



## OPEN ACCESS

## EDITED BY

Luuk Van Waes,  
University of Antwerp, Belgium

## REVIEWED BY

Thomas Quinlan,  
Independent Researcher, Princeton,  
United States  
Scott Beers,  
Seattle Pacific University, United States

## \*CORRESPONDENCE

Sanna Kraft  
✉ [sanna.kraft@svenska.gu.se](mailto:sanna.kraft@svenska.gu.se)

## SPECIALTY SECTION

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Education

RECEIVED 29 December 2022

ACCEPTED 08 March 2023

PUBLISHED 23 March 2023

## CITATION

Kraft S (2023) Revisions in written composition:  
Introducing speech-to-text to children with  
reading and writing difficulties.  
*Front. Educ.* 8:1133930.  
doi: 10.3389/feduc.2023.1133930

## COPYRIGHT

© 2023 Kraft. This is an open-access article  
distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Revisions in written composition: Introducing speech-to-text to children with reading and writing difficulties

Sanna Kraft\*

Department of Swedish, Multilingualism, Language Technology, Gothenburg University, Gothenburg, Sweden

The ability to perform revisions targeting the content of the text is important for text quality improvement, and it is hypothesized that lower-level transcription processes need to be automatized in order to free up capacity for higher-level processes such as revision. However, for people with reading and writing difficulties due to underlying difficulties with decoding and spelling, the transcription process is rarely automatized because of their troubles with spelling. One possible way to circumvent spelling difficulties, and possibly gaining capacity for higher level processes such as revision, is to write using speech-to-text (STT). This study investigates the revisions performed when children with reading and writing difficulties ( $n = 16$ ), and a reference group without such difficulties ( $n = 12$ ), compose text using STT and using a keyboard. More specifically, the study investigates whether, and if so how, revisions at various levels, errors left in the final text product, and text quality differ between conditions and between groups. The compositions were logged using keystroke logging (keyboard) and audio- and screen-recording (STT). The level of revisions were manually coded. The results showed that children with reading and writing difficulties gain more from composing with STT compared to keyboard than the reference group. They leave fewer errors in their final text product when composing by means of STT, even though they need to engage more in the correction of surface errors because of the large number of STT errors. Despite the numerous STT errors, neither the proportion of meaning-related revisions nor text quality decreased in composing with STT (for either of the groups). Taken together, the results suggest, albeit not emphatically, that STT may be appropriate as a facilitatory tool for children with reading and writing difficulties. However, more research is needed to investigate instruction that addresses strategies for STT transcription and highlights the shortcomings of the tool in the target language, and also focuses specifically on higher-level aspects of composition such as planning or revising, in order to gain further knowledge about the feasibility of using STT as a means of composition for children who struggle with writing, and its possible effects over time.

## KEYWORDS

reading and writing difficulties, writing process, speech-to-text, children, keystroke logging, revisions, text quality, assistive technology

## 1. Introduction

From a cognitive point of view, there is general agreement that three main processes interact recursively during composition: planning, formulation and revision (Alamargot and Fayol, 2009). Revision involves the interaction of the subprocesses of evaluation and revision (Flower and Hayes, 1981), and it is considered to be a demanding high-level process (De La Paz et al., 1998) that emerges later than both transcription and pre-planning (Berninger and Swanson, 1994). See also Torrance et al. (2007) for a discussion.

Revisions can be made to improve a text or to correct errors in it, and they are important for the writing process as well as its output (Conijn et al., 2022). According to Fitzgerald (1987), revisions can be made at any point during the writing process and can include any type of changes—and they can also be made in the writer's mind before the text is written. Revisions made during writing can thus include lower-level corrections, such as the editing of spelling errors or the paraphrasing of an expression, as well as higher-level revisions of the content of the text that affect the substance of what is being said. The present study investigates revisions at different levels made during composition using speech-to-text by children with reading and writing difficulties.

The ability to perform revisions targeting the content of the text is important for text improvement, and it has been hypothesized that transcription processes need to be automatized in order to free up capacity for higher-level processes such as evaluation and revision (McCutchen, 2000). However, for people with reading and writing difficulties due to underlying difficulties with decoding and spelling, the transcription process is rarely automatized. In fact, concerns about spelling often persist until the university level for members of this group (Sumner and Connelly, 2020), who make more spelling revisions than writers without such difficulties (Wengelin, 2007; Sumner and Connelly, 2020). When the spelling process is demanding, the transcription process (involving the interaction of spelling and handwriting/typing) places a great load on overall cognitive capacity, causing text quality to suffer (Sumner et al., 2013). Facilitating the spelling process so as to free up capacity for higher-level processes is of great importance when it comes to this group.

One possible way to circumvent spelling difficulties is to write using speech-to-text (STT). This has been claimed to free up cognitive capacity by allowing children to write without having to focus on spelling (MacArthur, 2009). However, even though an STT tool does not make any spelling mistakes as such, it may “mishear” words and thus make semantic errors (referred to below as “STT errors”) instead. In a worst-case scenario, those resources that were freed up in theory could still end up being used for lower-level processes, namely to proofread and to correct words, and the writer might end up being forced to spell some words herself anyway.

There is thus clearly a need to investigate how revision processes (pertaining both to content/meaning-related revisions and to surface revisions such as lower-level revisions of STT errors) are affected when children with (as well as without) reading and writing difficulties compose using STT. To my knowledge, this has not yet been investigated.

