

Why Bother with Writers? Towards “Good Enough” Technologies for Supporting Individuals with Dyslexia

Vibeke Rønneberg, Christer Johansson, Margunn Mossige, Mark Torrance and Per Henning Uppstad

Introduction

Existing technological solutions to aid writing do not consider *how* the text is produced. Most available tools are targeting a final product and the correctness of this product, and often fail to support the production of the text. Popular alternatives such as LaTeX may help the writer to focus on the text, and are one step towards separating writing and formatting. However, there is no explicit support for writing flow. The system needs to consider *when* a writer needs feedback. No writer likes to be interrupted when searching for an expression, or finishing a sentence or a phrase. However spell checkers do not wait until a writer has finished a sentence before a spelling error is underlined with red. This is related to another problem; the abundance of negative feedback and a matching scarcity of positive feedback given by spell checkers. We argue that the potential for development of writing skills is not currently supported by existing tools, and this may be a serious shortcoming for people with Specific Learning Disabilities (SLD) associated with reading and writing.

In a broad sense, writing always has involved technology. The term technology—originating from the Greek word *Technē*, “craftsmanship”—has been applied to historical phenomena that span from the introduction of the alphabet to digital solutions that underlie writing tools. The digital age has put forward new possibilities for the support of writing, mainly through options linked to digital word processors. Word processing makes changing and revising text a much easier and less costly activity. However, there has been a tendency for the design and development of digital writer-support tools to be driven by what technology permits rather than what writers need. Dixon et al. (2004, chapter 4) discuss this in relation to computer interfaces for text production. The WYSIWYG has become the model for direct manipulation interface that arguably has design advantages as well as psychological advantages, but for text production it puts quite high expectations on the user for understanding and handling the

full complexity of text production in a simplistic model based on direct manipulation. The system could arguably do more to help the writer to express ideas through words, but such a model would incorporate the system as a writing partner.

In this chapter the short history of digital writing technology is described as having a focus on getting single words written in a correct manner. Providing a full menu of visible choices (for the writer to read) and showing the writer what the end product will look like throughout the whole writing process have been guiding principles for human-machine interaction since early graphical interfaces became widely available (most notably through the first Macintosh computers), but there were earlier attempts at a “writer’s workbench” (cf. de Smedt, 2009). This design principle is often abbreviated as WYSIWYG—*What you see is what you get*. Typical of the tools that have emerged during the digital shift is the spell checker, with its focus on detecting spelling mistakes and giving choices regarding revision.

We approach the issue of how technological aids can meet the needs of writers with Specific Learning Disabilities (SLD) and writers with dyslexia in particular by asking *what is wrong with WYSIWYG*, in order to pinpoint challenges related to overarching ideas in the field. Our intention with the chapter as a whole is to draw the contours of a different principle for future technological tools for writers with SLD as well as ordinary writers, who need to improve their writing fluency: We propose *what you get is what you need*—WYGIWYN.

Before we discuss WYSIWYG as being the dominant premise provider for text tools, let us highlight some of the challenges people with dyslexia face every day. First, they struggle with spelling: their written texts in general contain more spelling errors than texts written by typical writers (Coleman, Gregg, McLain, & Bellair, 2009; Tops, Callens, van Cauwenberghe, Adriaens, & Brysbaert, 2013) and their struggling with spelling seems to persist (Bruck, 1993). However, this might not be their most persistent problem. Tops and colleagues (Tops et al., 2013) find that the overall *text quality* of students with dyslexia is generally poorer. Moreover, as a group, people with reading and writing difficulties write more slowly than typical writers (Wengelin, 2002; Wengelin & Strömqvist, 2000). Their writing behaviour is characterized by an attention towards the word currently being written or just written (Wengelin, 2007). This focus on single words might distract writers from focusing on other parts of text production. To write a coherent and fluent text implies that you are able to remember what you have written and plan ahead, while simultaneously dealing with the current situation. Confident writers know that they can handle the situations that are likely to emerge, and they spend their effort on planning ahead rather

than worrying about what can go wrong. For some writers, fluency is limited by focusing on single words and the possibilities and choices that follow.

A substantial aspect of writing is conveniently switching between focus on a larger plot or meaning, where one is heading, and monitoring what is about to be produced here and now. The alternation between looking ahead and monitoring what is close can be described as a ground-breaking hermeneutic for considering a whole, while continuously checking whether this whole is challenged by the evolving parts. There are reasons to believe that individuals with dyslexia face problems in such switching of focus, as detected by for example Rapid Automatic Switching (RAS), (Berninger, Abbott, Thomson, & Raskind, 2001, as reported in Berninger, Nielsen, Abbot, Wijsman & Raskind, 2008).

We know that persistent problems are intensified by lack of experience with text, i.e., problems in decoding orthography make those with SLD read less than readers without difficulties. We also know that the grand leap in reading development comes when children start to read a lot—input matters. It is possible to reason in similar ways regarding writing, which makes extensive writing both a goal and a means to the goal. A possible pathway to support this development is a theory that supports the development of concrete procedures or tools that value writing fluency. Following this line of arguments, we argue that better timed, more focused and positive feedback may give the writer the option of choosing more convenient paths, paths that give the writer new and richer experience with writing a text. The effect of actually enjoying the activity should not be underestimated as it is beneficial on many levels, not least in making the writer seeking out more opportunities to write rather than avoid them.

