOI: http://dx.doi.org/10.1145/1168987.1169027

[16] Mary C Dyson and Mark Haselgrove. 2001. The Influence of Reading Speed and Line Length on the Effectiveness of Reading from Screen. *Int. J. Hum.-Comput. Stud.* 54, 4 (April 2001), 585–612. DOI: http://dx.doi.org/10.1006/ijhc.2001.0458

[17] Maxine Eskenazi, Yibin Lin, and Oscar Saz. 2013. Tools for non-native readers: the case for translation and simplification. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*. Association for Computational

Linguistics, Atlanta, Georgia, 20–28. https://www.aclweb.org/anthology/W13-1503

[18] Organization for Economic Co-operation and Development. 2012. *Survey of Adult Skills (PIAAC)*. http://skills.oecd.org/SurveyofAdultSkillsUS.pdf

[19] Matthew Ford, Peta Wyeth, and Daniel Johnson. 2012. Self-determination Theory As Applied to the Design of a Software Learning System Using Whole-body Controls. In *Proceedings of the 24th Australian Computer- Human Interaction Conference (OzCHI '12)*. ACM, New York, NY, USA, 146–149. DOI:http://dx.doi.org/10.1145/2414536.2414562

[20] Brittany L Freel, M Diane Clark, Melissa L Anderson, Gizelle L Gilbert, Millicent M Musyoka, and Peter C Hauser. 2011. Deaf Individuals' Bilingual Abilities: American Sign Language Proficiency, Reading Skills, and Family Characteristics. *Psychology* 2, 1 (2011), 18–23.

[21] Batya Friedman and Helen Nissenbaum. 1996. User Autonomy: Who Should Control What and when?. In *Conference Companion on Human Factors in Computing Systems (CHI '96)*. ACM, New York, NY, USA, 433–. DOI:http://dx.doi.org/10.1145/257089.257434

[22] Carrie Lou Garberoglio, Stephanie Cawthon, and Adam Sales. 2017. *Deaf People and Educational Attainment in the United States: 2017*. Technical Report. Washington, DC.

[23] Florian Güldenpfennig, Peter Mayer, Paul Panek, and Geraldine Fitzpatrick. 2019. An Autonomy-Perspective on the Design of Assistive Technology Experiences of People with Multiple Sclerosis. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 127, 14 pages. DOI: http://dx.doi.org/10.1145/3290605.3300357

[24] Kentaro Inui, Atsushi Fujita, Tetsuro Takahashi, Ryu Iida, and Tomoya Iwakura. 2003. Text Simplification for Reading Assistance: A Project Note. In *Proceedings of the Second International Workshop on Paraphrasing - Volume 16 (PARAPHRASE '03)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 9–16. DOI: http://dx.doi.org/10.3115/1118984.1118986

[25] Sushant Kafle, Abraham Glasser, Sedeeq Al-khazraji, Larwan Berke, Matthew Seita, Matt Huenerfauth. 2019. Artificial Intelligence Fairness in the Context of Accessibility Research on Intelligent Systems for People who are Deaf or Hard of Hearing. The *ACM ASSETS 2019 Workshop on AI Fairness for People with Disabilities*. arXiv:1908.10414 [cs.HC]

[26] Jumpei Kobayashi and Toshio Kawashima. 2019. Paragraph-based Faded Text Facilitates Reading Comprehension. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 162, 12 pages. DOI: http://dx.doi.org/10.1145/3290605.3300392

[27] Poorna Kushalnagar, Scott Smith, Melinda Hopper, Claire Ryan, Micah Rinkevich, and Raja Kushalnagar. 2018. Making cancer health text on the Internet easier to read for deaf people who use American Sign Language. *Journal of Cancer Education* 33, 1 (2018), 134–140.

[28] John L Luckner and C Michele Handley. 2008. A summary of the reading comprehension research undertaken with students who are deaf or hard of hearing. *American Annals of the Deaf* 153, 1 (2008), 6–36.

[29] Mounica Maddela and Wei Xu. 2018. A Word-Complexity Lexicon and A Neural Readability Ranking Model for Lexical Simplification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Brussels, Belgium, 3749–3760. DOI:http://dx.doi.org/10.18653/v1/D18-1410

[30] Marc Marschark, John A. Albertini, and Harry G. Lang. 2002. *Educating deaf students: from research to practice*. Oxford University Press.

[31] Lourdes Moreno, Rodrigo Alarcon, Isabel Segura-Bedmar, and Paloma Martínez. 2019. Lexical Simplification Approach to Support the Accessibility Guidelines. In *Proceedings of the XX International Conference on Human Computer Interaction (Interacción '19)*. ACM, New York, NY, USA, Article 14, 4 pages. DOI: http://dx.doi.org/10.1145/3335595.3335651

[32] Carol Padden, Tom Humphries, and Carol Padden. 2009. Inside deaf cul- ture. Harvard University Press.

[33] Gustavo Paetzold and Lucia Specia. 2016. Understanding the Lexical Simplification Needs of Non-Native Speakers of English. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. The COLING 2016 Organizing Committee, Osaka, Japan, 717–727. https://www.aclweb.org/anthology/C16-1069

[34] S. J. Parault and H. M. Williams. 2010. Reading Motivation, Reading Amount, and Text Comprehension in Deaf and Hearing Adults. *Journal of Deaf Studies and Deaf Education* 15, 2 (2010), 120–135. DOI: http://dx.doi.org/10.1093/deafed/enp031

[35] LeAdelle Phelps and Barbara Jane Branyan. 1990. Academic achievement and nonverbal intelligence in public school hearing-impaired children. *Psychology in the Schools* 27, 3 (1990), 210–217.