In this study, I address this knowledge gap by exploring the revisions performed when children with reading and writing difficulties, and a reference group without such difficulties, revise text during composition using STT and using a keyboard, respectively. More specifically, this study investigates whether, and if so how, revisions at various levels, errors left in the final text product, and text quality differ between conditions and between groups. Further, as noted above, the STT tool sometimes fails to transcribe the words intended by the writer, meaning that the writer may have to type (and thus spell) words herself. Hence concerns about spelling may arise even during composition using STT. For this reason, instances of spelling management due to inaccuracies caused by the STT tool are explored qualitatively.

### 1.1. Revisions tax working memory

Human cognitive capacity is limited, and all revisions<sup>1</sup> will tax working memory. When lower-level processes such as spelling or typing are not yet automatized, they will demand resources from working memory, leaving less capacity to be directed toward higher-level processes such as evaluation and revision (McCutchen, 1996). The latter types of revisions, which affect the substantial content of the text, can, but do not necessarily, improve quality (De La Paz et al., 1998).

The process of revision is not a distinct process, but rather the interaction of several processes during composition (Hayes and Berninger, 2014). In fact, if the thoughts and ideas that the writer planned (internally) to express are not fully reflected in the linguistic formulation achieved (which is assessed by means of reading and evaluation), there is a need to make a diagnosis of the problem (by means of more reading and evaluation), whereupon, if the writer so chooses, a schema-oriented solution is implemented, meaning that the revision process itself includes planning, evaluation, and goal-orientation. Hence the revision process is a meta-cognitive activity where the writer has to consider not only the text written so far but also the forthcoming, emerging text. When other processes, such as transcription, are not yet automatized, there is a risk that the overall plan kept in working memory will be interrupted (McCutchen, 1996). Further, the ability to successfully revise text is dependent, in turn, on a large number of related processes, for example, reading (Alamargot et al., 2006; Wengelin et al., 2010), executive functions (De La Paz et al., 1998) and language ability (Chenoweth and Hayes, 2001). The development of this ability requires exercise and the learning of strategies (Scardamalia and Bereiter, 1987; De La Paz et al., 1998).

<sup>1</sup> Note that I use the term *revision* for all kinds of editing, regardless of its character. In previous research, a distinction has sometimes been made between editing and revising, where editing is a lower-level skill or activity and revising is a higher-level one. In the present paper, however, I instead refer specifically to different levels of revisions in order to make a corresponding distinction.

## 1.2. Development of revising skill

Revising behavior changes with development. In general, developing writers engage overall in less revision than more skilled writers (Torrance et al., 2007). Further, younger children's revising is mostly directed toward surface editing, such as mechanical changes and word substitutions (Plumb et al., 1994; Chanquoy, 2001).

It has also been suggested that revisions enable changes to thoughts or ideas (McCutchen et al., 1997), and revisions might indeed be made to comply with new ideas that may possibly have been created or generated by the process of writing itself (Galbraith, 2009). This is why revisions that target content and ideas have the capacity to improve the quality of the text. With development, surface revisions tend to decrease while meaning-related revisions grow more common. Limpo et al. (2014) reported that, for handwritten texts, the use of meaning-changing revisions predicted the quality of a text in children in grades 7–9 but not in children in grades 4–6. However, despite the obviously great importance of revising skill, we still lack knowledge about how revisions at different levels manifest themselves when children with reading and writing difficulties compose in general, and by means of STT in particular (but see Sumner and Connelly, 2020, for revisions in university students with reading and writing difficulties when composing by hand).

The main goal for teachers in supporting children's writing is to help them acquire strategies to improve their texts. Hence helping children become skilled, goal-oriented editors of content is of great importance. For children with reading and writing difficulties, the inadequacy of lower-level skills of decoding and spelling reduces the ability to direct attention toward the content and goal of the text being written. Such children tend to be local in their writing process, directing a large amount of their capacity toward spelling revisions (Wengelin, 2002, 2007; Sumner and Connelly, 2020). One way to circumvent spelling difficulties, and hence to reduce surface revisions, is to use STT as a means of composition, since this might circumvent the low-level transcription process for spelling and possibly free up more cognitive capacity for higher-level processes such as evaluating the text written so far (possibly followed by revision if any errors are detected) and text improvement. Thus, there is clearly a need for research about the effect of STT use on children's overall writing processes, including revisions.

## 1.3. Investigating revisions

Revisions in writing have been researched in a wide range of settings, such as by means of think-aloud protocols (see, e.g., Chenoweth and Hayes, 2001) or by manipulating text that participants are made to interact with (see, e.g., Limpo et al., 2014). Another possibility is to investigate revisions as they emerge during functional text composition, which can be done if the revisions are traced using keylogging software such as Scriptlog (Frid et al., 2014) or Inputlog (Leijten and Van Waes, 2013). This latter method was used for keyboard processes in the present study. Unfortunately, there is at present no keystroke-logging software capable of tracing STT data from built-in tools and automatically generating revision

output<sup>2</sup>, which is why manual tracing had to be performed for the STT data.

How revisions are analyzed can also differ. For example, the approach may be based on where in the text process they are performed, on the size of the text segments they involve, or on the level concerned by revisions, where a distinction may be made, for example, between surface edits and meaning-changing revisions. The latter approach is the one used in the present study. Faigley and Witte (1981) constructed a taxonomy for these kinds of revisions which has been used to identify differences in revising behavior between skilled and less skilled writers (Faigley and Witte, 1981). Recently, a tagset combining process and product measures in writing has been developed (Conijn et al., 2022), adapting the taxonomy proposed by Faigley and Witte (1981). In the present study, the taxonomy from that tagset was used for annotating revision levels.