What is Wrong with WYSIWYG

WYSIWYG imposes large cognitive demands on the writer to take control of the final output. This involves identifying what action to take or tool to use and then, often non-trivially, navigating the systems menus to make this happen. The interface affords in this way a situation where the writer is expected to produce a text that is simultaneously correct, coherent, and aesthetically presented. The number of choices provided during writing may be a challenge to both working memory and visual capacity of the typical writer, and even more so for a challenged writer. On top of this is the substantial number of tools in typical word processors that are devoted to revision and correction. However, writers with SLD face difficulties at a more basic level; to get their thoughts

and ideas down on paper in an understandable manner. The problem is not primarily that their texts are not understandable to readers—the problem is that the writer may experience uncertainty, stress and lack of self-confidence when faced with written tasks (Mossige, Rønneberg, Johansson, Uppstad, & Torrance, n.d.). While there are few investigations on the relation between stress, self-confidence and outcome on writing performance, Alexander-Passe (2006) discuss this in terms of coping mechanisms. Because of the focus on the final product, the design principle for text tools runs the risk of being more concerned with the needs of the readers than the needs of the writer.

Word processing programs are used extensively in Norwegian schools and work life, and writers use word processors from an increasingly younger age. In addition to producing legible letters and allowing for easy revision, most word processors today include some sort of a spell checker. Three meta-studies have found that use of word processors improves writing quality for students, compared to writing by hand (Bangert-Drowns, 1993; Goldberg, Russell, & Cook, 2003; Graham & Perin, 2007).

Because writers with SLD struggle with transcription and revision, word processors and spell checkers should prove to be advantageous for this group in particular. However, the spell checkers typically integrated in word processors may trigger more uncertainty by only providing negative feedback, which makes it hard to see what is good enough. Morphy and Graham (2012), in a meta-analysis of the effects of word processing for students who were weak writers, weak readers, or both, in grades 1–12, found that overall word processors had a positive effect on the quality of the produced texts. Comparing different kinds of word processors, they found that those that included additional support in *planning*, *drafting*, and *revision* showed significant gains in writing quality. Morphy and Graham conclude that weak writers/readers should use word processing as their primary tool when composing, especially those word processors that provide feedback on the quality of the text, or support planning, drafting and revision.

Although word processors and other computer tools have their limitations, MacArthur (2006) argued that all transcription tools remove one burden, but add a burden on working memory, or require training. Typically, the spell checker gives immediate and frequent feedback about errors. Uncertainty comes when the writer's attention is guided away from meaning production to a focus on a word's orthography. MacArthur argues that although spell checkers may help struggling writers correct more errors than they do with no such help, struggling writers are not always able to recognize the correct spelling. Having to read and choose from a list of alternatives that typically differ by only one or two letters might impose a new burden for struggling writers (MacArthur,

2006; MacArthur, Graham, Haynes, & De La Paz, 1996) and severely misspelled words may not even be listed (which occurred in 42% of cases). The difficulty of choosing the right word is rather evident, giving what we know about poorer reading skills on average for individuals with dyslexia. Another limitation is that homonyms or other real words that are the result of spelling errors are not detected (which occurred in 37% of cases).

It is striking how *phonologically close* some severe misspellings may be (cf. *fernitcer* / *furniture* (MacArthur, 2006), which is close if pronounced but far in edit distance, meaning that several operations must be done to the misspelling in order to reach the target word), in light of many students being diagnosed with dyslexia on the basis of difficulties with phoneme-grapheme mapping or other phonological deficits. We find many similar examples in our Norwegian data, although there are other error types as well. This highlights the fact that sound coding, rather than orthographic coding, may help catch some errors. More advanced word prediction may help here, by predicting the intended word based on context words to the left and right, as well as the characters in the word to be corrected. Additionally, the likelihood of the correct word is affected by topical context words in the text to be corrected.

Some word prediction programs incorporate phonetic spellings of words in their algorithms, and these could be helpful for struggling writers. However, research on early word prediction programs shows that these slow writers down (Lewis, Graves, Ashton, & Kieley, 1998). Also in using these programs, as with spell checkers, writers do not always recognize the correct spelling, unless the system also provides speech synthesis to help writers read the text. The solution to the limitations, especially the demand on reading skills, of word prediction has been to implement text-to-speech synthesis, where the writer may listen to the words in the list in order to decide the right one. Speech synthesis reads the written text aloud, and can therefore support students, with or without reading disabilities, when they revise their texts. Writers get helpful auditory feedback on misspelled words, as the program attempts to pronounce the words phonetically. If a pronunciation seems strange, it is an indication that there is a spelling error. However, using speech synthesis to correct individual words is time consuming and imposes the burden of listening for misspellings, and if done while producing text may interrupt the flow of writing.

Still, research supports the use of word prediction programs (referred to as “word completion” by MacArthur) for students with spelling difficulties. MacArthur (2006; pp. 17–18) reports some positive results, although when the algorithm used a larger vocabulary, the complexity of selecting the correct word completion increased to a level that was difficult for students with reading difficulties. Matching the active vocabulary to the immediate needs of the

students seems to be a key ingredient in making word completion work. We argue that the task of predicting a word can be saved until an editing stage where more restrictive context is available. The aim is to be able to either determine the intended word or provide very few, but highly likely, alternatives, to make the correction procedure less demanding of reading and monitoring skills. Recently, word completion and word prediction programs have benefited from using extremely large data sets and machine learning algorithms (van den Bosch 2005, 2006, 2011). The task of predicting the intended words using cues that are available from either the product (e.g., the lexical material around the word in focus) or the process (e.g., keystroke latency intervals) is intrinsically a task within the scope of computational linguistics and natural language processing.