[36] Luz Rello, Ricardo Baeza-Yates, Stefan Bott, and Horacio Saggion. 2013. Simplify or Help?: Text Simplification Strategies for People with Dyslexia. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility (W4A '13)*. ACM, New York, NY, USA, Article 15, 10 pages. DOI:http://dx.doi.org/10.1145/2461121.2461126

[37] Luz Rello, Roberto Carlini, Ricardo Baeza-Yates, and Jeffrey P. Bigham. 2015. A Plug-in to Aid Online Reading in Spanish. In *Proceedings of the 12th Web for All Conference (W4A '15)*. ACM, New York, NY, USA, Article 7, 4 pages. DOI:http://dx.doi.org/10.1145/2745555.2746661

[38] Luz Rello, Martin Pielot, and Mari-Carmen Marcos. 2016. Make It Big!: The Effect of Font Size and Line Spacing on Online Readability. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16). ACM, New York, NY, USA, 3637–3648. DOI: http://dx.doi.org/10.1145/2858036.2858204

[39] Naomi B Robbins, Richard M Heiberger, and others. 2011. Plotting Likert and other rating scales. In *Proceedings of the 2011 Joint Statistical Meeting*. 1058–1066.

[40] Horacio Saggion, Sanja Štajner, Stefan Bott, Simon Mille, Luz Rello, and Biljana Drndarevic. 2015. Making It Simplext: Implementation and Evaluation of a Text Simplification System for Spanish. *ACM Trans. Access. Comput*. 6, 4, Article 14 (May 2015), 36 pages. DOI: http://dx.doi.org/10.1145/2738046

[41] Tenace Setor, Damien Joseph, and Shirish C. Srivastava. 2015. Professional Obsolescence in IT: The Relationships Between the Threat of Professional Obsolescence, Coping and Psychological Strain.. In *Proceedings of the 2015 ACM SIGMIS Conference on Computers and People Re- search (SIGMIS-CPR '15)*. ACM, New York, NY, USA, 117–122. DOI: http://dx.doi.org/10.1145/2751957.2751962

[42] Matthew Shardlow. 2014a. Out in the Open: Finding and Categorising Errors in the Lexical Simplification Pipeline. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. European Languages Resources Association (ELRA), Reykjavik, Iceland, 1583–1590. http://www.lrecconf.org/proceedings/lrec2014/pdf/479 Paper.pdf

[43] Matthew Shardlow. 2014b. A survey of automated text simplification. *International Journal of Advanced Computer Science and Applications* 4, 1 (2014), 58–70.

[44] Svetlana Sheremetyeva. 2014. Automatic text simplification for handling intellectual property (the case of multiple patent claims). In *Proceedings of the Workshop on Automatic Text Simplification-Methods and Applications in the Multilingual Society (ATS-MA 2014)*. 41–52.

[45] C. B. Traxler. 2000. The Stanford Achievement Test, 9th Edition: National Norming and Performance Standards for Deaf and Hard-of-Hearing Students. *Journal of Deaf Studies and Deaf Education* 5, 4 (Jan 2000), 337–348. DOI:http://dx.doi.org/10.1093/deafed/5.4.337

[46] Shari Trewin, Diogo Marques, and Tiago Guerreiro. 2015. Usage of Subjective Scales in Accessibility Research. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS '15),* 59–67. https://doi.org/10.1145/2700648.2809867

[47] Gerard G Walter. 2010. *Deaf and Hard-of-Hearing Students in Transition: Demographics with an Emphasis on STEM Education*. National Technical Institute for the Deaf, Rochester, NY.

[48] Willian Massami Watanabe, Arnaldo Candido Junior, Vinícius Rodriguez Uzêda, Renata Pontin de Mattos Fortes, Thiago Alexandre Salgueiro Pardo, and Sandra Maria Aluísio. 2009. Facilita: Reading Assistance for Low-literacy Readers. In *Proceedings of the 27th ACM International Conference on Design of Communication (SIGDOC '09)*. ACM, New York, NY, USA, 29–36. DOI:http://dx.doi.org/10.1145/1621995.1622002

[49] Netta Weinstein, Andrew K Przybylski, and Richard M Ryan. 2012. The index of autonomous functioning: Development of a scale of human autonomy. *Journal of Research in Personality* 46, 4 (2012), 397–413.

[50] Gary S Wilkinson and Gary J Robertson. 2006. Wide range achievement test (WRAT4). *Lutz, FL: Psychological Assessment Resources* (2006).

[51] Seid Muhie Yimam and Chris Biemann. 2018. Par4Sim – Adaptive Paraphrasing for Text Simplification. In *Proceedings of the 27th International Conference on Computational Linguistics. Association for Computational Linguistics*, Santa Fe, New Mexico, USA, 331–342. https://www.aclweb.org/anthology/C18-1028

[52] Chen-Hsiang Yu and Robert C. Miller. 2010. Enhancing Web Page Readability for Non-native Readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2523–2532. DOI:http://dx.doi.org/10.1145/1753326.1753709