## 1.4. Previous research on composing by means of STT

Previous research has shown promising, but somewhat diverse, results regarding composing by means of STT for people with various writing difficulties. For example, Quinlan (2004) showed that children aged 11–14 years with writing difficulties wrote longer texts with STT than when writing by hand, but with no gain in text quality. In somewhat older children (secondary-schoolers) with learning difficulties, MacArthur and Cavalier (2004) investigated writing in three conditions: writing by hand, dictating to a scribe, and dictation with STT. They found that dictation, especially to a scribe, improved text quality. As regards post-secondary students with a learning disability, Higgins and Raskind (1995) showed that they wrote texts with a higher proportion of long words; what is more, the study participants pointed out that they did not have to substitute words that were hard to spell when dictating. However, all of these studies were conducted almost two decades ago, and speech-recognition technologies have improved considerably since then (Lu et al., 2020). A more recent study (Haug and Klein, 2018) showed that, for children aged 10–11 years with no relevant difficulties, composing by means of STT was as good as keyboarding for strategy instruction, because both conditions showed similar gains in argumentation, text length, and text quality. Kraft et al. (2019) showed that composing by means of STT could have the potential to be beneficial for children aged 10–13 years with spelling difficulties, since their (Swedish-language) texts produced by means of STT contained significantly fewer spelling errors. The same authors also reported that the texts produced by means of STT or a keyboard were similar in lexical diversity, lexical density, and text length, whereas both of those conditions differed from spoken production in terms of lexical density and the proportion of long words. This was interpreted as suggesting that the children were able to adapt to written-language norms when composing by means of STT. In other research, more qualitative approaches have highlighted the

<sup>2</sup> It should be noted that Inputlog has the possibility to log speech input from Dragon Naturally speaking and combine it with keystroke logging.

need to consider personal preferences and technical challenges when implementing STT as a writing tool for children (Ok et al., 2022). In addition, previous research into the use of STT in professional writers (without reading and writing difficulties) with and without dictation experience, has showed that writers use different adaptation strategies when they interact with the STT tool, and that the writing mode itself can influence the organization of writing processes (Leijten and Van Waes, 2005). Given the increased availability of STT and the improvement of STT tools, there is clearly a need to investigate its potential usefulness for children with reading and writing difficulties during composition, including revision, and to explore its impact on text-product characteristics, such as the errors left in the final text product as well as text quality.

### 1.5. Opportunities (and possible obstacles) when composing with STT

As mentioned earlier, composing by means of STT could potentially circumvent the spelling process for children with reading and writing difficulties, possibly facilitating transcription and reducing the need for (spelling-related) surface revisions. This is because spelling is not a problem for an STT tool, although it may misinterpret its input owing to homophones or other similar-sounding words, especially when words are dictated one by one and so lack context. It should be pointed out that most previous studies on the usefulness of STT as a writing tool have concerned English-language writing and hence English orthography. There are two reasons why this matters. First, English is a widely used language and the vast amount of available data enables the development of better STT tools. Second, different orthographies are associated with different spelling difficulties. Reading accuracy and speed tend to develop earlier in orthographies with strong grapheme-phoneme correspondence, such as Turkish and Finnish, than in “deep” orthographies that are more dependent on orthographic knowledge, such as English (Aro, 2006). Swedish—the language of study here—is somewhere between those extremes (Seymour et al., 2003). In fact, the same phenomenon can be said to apply within languages in that words with a stronger phoneme-grapheme correspondence are easier (quicker) to write than words that are more dependent on morphological, syntactic, or orthographic knowledge (Delattre et al., 2006).

Factors such as word length and exposure (and, relatedly, word frequency) also influence the level of difficulty in spelling a word. As regards length, children with reading and writing difficulties often have difficulties within the phonological-processing system in terms of problems processing long words and problems placing phonemes in correct order, both in speech and in writing. As regards exposure/frequency, highly frequent words are easier to spell, which could be explained by a statistical-learning effect (Treiman, 2018). Hence words that are both long and infrequent are particularly hard to spell. Here it is interesting to note that Kraft et al. (2022) found that an STT tool (used for Swedish) was better at transcribing long words than short ones when they were dictated one by one, as typically happens after a misinterpretation by the

STT tool. Hence the STT tool was better able to provide the correct spelling for long words than for short ones in that context.

As noted above, aside from the exception involving homophones, orthographic knowledge is unproblematic for an STT tool, which is why it might be extra useful when writers are composing long words that are especially hard to spell, because of knowledge that goes beyond phoneme-to-grapheme conversion. This, in turn, suggests that composing by means of STT could be even more facilitatory in languages with “deep” orthographies, such as English. To more fully understand the effects of STT on text production across languages, there is thus a great need for more research.

Further, it seems reasonable to assume that various kinds of revising behavior are likely to differ between writing conditions and between groups with and without reading and writing difficulties (and, as previous research into adult writers has shown—also between writers with different dictating experience, see Leijten and Van Waes, 2005). First, revisions of spelling might differ: when keyboarding, writers need to spell all the words themselves, meaning that the success of spelling revisions depends mainly on spelling ability. By contrast, when STT is used, some of the words that are otherwise hard to spell will be correct. However, if and when the tool misinterprets its input, writers will have to correct either semantic STT errors or spelling errors produced by themselves—if the tool persistently transcribes its input incorrectly and the writer therefore must spell certain words using the keyboard. Second, revisions of meaning might differ both between conditions and between groups. In particular, if the STT tool succeeds at freeing up more cognitive capacity by reducing the amount of energy needed for spelling, the children with spelling difficulties may end up engaging more in meaning-level revision.

Previous research has shown that the orchestration of cognitive processes differs across conditions. For example, the greater ease of performing revisions in the text written so far when keyboarding than when handwriting causes revisions to be more common in the former condition (MacArthur and Graham, 1987). However, we do not know whether, and if so how, revisions differ between composing by means of STT and other writing conditions such as keyboarding.