Current solutions to word completion and speech synthesis can be considered unfortunate if the goal is to help the writer focus on developing content. Word completion can be challenging in that it forces writers with dyslexia to consider orthography, while speech synthesis is time consuming. Both solutions are, in accordance with the WYSIWYG-principle, in guiding the writer to correct here and now, using all means possible to get the text looking perfect. They can however block the writer from completing the meaning that is under construction.

WYGIWYN—What You Get is What You Need

When we here suggest a new standard—WYGIWYN (*what you get is what you need*), we claim that the writer, and the writer with SLD in particular, needs *less* rather than more information. A writing tool that emphasizes fluency therefore needs to find ways of reaching this goal that are in line with a WYGIWYN philosophy.

As stated earlier, we aim to draw the contours of a principle that differs from the way in which writer support is currently typically conceptualised. Figure 7.1 shows the shift in emphasis from correction to fluency achieved by WYGIWYN. This might appear at first sight as splitting hairs, but the consequences of the perspective shift become evident if we think of how a technological aid would have to be different if it were built so that fluency were prioritised over correction. Even though word processors today have a WYSIWYG design with a focus on correcting rather than supporting fluency, the field is aware that a more cognitively motivated design is desirable, as noted by de Smedt (2009, p. 4): “[A] holistic approach to writing is needed, where NLP [Natural Language Processing] research better interacts with the study of cognitive aspects of the

WYSIWYG <i>What you see is what you get</i>	CORRECTION	fluency
WYGIWYN <i>What you get is what you need</i>	correction	FLUENCY

FIGURE 7.1 Schematic presentation of two principles which differ qualitatively in the accentuation of fluency versus correction. Capital letters indicate the main element of the principle.

writing process (including first and second language learning and language disorders)". There is still much work to be done to understand how a writing tool can interact better with the writer.

Why Fluency?

In accordance with theories of flow (Csikszentmihalyi, 1997), we consider fluency as a span, in which there is a threshold that needs to be passed to be fluent. This is different from the idea that fluency equals speed. Fluency has to do with the efficiency of coding and decoding. The threshold can be defined as a point where the thought moves from being delayed or interfered with by various constraints (e.g., on coding and decoding, or confidence) to being pulled ahead by the semantic networks the current words activate. The act of writing, accompanied by formal instruction, typically involves writing down an idea. However, writers soon experience that the words they write lead their thought onto new tracks, often ending up somewhere else than initially planned. Hecht-Nielsen (2007) presents a model called Confabulation Theory, where the activation of knowledge links may create similar stories that are spun out from seed concepts, and this may also explain how false memories are created through activation and co-activation of knowledge links. This kind of progress in terms of words following words, where interpretations continuously are challenged by new meanings, is a common denominator for both reading and writing. Still, the elements that hamper this kind of progress in writing are different from reading. Although a reader manages to decode a word, there is no guarantee the words are already in the reader's lexicon. In writing it is first and foremost *spelling a word that is already in your lexicon* correctly, and then *recognizing* the written words as if they were old friends when you see what you wrote. Therefore, in reading, the problem involves lack of access to the word and concept; in writing, however, it is about knowing the orthographic representation of a word for which the writer already has an active phonological representation

and semantic associations (i.e., the writer knows what he or she wants to write). Given this, we can investigate it experimentally by having individuals write in a blinded condition where the visual feedback is degraded by having the letters of the words crossed out before presented to the screen.

Text linguistics (e.g., Halliday & Hasan, 1976) focuses on how activation of semantic fields contributes to text cohesion. A central insight is that one word activates a whole field of new words which are rooted in a variety of contexts and as parts of different expressions. Activation effects are also investigated using priming with related words or pictures, and effects of both semantic and phonological priming are robust (cf. Levelt, Roelofs, & Meyer, 1999; Levelt 2001, *inter al.*).

Research exploring how words cue or prime meaning has to a large extent been focused on readers or listeners, and less on speakers or writers. In focusing on writers we link these effects to the notion of fluency. In the literature on teaching writing, these aspects are essential in the approach termed *free writing* (Elbow, 1973). According to Elbow, free writing is writing without stopping, editing, worrying about grammar, thinking, or rushing. Free writing is a technique extensively used to overcome writer's block, and appears as an element in many approaches to writing instruction. From our point of view free writing is useful because it allows the writer to follow the emergent properties of words, to follow new paths of meaning, without the heavy burden imposed by the WYSIWYG interface.