Since composing by means of STT is based on speech, the text-production process most likely differs from that for composing by hand or by keyboard. It has been proposed that, when translation processes are fluent, the load on working memory in writing will be reduced (McCutchen et al., 1994). Further, language skill has been shown to be important when it comes to reducing the need for internal revisions of linguistic formulations prior to transcribing them externally, when composing by hand (Chenoweth and Hayes, 2001). It is possible that, at least for some writers, composition by means of STT will create a need for revisions of wording/phrasing, namely if the translation of language from the writer to the STT tool is disfluent. On the other hand, if this process *is* fluent, there would be no need to revise an idea that has been packaged in a linguistically clear manner and is in line with the internally proposed idea (also because that idea has then been expressed without the need to spell words). In practice, however, this may be difficult to achieve when composing by means of STT, especially if the sentence is long and exists of a complex structure that normally



is not used in speech. For this reason, it could be the case that revisions of wording/phrasing will be more common, or at least be of a different character, in STT composing than in keyboarding.

Finally, as mentioned earlier, the STT tool sometimes misinterprets its spoken input. Composing with STT involves an interaction between the technical aspects of the tool and the behavior of the writer during composing, including, for example, burst length and accuracy (Kraft et al., 2022). Better interaction between the writer and the STT tool will reduce the frequency of misinterpretation and hence the need for surface revisions of STT errors. Detecting errors requires reading (Alamargot et al., 2006; Hayes, 2012), and it is possible that the STT condition will encourage writers to look at their emerging text to a large extent, since there will be no need to look anywhere else (such as looking for the correct key when composing by keyboard). Against that background, the present study investigates how STT errors are monitored when children compose by means of STT as well as exploring whether any STT errors are left in the final text. Further, since there is a risk that some of the tool's shortcomings will force writers to spell words by themselves, spelling errors may occur even in the case of composing by means of STT. For this reason, spelling revisions and spelling errors left in the final text are summarized and investigated in this condition as well as in the keyboarding condition. In fact, in order to be able to gain insight into the feasibility of different writing conditions, such as STT, it may be useful to compare revisions of spelling and spelling errors left in the final text from the keyboarding condition with revisions of STT errors and STT errors left in the final text from the STT condition.

To sum up, revising behavior most probably differs between languages of different orthographic opacity, between composing condition, and between children with and without reading and writing difficulties. What I investigate here are revisions performed by children with and without such difficulties who compose in Swedish in two different conditions—STT and keyboard. Since revising skill is of great importance for text-quality improvement, there is a need to investigate whether composing by means of STT affects revising behavior at different levels (such as meaning-changing revisions vs. surface revisions of spelling, wording, and STT errors) as well as measures of text quality. There is an overall lack of knowledge about STT composing processes for children; to my knowledge, there is no research at all into the revision processes seen when children compose using STT (but see Leijten, 2007, regarding STT writing processes in adults). To address this gap, I investigate revision processes, errors left in the final text, and text quality in children composing text by means of STT and on a keyboard, respectively.

## 1.6. Purpose

The overarching purpose of this study is to contribute insights into the feasibility of using STT as a facilitatory writing tool for children with reading and writing difficulties. This is operationalized by investigating whether there are differences in the characteristics of the composition process and the final text products for children with and without reading and writing difficulties composing texts using STT, which is a new condition

to them, and keyboarding, respectively. Regarding the composition process, I quantitatively investigate whether there are differences between the conditions and the groups with respect to revisions at various levels. A further aim is to qualitatively identify obstacles to spelling revision that may emerge (even) during composition with STT (due to STT errors), and to describe how the children deal with these obstacles. Regarding final text products, I investigate whether there are differences between the conditions and groups in terms of errors (spelling errors and STT errors) left in the text and in terms of overall text quality.

### 1.6.1. Research questions

1. What are the main differences and similarities in terms of revisions at different levels (surface and meaning-related revisions) between STT and keyboard composing, and do those differences and similarities differ between the groups of children with and without reading and writing difficulties?
2. What revising difficulties (related to spelling and STT errors) does the STT condition yield, and how are those difficulties dealt with?
3. Does the final text product differ by writing condition and group in terms of errors left and in terms of text quality?

## 2. Method

### 2.1. Participants

The participants ( $n = 28$ ), aged 10–13 years, were recruited and divided into groups (with and without reading and writing difficulties) by special educators and classroom teachers from seven schools in southwest Sweden. Group belonging was confirmed by using the threshold of stanine 3 or below on a standardized test of spelling or percentile 22 or below on a standardized test of decoding words and nonwords. All but three participants remained in their initial group. This resulted in one group with reading and writing difficulties ( $n = 16$ ), referred to below as *Spell*, and one group without such difficulties ( $n = 12$ ), referred to below as *Ref*. Table 1 shows descriptive statistics for the two groups. All composition tasks and the assessment of background measures were administered by the author of this study and took place individually at each participant's school.

The material used in the present study has been retrieved from a broader set of data collected as part of a research project on reading and writing difficulties and text production with speech recognition. The subset in question includes all participants for whom there were complete composition-process data. The participants performed several tasks, but only those relevant to this study are reported here. The study has received ethical approval from the Swedish Ethical Review Authority (Ref. No. 702–17). Written assent/consent was collected from the participants and their caregivers. The participants were informed that they could end their participation at any time without giving a reason.

TABLE 1 Age and scores on measures of individual abilities by group.