Flower & Hayes' (1981), and later Hayes' (2012) model of writing both emphasize that writing takes place in distinct processes, that is, that planning takes place in sentence boundaries, during longer pauses. Further, the execution of what is planned takes place in separate sequences, called bursts. To a large extent, this model is in line with common sense about the writing process, particularly a production model that assumes that we plan first, then translate/transcribe our thoughts into the written mode. Ofstad Oxborough, in her doctoral research, shows that a large part of the planning activity happens in parallel with the writing bursts (Ofstad Oxborough, forthcoming; Torrance, 2015). The competent undergraduate students in her samples wrote coherent text with remarkable fluency, with average pauses of two seconds followed by writing 13 words, on average, before pausing again (and for 25% of two second pauses, a mean of 18 words). Pauses at the start of sentences were, on average, around 2.5 seconds long. Other researchers report similar values (Schilperord, 1996), with only marginally longer times for younger writers (Wengelin, 2002), suggesting that a large part of planning happens concurrently with output rather than during pauses, and that this planning is a largely implicit, low level process (rather than explicit thinking and reasoning). Two or three sec-

onds is not sufficient time to engage in any but the most rudimentary higher-level processing. At minimum, the surface form of the emerging text must be planned incrementally, in parallel with output. Even if 13 or more words can be phonologically encoded in two or three seconds, retaining these during output would be substantially beyond working memory capacity.

These kinds of findings suggest, therefore, that an essential part of meaning construction takes place *during* writing, word by word. In writers whose output is fluent, unhindered by the need to attend to spelling, central long term memory retrieval processes remain available for meaning making (i.e., for activation to spread from what is currently being expressed to associated semantic content). This is, however, a demanding, focussed and precarious process. If retrieval processes are demanded by the need to spell the current word, meaning-related processing immediately grinds to a halt, and must be picked up again once the spelling problem has been resolved (probably with recourse to some deliberate cue-searching, which further disrupts flow). If there are sufficient interruptions to the parallel output / meaning making processing, the writer loses the plot (quite literally, if the task is to write a narrative).

Supporting writing fluency therefore requires maintaining, as much as is possible, the dual meaning making and output processing that occur when writing is flowing. Help with spelling that interrupts writers mid-sentence is likely to achieve correctness at the expense of meaning making. There are, however, points during production where the writer would naturally stop to take breath—when the ongoing flow of the composition would naturally be halted while the writer considers what has already been written, or just has a bit of a rest. This is more likely to occur at sentence boundaries than within sentences (Wengelin, 2006), and much more likely to occur at paragraph boundaries. If a writer support tool must interrupt the writer, these contexts are likely to be much less disruptive points at which to do so.

Preventing Reading during Writing

In a recent study (Torrance, Rønneberg, Johansson, & Uppstad, 2016) we asked upper secondary school writers with dyslexia to write—by keyboard—short argumentative essays under each of two conditions: Blind Writing, in which reading was prevented by displaying Xs on the screen for each letter keystroke (the actual text being stored for analysis), and Normal Writing, in which they could see what they were writing. The writers with dyslexia ($N = 26$) were compared with a group of students without dyslexia ($N = 26$) matched on age and math performance (as a measure of non-verbal ability).

As might be expected, the individuals with dyslexia made substantially and significantly more spelling errors than the students without dyslexia, in both

conditions. The extent to which the spelling for a particular word deviates from correct spelling can be assessed in terms of Levenshtien distances (Levenshtein, 1966), sometimes called *edit distance* as measured by the minimal edit operations on characters necessary to go from one word to the other. Preliminary analysis of the data indicated that for the writers with dyslexia this distance was not larger than for the control group, indicating that the individual errors are not generally more severe, only that there are more of them. This is partly good news in terms of correcting the spelling mistakes, but since there are more word errors there is a challenge for automatic correction procedures: if we use local lexical context for correction, there is an increased probability that the context is degraded. Writers with dyslexia also tended to write fewer words, although this difference was not statistically significant.

The most surprising finding, however, was that in the blind condition both groups managed to produce coherent text despite not being able to read what they had written. Final texts contained more spelling errors, as might be expected given that preventing reading prevents both error monitoring and error correction. However, there were only small differences between conditions in holistic quality, and texts were longer in the blind condition. This increased length can most easily be explained by time saved both in error correction and by the fact that they could not spend the last few minutes of the writing period reading what they had written. Note, though, that preventing this final read-through did not adversely affect quality. The net result of preventing reading, therefore, was to improve productivity, with no substantial detrimental effects to quality. This at minimum suggests the feasibility of a tool helping writers with dyslexia which involves them writing, for some of the time, without being able to monitor what they are producing.

One possible conclusion from the fact that preventing reading did not seem to impair writing in either group is that writers do not tend to perceive, or at least process to any depth, the words that they are currently typing. This is possibly because their attention is directed to something else. This could be very literally the case; if writers look at the keyboard while typing, they will not read the words currently being typed. However, even if students in our study did fixate on the screen when typing (which is anecdotally increasingly the norm for a generation for whom keyboarding is a common, highly practiced activity), it may be that attention was not focussed on the emerging word: writers' focus on what to write next means that they do not monitor what is currently being written. This is probably an adaptive and desirable state of affairs. Once text has been written, there is no particular benefit to revisiting it and flow is partly achieved by attention always being ahead of what is appearing on the screen. In studies exploring writers' eye movements, we have observed, informally, that

some writers tend to habitually fixate slightly above or below the text that they are typing, a markedly different pattern from reading eye movements, where fixation is directly on the word being read. One possible explanation is that writers degrade visual input from the words as they appear on the screen, thus reducing the chances that these words will steal attention that is better focused on what is to come next. The blind condition in our study simply takes this situation one step further.

How to Create a Better Writing Tool

Most of the ideas here are intended as ideas for a general writing tool, and not only for individuals with dyslexia. The idea is to work with feedback to the writer, and help motivate the writer to write more and to rehearse, rather than avoid, difficult words.