Background measure	Assessment	Description of assessment	Spell ( <i>n</i> = 16)		Ref ( <i>n</i> = 12)	
			M (sd)	Min–max	M (sd)	Min–max
Age			11.44 (0.95)	10.00–13.00	11.42 (0.73)	10.42–12.83
Spelling	DLS 4–6 <sup>a</sup>	Spelling words from sentences read aloud	-1.89 (1.01)	-3.89– -0.39	0.16 (0.61)	-0.64–1.24
Decoding words	LäSt <sup>b</sup>	Decoding real words from a word list	-1.17 (1.14)	-3.15–0.39	0.13 (1.00)	-1.45–1.39
Decoding nonwords	LäSt <sup>b</sup>	Decoding nonwords from a word list	-1.07 (0.69)	-2.61– -0.10	0.10 (1.20)	-2.40–2.05

<sup>a</sup>DLS 4–6 (Järpsten and Taube, 2010), <sup>b</sup>LäSt (Elwér et al., 2011).

For ease of comparison, the raw scores have been converted into z-scores ( $M = 0$ ,  $SD = 1$ ). Means, standard deviations, and min-max values.

## 2.2. Data collection

### 2.2.1. Text composition processes

The participants wrote in MS Word using an Apple computer with the built-in STT system and with a keyboard. The spell-checker was turned off. The participants received individual instructions on how to compose with STT: they were presented to a 5-min film clip on how to use the STT tool for composing and editing, whereupon they had 10–15 min training to practice the tool. No participant had prior experience of composing text with STT, but some of them had used STT to search on the Internet.

The participants wrote expository texts and the texts were elicited by a short, silent film clip presenting a moral dilemma (Berman and Verhoeven, 2002). The participants were asked to reason about what a superhero would do if he or she saw what happened. Two different moral dilemmas (stealing and cheating) were used. Order and topic were counterbalanced. Participants' compositions by means of STT were audio- and screen-recorded using Camtasia (Techsmith, 2018), to enable analysis of the composition process. The compositions by keyboard were logged using the Scriptlog keyboard-logging software (Frid et al., 2014). The revisions in the keyboard data were extracted by exporting the Scriptlog data (.idfx-file) to the Inputlog keyboard-logging software (Leijten and Van Waes, 2013). In Inputlog, all revisions from each composition were automatically extracted, and the level of revisions was manually coded; see Section 2.2.2. For the STT compositions, all compositions were exported to ELAN (2019), where the revisions were manually annotated, using the same taxonomy as for the keyboard data. The composition and revision data were then exported from ELAN to R Studio (R Core Team, 2019) for further analysis.

### 2.2.2. Definition and annotation of revisions

For the present study, I used the *orientation* category from the process- and product-oriented tagset of Conijn et al. (2022) when annotating the level of revisions. That tagset is a further development of the widely used taxonomy proposed by Faigley and Witte (1981). The orientation category enables categorization of the levels of revisions, and it distinguishes two main categories: surface changes (referred to below as *surface revisions*) and

semantic changes (referred to below as *meaning-related revisions*). The revisions can be substitutions, reorganizations, additions, or deletions, and they can be made below word level, or on the word, phrase, clause, or paragraph level. See Table 2 for annotation examples. In accordance with Conijn et al. (2022), surface revisions were defined as formal changes involving typos, spelling<sup>3</sup>, grammar, punctuation, capitalization, presentation, and no change (that is, deleting and then adding the same formulation), as well as meaning-preserving changes involving wording and phrasing (referred to below as *wording*). Because of the use of STT, a subcategory was added to *surface revisions*: *STT-error revisions*, that is, revisions of misinterpretations produced by the STT tool (cf. Leijten et al., 2010), such as when a participant dictated *stjåla* “to steal” and the tool transcribed *valla* “to herd.” In the present study, *surface revisions* were first analyzed with regard to the main category of surface revisions—including all of the above-mentioned subcategories—and then further analyzed into the subcategories of *wording revisions*, *spelling revisions* and *STT-error revisions*.

The other main category, that of *meaning-related revisions*, involves micro-structure changes (such as adding or removing supporting information, changing emphasis, understatement, coherence, or cohesiveness) as well as macro-structure changes (such as changing the overall aim or adding or removing subtopics). Since the meaning-related revisions in the present study were few, they were not further divided into subcategories.

What is considered a revision that does, or does not, change the meaning of a text can differ between raters and contexts, and presumably also between genres. Because of this, it is hard to determine whether a revision is a revision of wording or a revision of meaning. For this reason, an inter-rater reliability test was carried out. Twenty percent of the revisions that had been coded as revisions of either meaning or wording were re-coded by an independent rater according to the criteria from the tagset (Conijn et al., 2022). It turned out that the two raters agreed in 87.7% of the cases; the independent rater consistently coded the revisions as wording rather than meaning compared to the first rater. The agreement, measured by Krippendorff's alpha was  $\alpha =$

<sup>3</sup> Note that spelling errors can also occur even when composing using STT, since the writer sometimes corrects errors produced by the STT tool through typing, which in some cases results in spelling errors.

TABLE 2 Annotation examples of surface and meaning-related revisions, and their subcategories.