So far in this chapter we have identified some of the ways in which writers with dyslexia might struggle, relative to peers without dyslexia and made the following arguments: Aiming to promote fluency rather than complete accuracy will benefit both written products and the writers' learning. "Good enough" is an appropriate fluency-maintaining attitude toward the emerging text, and WYNIWYG is a design principle that can facilitate delayed error-correction. We have shown that preventing revision during writing did not adversely affect quality. By separating revision from transcription, we argue that writers are better able to plan in parallel with writing. Still, writers with dyslexia make many errors that need to be corrected. For a spell checker these errors need to be detected and feedback should be given to the writer at a point where the writer is ready for feedback. In this section, we suggest features that might be present in more effective writer-support tools for writers who struggle with spelling.

Timing of Feedback

The timing of feedback from a writing tool depends on where the writer is in the writing process. Feedback should aim at being nonintrusive; disturbing the writer can distract from other goals in the writing process. The writing tool has information about the typical writing speed and pause pattern, and may use machine learning techniques to find out when it is safer to provide feedback, with different thresholds for positive and negative feedback. Not giving immediate feedback on spelling errors will have positive consequences for the writer as well as the writing tool.

Interviews with individuals with dyslexia (Mossige, Johansson, Rønneberg, Torrance, & Uppstad, in preparation) show that this group is aware of negative

feedback during writing, and this is often reported as being overwhelming. Many also report that they try to correct as soon as they get the red underline for misspelled words. However, this may in fact disrupt the writing process, and should therefore be avoided; in addition, there are cases where the currently-written error is unlikely to be identified, because the material is too far off from known words. These cases may benefit from the tool asking for a second or third try at the word. Such tries are likely to circle around the intended word even if the intended word is still elusive. Extra tries give valuable information about what the writer is more certain about, for example, that some letters should be near the beginning of the word, which is the kind of simple information that would help an automatic correction procedure to restrict the search space.

Positive learning happens when the learner enjoys the activity. Positive feedback is one way to enhance the experience, so it is surprising that there is scarcely any positive feedback in writing tools although negative feedback is abundant. There is little reported research on positive feedback, but teachers who work with students with dyslexia are commonly advised to give positive feedback to encourage the students to write, learn, and not give up. This may also be good advice for computational writing tools, although we find very little research that systematically evaluates outcomes. Most of the writing aids are aimed at helping students correct errors, rather than encouraging them to write and focus on the important tasks of planning, structuring, presenting and arguing in the text. For a hesitant writer it would be reassuring to get feedback that indicated that the written words are good enough (see also Serlin & Lapsley 1985), meaning that the written words are either correctly spelled or can very likely be corrected by a spell-checker (i.e., they are close enough to existing words).

There is very little evidence that individuals with dyslexia benefit from immediate correction. This supports an idea of having them write in stages that separate text production from editing, to wait and edit in a second stage. Thus, when it is time for editing, the mechanism has access to a *history* of how the text was produced. In much of mainstream Natural Language Processing the text (input) appears as if it just materialized. However, as we have argued, the text is produced over time, and has gone through a process where the writer has stopped and paused at various points, and where words have been written and rewritten as the writer has struggled to find the right expression or spelling. This information can be used to aid detection of problems in the product (the text to be edited). It can also be argued that it is a simpler task to provide correction assistance for a text after it has been written, rather than while it is being written, because we have access to more information, such as information about the user behaviour while writing. We would also have access

to the words in the entire text, which could be used to detect more higher-order information (topic, semantic relatedness between words) and to restrict the search space of words to consider when alternatives have to be determined.

Another reason to edit in a second stage is that the text could be read out by text-to-speech software, and the user could be asked to press a key when a problem is detected. This will give a search range for detecting problems. All hearing users will be able to interact with the program in this way, and problems such as confused words will be easier to detect when read aloud than when visually scanning through the text, as even when scanning we typically read for meaning. One common error, familiar to most writers, is that it is hard to detect duplicated unstressed function words (e.g. the the) by scanning or reading, but more easily detected if read aloud by automatic text-to-speech software.

The use of context words is highly likely to positively influence automatic word prediction, which is the useful task of predicting the intended word given textual input from the writer. We will look at some common types of lexical errors later in the section. There are many statistical techniques for finding out which words are correlated in usage. The challenge is to weigh correlations from large general text sources together with the individual preferences of the user. For example, writing the word “dog” will influence, or prime, the use for other words for domestic animals such as “cat”.

Spelling errors can be separated into two categories, non-word errors and confusables (Reynaert 2009; Stehouwer & van Zaanen, 2009). In addition to spelling errors, we need to handle out-of-vocabulary words (OOV) in a graceful way, e.g., through using the internet. Try googling “she bought two books” and “she bought too books”. You find significantly more documents with “two” than “too”; the same technique is useful for OOV.

The advantage of the internet is that popular names (artists, etc.) and new words are available more quickly than in edited word lists and dictionaries. The idea is also to use context words to see how alternative words behave in the given context. Web material can have reasonable coverage for words in context, simply because there are such vast amounts of data. Rømcke & Johansson (2008) solved a classification of named entities task using frequency ranking of web queries such as “his name is x” vs “her name is x” vs “hotel in x” etc., where x is replaced by the item to classify for a large proportion of cases. Our largest corpora cannot compare in size with having the entire web to search, and as mentioned the web will be more up to date on recent events and the associated lexical material.