	Initial composition	Revision	Result	Description
<b>Surface</b>				
Spelling	råd (space and pause)	Deletes space, writes d	rådd “afraid”	Incorrect>correct
	super hjälte	Deletes space	superhjälte “superhero”	Incorrect>correct
	jelte	Adds h and substitutes e	hjälte “hero”	Incorrect>correct
	hene	henne	henne “her”	Incorrect>correct
	sett	Substitutes e>ä>e	sett “seen”	Correct>incorrect>correct
	superhjälte	Adding a space	super hjälte “super hero”	Correct>incorrect
Wording	beholla “keep”	Substitutes word	ta “take”	No meaning change
	prov “test”	Substitutes word	test “test”	No meaning change
	hade han sagt till	Substitutes phrase	skulle han sagt till	No meaning change
STT error	“he would have told”		“he would have told”	
	valla “herd”	Redictates intended word	stjåla “steal”	Correction STT error
	och inte skriva positiva	Redictates intended phrase	skriva en fusklapp	Correction STT error
	“and not write positive”		“write a cheat sheet”	
<b>Meaning</b>				
Meaning	saker “things”	Substitutes word	problem “problems”	Changes meaning
	med två kvinnor	Deletes phrase	om en kvinna och en flicka	Changes meaning
	“with two women”		“about a woman and a girl”	
	och då svarar en kille	Deletes phrase	killen tittar på superhjälten	Changes meaning
	“and then a guy answers”		“the guy is watching the superhero”	

0.54. This relatively low score can be compared to Conijn et al. (2022), where the (trained) raters reached  $\alpha = 0.59$ . The inter-rater reliability will be discussed.

For reasons of transparency, some key aspects of the approach taken to annotation will be described in the following. When an edit was made below word level, the revision was classified as a wording revision, not as a meaning-related one, since it is impossible to determine with certainty what word the writer intended. For example, if *för att hon skulle kö* “because she was going to buy” was changed to *för att hon skulle göra* “because she was going to make,” this was classified as a wording revision. However, if such an editing operation involved a letter located next to the replacement letter on the keyboard, it was classified as a typographic correction (typo) and analyzed only in the overarching surface-revision category.

The classification of typographic and spelling revisions was performed in accordance with guidelines previously used for manual annotation (Wengelin, 2002; Stevenson et al., 2006). These categories are also sometimes hard to distinguish. A revision was considered to be a typographic correction where it involved the correction of an error due to a slip on the keyboard, for example a substitution, an omission, or an insertion involving a key adjacent to the target one, where the word prior to the edit did not conform to the orthographic rules or to a likely pronunciation of a word, or where the semantic context indicated that the form could not have been intended; this is in line with the guidelines given in Stevenson et al. (2006). To this taxonomy were added revisions that followed a deletion of the last letter(s) of a word, where a participant deleted

a word or phrase that the STT tool had transcribed incorrectly, (presumably) unintentionally deleted one (or more) letter(s) of the preceding word, and then immediately added the same letter(s) as the deleted one(s). As mentioned above, there are cases where it is hard to determine whether a revision is a typographic revision or a spelling revision, but on a general level it is likely that correcting a typographic error involves less thinking than making a spelling revision (Stevenson et al., 2006). For this reason, consideration of the composing process may make it easier to define such errors. In uncertain cases, the recordings of the composing process were allowed to guide the annotation.

In contrast to typographic errors, where the writer knows how to spell the target word, spelling errors are due to uncertainty about spelling. The following were considered to be spelling revisions: revisions of (a) an incorrectly spelled word into a correctly spelled word, (b) an incorrectly spelled word into another incorrectly spelled word, or (c) a correctly spelled word that was partially edited first into an incorrectly spelled word and then back into a correctly spelled word. All instances that were not considered to be typographic errors and did not match the wordlist of the Swedish Academy (the gold standard for Swedish spelling) were considered to be spelling errors; this follows the guidelines used in previous manual annotation in Swedish (Wengelin, 2016).

When annotating spelling revisions in children with spelling difficulties, one sometimes encounters revisions that could be the consequence of spelling difficulties, such as a revision from a long, incorrectly spelled word with a complex structure into one

or two words that are easier to spell (that is, if the initial word is not solely dependent on phoneme-grapheme correspondence, but also on orthographic, morphological, and/or morpho-syntactic knowledge). However, it cannot be ruled out that such revisions are instead revisions of wording. In such cases, a revision was classified as a spelling revision if there was any revising behavior prior to the word change. If the word was substituted without any prior revision, it was instead classified as a wording revision. This type of revisions will be discussed below.

## 2.3. Data analysis

I used a mixed-methods approach to investigate how revisions are manifested in children's writing. To answer the first and third research questions, about whether there were differences between groups and conditions in terms of revisions at different levels during the process, errors left in the final text product, and text quality, two-way ANOVAs were carried out for each category. For the STT errors, a Mann-Whitney *U*-test was used, since this only included group comparison. To answer the second research question, about how children manage spelling-related obstacles when composing with STT, instances involving spelling during the process were annotated and classified into representative categories.

Prior to analysis, the revision frequencies were standardized for all levels in order to correct for variation in text length. Standardization was carried out on the basis of the proportion of revisions, which was calculated by dividing the frequency of the revision type in the composition process by the total number of words in the final text, following the procedure in [Stevenson et al. \(2006\)](#). The revision types analyzed were overall revisions; surface revisions with the subcategories of wording revisions, spelling revisions, and STT-error revisions; and meaning-related revisions with no subcategories. The proportions of spelling errors and STT errors left in the final text product were calculated by dividing the number of spelling and STT errors in the final text product by the number of words in the final text product.

Text quality was assessed using comparative judgment by four raters who had not been involved in data collection. This was done on the [Nomoremarking.com](#) website ([No More Marking, 2021](#)), where the assessor was presented with two texts from the material, side by side, and asked to decide which was better. According to [Verhavert et al. \(2019\)](#), 19–20 assessments of each text are needed to reach “Scale Separation Reliability” (SSR) of at least 0.80. For this reason, each text was assessed at least 20 times. The judgments yield a scaled score ranging from 0 to 100 for each included text. SSR, which is considered to correspond to Cronbach's alpha ([Pollitt, 2012](#)), reached 0.91.

## 3. Results

The results are presented below by research question. First, the results regarding process data on revisions are presented for each level (first research question). Then the qualitative analysis of obstacles related to STT errors and spelling is presented (second research question). Finally, results regarding differences in the

characteristics of the final text product—spelling errors, STT errors, and text quality—between the conditions and groups are presented (third research question). Notably, the amount of error correction needed was dependent on the accuracy (and length) of the segments (bursts) produced. The mean length of production bursts was  $M = 3.30$  (3.24) for *Spell* and  $M = 3.05$  (2.63) for *Ref*, but this will not be discussed further here. See [Kraft et al. \(2022\)](#) for more details about what strategies for transcription the participants used and its contribution to fluency.

[Table 3](#) shows descriptive statistics for both process and product measures, by group and composing condition. For readability, the proportion has been multiplied by 100, thus generating the number of revisions per 100 words. A max value for revisions exceeding 100 means that the number of revisions made during the composition process exceeds the total number of words ending up in the final text product.

## 3.1. Revisions during the composition process

[Figure 1](#) shows the investigated measures from the composition process: total revisions, surface revisions, wording revisions, spelling revisions, STT-error revisions, and meaning-related revisions, by group and composing condition (since STT-error revisions occur only in the STT condition, there is no interaction plot for those revisions, but a boxplot showing group differences). Note that spelling errors thus occur in the STT condition as well.

### 3.1.1. Total revisions

[Table 3](#) shows that, overall, both groups made more revisions in the STT condition than in the keyboard condition. There was no significant interaction effect between group and condition for the proportion of revisions ( $p = 0.935$ ), and no main effect of group ( $p = 0.685$ ). However, there was a significant main effect of composing condition on the proportion of revisions [ $F_{(1,52)} = 9.839$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.16$ ]. This means that both groups revised more in the STT condition; see [Figure 1A](#).

### 3.1.2. Total surface revisions

Surface revisions as a main category was the predominant revision type in both modalities, for both groups. Similarly to the finding for total revisions, both groups made more surface revisions in the STT condition. Just as for overall revisions, there was no interaction effect between group and condition ( $p = 0.86$ ). Main-effect analysis showed that there was a significantly higher proportion of surface revisions in the STT condition [ $F_{(1,52)} = 10.41$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.17$ ], but there were no differences between the groups ( $p = 0.557$ ). This means that both groups had a higher proportion of surface revisions in the STT condition; see [Figure 1B](#).

### 3.1.3. Wording revisions

Regarding the proportion of wording revisions, there was no interaction effect between group and condition ( $p = 0.489$ ). Main-effect analysis showed that neither composing condition ( $p = 0.059$ )



**TABLE 3** Descriptive statistics relating to revisions at different levels performed during composition, to text length, to errors in the final text product (spelling errors and STT errors), and to text quality.

Spell	STT			Keyboard		
	M(sd)	Median	Min-max	M(sd)	Median	Min-max
<b>Process measures</b>						
Total revisions/100 words	50(24)	48	15–89	30(19)	27	0–78
Total surface revisions/100 words	47(24)	47	13–85	26(15)	27	0–66
Wording revisions/100 words	2(2)	3	0–5	4(3)	3	0–9
Spelling revisions/100 words	1(1)	0	0–5	5(5)	3	0–14
STT-error revisions/100 words	40(23)	42	11–80	NA	NA	NA
Meaning-related revisions/100 words	3(3)	3	0–12	4(6)	2	0–21
<b>Product measures</b>						
Text length	92.00(47.56)	86	20–190	79.38(46.35)	68.5	15–173
Spelling errors left/100 words	2(2)	0	0–7	13(8)	15	0–28
STT errors left/100 words	3(3)	3	0–11	NA	NA	NA
Spelling errors + STT errors/100 words	5(4)	4	0–18	13(8)	15	0–28
Text quality	51.94(9.82)	48.00	36.00–73.00	49.75(14.59)	46.00	30.00–87.00
Ref	STT			Keyboard		
	M(sd)	Median	Min-max	M(sd)	Median	Min-max
<b>Process measures</b>						
Total revisions/100 words	52(31)	46	19–133	33(19)	31	5–82
Total surface revisions/100 words	49(32)	45	15–131	31(19)	28	5–80
Wording revisions/100 words	3(4)	1	0–12	5(5)	4	0–15
Spelling revisions/100 words	1(2)	0	0–7	3(2)	2	0–8
STT-error revisions/100 words	43(30)	35	11–119	NA	NA	NA
Meaning-related revisions/100 words	3(2)	2	0–7	2(2)	2	0–6
<b>Product measures</b>						
Text length	106.42(61.19)	100.50	42–212	122.42(74.64)	106.50	34–245
Spelling errors left/100 words	0(1)	0	0–2	4(4)	3	0–14
STT errors left/100 words	1(2)	0	0–8	NA	NA	NA
Spelling errors + STT errors/100 words	2(2)	1	0–8	4(4)	3	0–14
Text quality	63.58(19.19)	65.00	32.00–94.00	65.92(17.35)	64.00	42.00–100.00

M(sd), median, min–max–values by group and composing condition.

or group ( $p = 0.157$ ) had a significant effect on wording revisions, meaning that neither group nor condition affected the proportion of wording revisions; see [Figure 1C](#).