Non-word errors are easy to detect since they are not in the vocabulary. However, they are not always easy to correct, as this depends on how different the input is from the correct word, i.e., how many simple edit operations are

needed to correct the input: such simple edit operations can be classified as *transpositions, *deletions, *insertions, and *substitutions (cf. Damerou, 1964; Reynaert, 2004).

Confusables are errors that result in existing words, and are often difficult to detect, e.g., homonyms (words with similar sound; e.g., two, too, to). If one imagines utilizing the internet to identify a likely error and constrain the appropriate choice, the same first letter will be particularly important for lexical search. However, these words often occur in very different lexical contexts, as if the homonymy is preserved because the lexical context provides enough information to choose the right word. However, not all confusables are homonyms. We will consider any error that leads to a real word as a confusable (e.g., from, form). Fortunately, confusables can often be detected using lexical context without full parsing (e.g., “She bought too books”, “I think so to”, “Two think is too do”), especially if we can predict the alternative. Kukich (1992) mentions that of all errors in documents 25–50% are confusables.

Compounding

A problem in Scandinavian languages is separating compound words into two words (e.g., “sær skrivning”). Reynaert (2004) introduced the idea of a Zipf-filter to identify divided compounds. One corpus-based solution is to detect that the frequency of the compound in a large corpus (e.g., Norwegian Newspaper Corpus) is significantly higher than the frequency of finding the first and the second word together; this is a signal that we might have an error and the user could be alerted to the possible problem.

Once a possible error has been identified, we need to suggest corrections, i.e., to find the closest attested match. The alternatives are commonly ranked by their Levenshtein distance (cf. Reynaert, 2009). A common OCR error is “m” instead of “in”. Correcting this takes two operations: substitute “n” for “m” and insert an “i”. For practical reasons most spell checkers do not consider Levenshtein distances above 5, since the number of possible candidate words increases significantly with each operation. In Norwegian, there are two written norms that sometimes produce problems for form based correction; for example when looking for the Nynorsk word “forespurnad” (request) and writing the bokmål word “forespørsel” (request) the alternative “fårepølse” (sheep sausage) was suggested, which looks far from both original words but contains the same first letter, the same consonants (f, r, s, p, l) and the same vowels (o/å, ø, e) illustrating that sound coding may indeed find unexpected alternatives. As illustrated here, the task of Norwegian spelling correction might need to include translation between spelling norms. Reynaert (2004, 2009) uses a numeric anagram to code word sequences for look-up, the idea being to repre-

TABLE 7.1 *One or two words? Frequencies and estimated effect size in each cell.*

	One word	Two words
tooth brush	33,100,000 / (+0.80)	4,240,000 / (-0.06)
is there	158,000 / (-0.07)	430,000,000 / (+0.01)

sent the group of letters (or bigrams or trigrams of letters) in the input, and then look up all attested word sequences (n-grams) having the same code. The list of alternatives is then ranked by their Levenshtein distance to find the closest match. It is possible to weight the choices, so that frequent patterns take precedence over less frequent ones, and choices that are in the vocabulary used in the current text could rank higher for that reason.

Reynaert's approach has proven effective at detecting and correcting spelling errors, especially confusables. He gives comforting statistics on error types and edit distances: most errors are corrected in less than three edits. Deletions are the most common error. Context-sensitive matching produces better texts, i.e., more errors are corrected than (re-) introduced by the algorithm. Reynaert (2004) reports a fitness value of 3.89 (for each erroneously corrected error there are 3.89 correct corrections) for Dutch text and 1.49 for English text. The commonly used ISPELL and Microsoft Proofing Tools both have scores under 1, i.e., they make the text worse if you trust their first suggestion.

Frequency counts, especially Google frequency estimates, can be very useful for deciding between alternatives. Johansson (2013) suggests a simple measure for calculating effect size, and this can be generalized to give the effect size contribution in each cell of a contingency table. The effect size is a relevant measure for ranking alternatives on higher or lower frequencies than expected. Each cell is marked positively for cells with a higher observed frequency than expected and vice versa (positive for lower than expected). One example is choosing between compounding or not between two words. Table 7.1 shows an English example, the choice between "tooth brush" or "toothbrush". According to Google frequencies, the choice is quite clear. The cell containing toothbrush in one word is indicated by a positive effect size of 0.8, and the other cells are more or less as expected: frequencies lower than expected would be alternatives to avoid given a better choice. The lexical sequence "is there" is indicated as two words that are occasionally mistakenly written as one word, a fact we use to estimate how often people slip up and fail to type a space between two words.

TABLE 7.2 *Effect of context words. Frequencies and estimated effect size in each cell.*

	Word	+Fruit +Banana
kiwi	71,200,000 / (0)	9,220,000 / (0.21)
wiki	905,000,000 / (0)	6,720,000 / (-0.02)

Table 7.2 shows an example where two real words (“wiki” and “kiwi”) might have been confused. Here the challenge is to decide if “wiki” or “kiwi” is the intended word. If we only have the word without context, we do not have any evidence for either word, although “wiki” is by far the most frequent alternative. Context words, content words that are generally infrequent but are overrepresented in the text at hand, may change preferences. For example if relevant context words (such as “fruit” and “banana”) were detected, the effect size indicates that “kiwi” might be the intended word—it occurs more frequently than expected and with an effect size nearly ten times larger than “wiki” in this context.

Access to frequency count estimations with good coverage is essential for this simple measure, and the frequencies estimated by Google have the advantage of high coverage and the capability of using lexical correlations and other information to estimate useful numbers for co-occurrence of lower frequency words and misspellings. However, frequency estimation is a challenge that can be tackled by many different methods; machine learning techniques combined with very large corpora may provide frequency estimates that do not depend on access to the internet or a search engine.

Gaming

When a writer has written a word incorrectly, what does it mean from a learning perspective? It could mean that this is a word that the writer is likely to use in the near future. That is, it is a word that it is worth putting some extra effort into, apart from just correcting it. This is valuable information for a writing tool with a goal of helping the writer learn to write better. For example, the word can be added to a set of practice words. A program could also award points when difficult (for example, low frequency) words are written correctly, or even close to correctly, and in that list of difficult words would be words that have been written incorrectly previously. The game aspect would have some chance of maintaining interest by encouraging writers to use words they have misspelled before, and thus may tend to avoid in writing.

Conclusion

Why bother with writers? One answer is that writers provide much more information than what one gets from just looking at their final products. We have shown that simple behaviour measures such as keylogging can give some information about where the writer is in the writing process. Long pauses before starting a sentence might be a good time to present feedback. The writer also continuously updates the vocabulary that is used in the document, and an intelligent writing aid would use these words to activate semantically related or frequently co-occurring words. The writer is a great source of information.

The second part of the answer is that an intelligent mechanism could figure out from the writers' activity what they need. Often writers only need a bit of positive feedback (that the text so far is good enough). The aim is to improve the writers' flow of writing by encouraging them to plan forward rather than worry about already-produced text. Not only is this a valuable addition to a text tool, it can make writing an enjoyable learning experience. There are several ways of achieving this, and one that we have discussed is to keep track of the vocabulary and provide positive feedback when previously hard items are produced correctly or closer to the correct word. Making part of this a game activity, where writers are awarded points for correctly writing "difficult" words, also would encourage the writer to spend time writing. Writers with SLD have great potential to learn and improve, and spending more time writing would certainly help, and this is likely if the activity is enjoyable.

Existing technological solutions to aid writing do not consider *how* the text is produced. Since tools target a correct final product, they miss supporting the writing process. We have argued that such goals should be rethought. The writer provides information that can be utilized. This has consequences for *how* and when a writing aid should interact with the writer. The writing flow and experience of the writer should have higher priority.

We have also argued that separating writing and editing may both support text flow and provide a better chance for semi-automatically correcting the product. Correction processes are supported by recent developments in the availability of extremely large collections of text, possibly through the internet, together with the availability of user-provided behaviour data and data on the vocabulary used in the document. An automatic correction procedure does not have to consider all possible words and word forms in the language (or languages), just those likely to occur in the document at hand and the contexts it provides.

This chapter has outlined some solutions within reach for current technology. Increasingly large amounts of text available via the internet can be used for word prediction in context, to determine co-occurrence patterns, and to activate words often found together. The writing process can thus be supported by building up expectations tailored to the needs of the writer.

A shift in design philosophy is relevant because it will allow the writer to focus more on producing text, learning to produce better texts, and becoming a better editor rather than worrying about often minor mistakes, which often can be more or less automatically corrected by a context sensitive editing mechanism.

Producing a readable text is hard; making a text readable, even enjoyable, for another human is a process that involves many steps of editing and reviewing. In addition, it is often necessary to have other people read and comment on the text; we are not arguing that we can eliminate readers from the writing process. However, when the writer enjoys the activity and seeks more of it, they are also more likely to adapt to their readers and with time and practise to produce even more readable texts.

References

- Alexander-Passe, N. (2006). How dyslexic teenagers cope: An investigation of self-esteem, coping and depression. *Dyslexia*, 12, 256–275.
- Bangert-Drowns, R.L. (1993). The word processor as an instructional tool: A meta-analysis of word processing in writing instruction. *Review of Educational Research*, 63(1), 69–93.
- Berninger, V., Abbott, R., Thomson, J., & Raskind, W. (2001). Language phenotype for reading and writing disability: A family approach. *Scientific Studies in Reading*, 5, 59–105.
- Berninger, V.W., Nielsen, K.H., Abbott, R.D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, 46(1), 1–21.
- van den Bosch, A. (2005). Scalable classification-based word prediction and confusable correction. *Traitement Automatiques des Langues*, 46(2), 39–63.
- van den Bosch, A. (2006, June). All-words prediction as the ultimate confusable disambiguation. In *Proceedings of the Workshop on Computationally Hard Problems and Joint Inference in Speech and Language Processing* (pp. 25–32). New York: Association for Computational Linguistics.
- van den Bosch, A. (2011). Effects of context and recency in scaled word completion. *Computational Linguistics in the Netherlands Journal*, 1, 79–94.

- Bruck, M. (1993). Component spelling skills of college students with childhood diagnoses of dyslexia. *Learning Disability Quarterly*, 16(3), 171–184.
- Coleman, C., Gregg, N., McLain, L., & Bellair, L.W. (2009). A comparison of spelling performance across young adults with and without dyslexia. *Assessment for Effective Intervention*, 34(2), 94–105.
- Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. New York: Basic Books.
- Damerou, F.J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 7(3), 171–176.
- Elbow, P. (1973). *Writing without teachers*. London: Oxford University Press.
- Flower, L., & Hayes, J.R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
- Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *The Journal of Technology, Learning and Assessment*, 2(1).
- Graham, S., & Perin, D. (2007). Effective strategies to improve writing of adolescents in middle and high schools. *A Report to Carnegie Corporation of New York*. New York: Alliance for Excellent Education.
- Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman
- Hayes, J.R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388.
- Hecht-Nielsen R. (2007). *Confabulation theory*. Springer-Verlag: Heidelberg.
- Johansson, C. (2013). Hunting for significance. In L. Hareide, C. Johansson, & M. Oakes (Eds.), *The many facets of corpus linguistics in Bergen—in honour of Knut Hofland*, in series *Bergen Language and Linguistics Studies*, Vol. 3(1) (pp. 211–220). Bergen, Norway: University of Bergen.
- Kukich, K. (1992). Techniques for automatically correcting words in text. *ACM Computing Surveys*, 24(4), 377–439.
- Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production, *Behavioral and Brain Sciences*, 22, 1–75.
- Levelt, W. (2001). Spoken word production: A theory of lexical access. *PNAS*, 98(23), 13464–13471.
- Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics—Doklady*, 10, 707–710.
- Lewis, R.B., Graves, A.W., Ashton, T.M., & Kieley, C.L. (1998). Word processing tools for students with learning disabilities: A comparison of strategies to increase text entry speed. *Learning Disabilities Research and Practice*, 13(2), 95–108.
- MacArthur, C.A., Graham, S., Haynes, J.A., & De La Paz, S. (1996). Spelling checkers and students with learning disabilities: Performance comparisons and impact on spelling. *Journal of Special Education*, 30, 35–57

- MacArthur, C. (2006). Assistive technology for writing: Tools for struggling writers. In L. Van Waes, M. Leiten, & C. Neuwirth (Vol. Eds.) & G. Rijlaarsdam (Series Ed.), *Studies in Writing*, Vol. 17. *Writing and Digital Media* (pp. 11–20). Oxford: Elsevier.
- Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: a meta-analysis of research findings. *Reading and Writing*, 25(3), 641–678.
- Mossige, M., Rønneberg, V., Johansson, C., Uppstad, P.H., & Torrance, M. (n.d.). How should we support the writing of students with dyslexia in upper secondary school: Interviews and perspectives.
- Ofstad Oxborough, G.H. (n.d.). *The Visual Writer* (unpublished doctoral thesis). University of Stavanger.
- Reynaert, M., (2004). Text induced spelling correction, In *Proceedings COLING 2004*. Geneva.
- Reynaert, M. (2009). Parallel identification of the spelling variants in corpora. In *Proceedings of The Third Workshop on Analytics for Noisy Unstructured Text Data 2009* (AND-2009) (pp. 77–84). New York: ACM.
- Schilperoord, J. (1996). *It's about time, temporal aspects of cognitive processes in text production*. Amsterdam: Rodopi.
- Serlin, R.C., & Lapsley, D.K. (1985). Rationality in psychological research: The good-enough principle. *American Psychologist*, 40, 73–83.
- de Smedt, K. (2009). NLP for writing: What has changed? In R. Domeij, S. Johansson Kokkinakis, O. Knutsson, & S. Sofkova Hashemi, (Eds.), *Proceedings of the Workshop on NLP for Reading and Writing—Resources, Algorithms and Tools (SLTC 2008)*. *NEALT Proceedings Series*, Vol. 3. (pp. 1–11).
- Stehouwer, H., & van Zaanen, M. (2009). Language models for contextual error detection and correction. In *Proceedings of the EACL 2009 Workshop on Computational Linguistic Aspects of Grammatical Inference*. (pp. 41–48). New York: Association for Computational Linguistics.
- Tops, W., Callens, M., van Cauwenberghe, E., Adriaens, J., & Brysbaert, M. (2013). Beyond Spelling: the writing skills of students with dyslexia in higher education. *Reading & Writing*, 26(5), 705–720.
- Torrance, M. (2015). Understanding Planning in Text Production. In C.A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd Edition) (pp. 1682–1690). New York: Guilford Press. <http://doi.org/10.1007/s13398-014-0173-7>
- Torrance, M., Rønneberg, V., Johansson, C., & Uppstad, P.H. (2016). Adolescent weak decoders writing in a shallow orthography: Process and product. *Scientific Studies of Reading*, 20(5), 375–388.
- Wengelin, Å., & Strömquist, S. (2000). Discourse level writing in dyslexics—methods, results, and implications for diagnosis. *Logopedics Phoniatrics Vocology*, 25(1), 22–28. doi: 10.1080/140154300750045876

- Wengelin, Å. (2002). Text production in adults with reading and writing difficulties. *Gothenburg Monographs of Linguistics*, 20. Gothenburg: Gothenburg University.
- Wengelin, Å. (2006). Examining pauses in writing: Theory, methods and empirical data. In G. Rijlaarsdam (Series Ed.) and K.P.H. Sullivan & E. Lindgren (Vol. Eds.), *Studies in writing*, Vol. 18, *Computer key-stroke logging and writing: methods and applications (Studies in Writing)* (Vol. 18, pp. 107–130). Oxford: Elsevier.
- Wengelin, Å. (2007). The word level focus in text production by adults with reading and writing difficulties. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 67–82). Amsterdam: Elsevier.