### 3.1.4. Spelling revisions

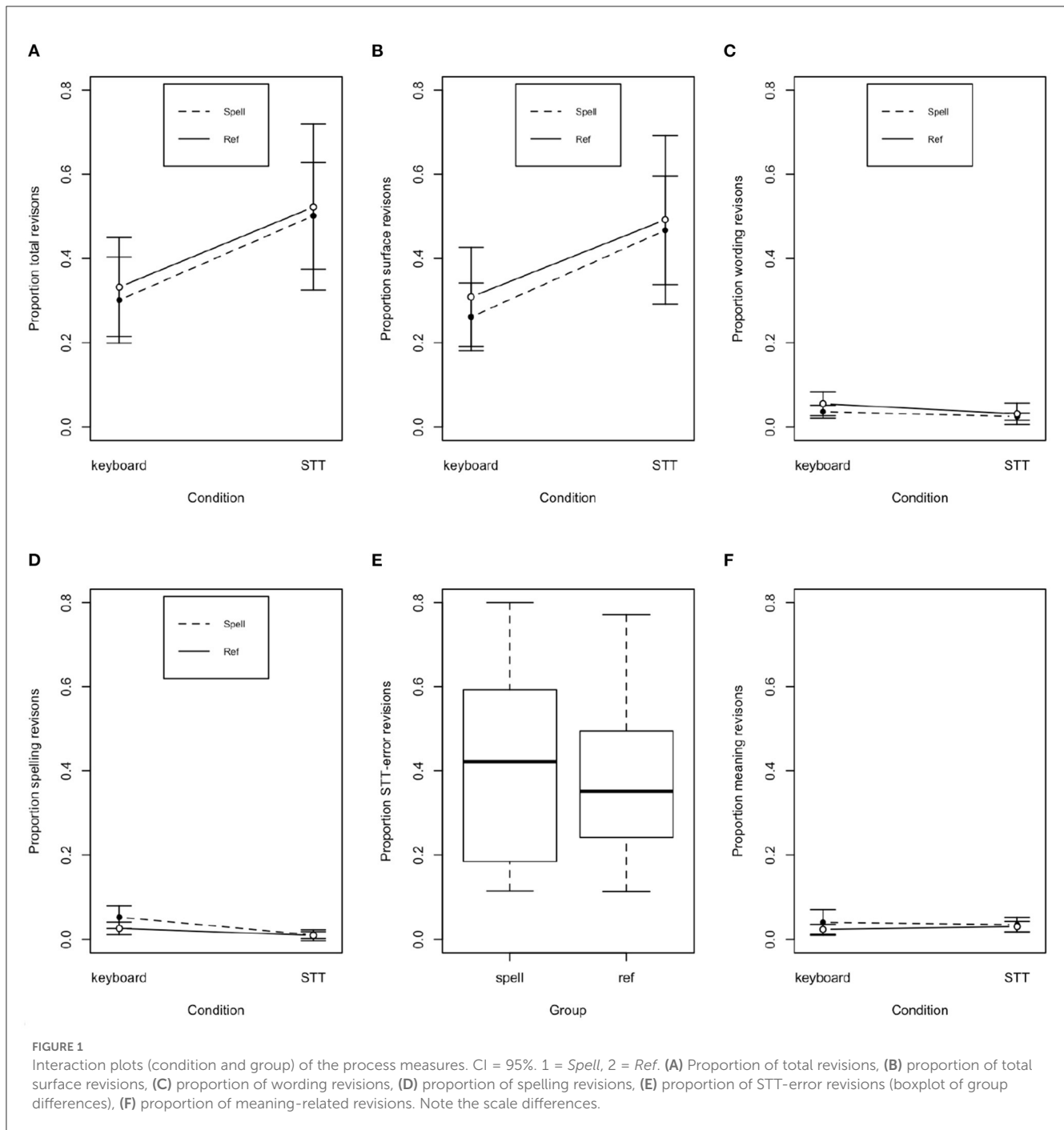
Regarding the proportion of spelling revisions, there was no interaction effect between group and condition ( $p = 0.14$ ). Main-effect analysis showed that composing condition had a significant effect on the proportion of spelling revisions [ $F_{(1,52)} = 13.65$ ,  $p \leq 0.001$ ,  $\eta_p^2 = 0.21$ ] but that there was no main effect for group ( $p = 0.11$ ), meaning that the groups did not differ in the proportion of spelling revisions but that both groups revised spelling significantly more often in the keyboard condition; see [Figure 1D](#).

### 3.1.5. STT-error revisions

Regarding proportion of STT-error revisions, the Mann-Whitney  $U$ -test did not show a significant difference between the groups ( $p = 0.830$ ), meaning that the groups did not differ in the proportion of STT-error revisions that they performed during composition; see [Figure 1E](#).

### 3.1.6. Meaning-related revisions

Regarding the proportion of meaning-related revisions, there was no interaction effect between group and condition ( $p = 0.549$ ) and main-effect analysis showed no effect of either group ( $p = 0.304$ ) or condition ( $p = 0.975$ ), meaning that neither group nor



condition affected the proportion of meaning-related revisions; see [Figure 1F](#).

### 3.2. Management of spelling when composing by means of STT

One main reason why composing by means of STT might be appropriate for children with spelling difficulties is the elimination of the spelling process. However, even though the STT tool does

not make any spelling errors as such, it sometimes misinterprets the writer's speech and produces semantic errors (an English-language example is that the writer dictates "if the wish is" and the tool transcribes this as "if the witches"). When the interaction between the child and the tool fails, the child must either re-dictate the words concerned, which might be hard to spell, or—if that does not work—type those words on the keyboard (and hence inevitably spell them). In this section, I will describe the categories of obstacles identified (regarding STT errors and spelling) as occurring during composition by means of STT as well as the solutions used by the children, sometimes in interaction with the tool. In this context,

I will suggest how these obstacles can be managed and point out some problems that need to be considered when STT is used as a means for composition. The categories identified are four in number: (a) The STT tool spells difficult words correctly, but not immediately; (b) The STT tool creates spelling errors; (c) The STT tool creates uncertainty; and (d) The writer incorporates an unintended STT transcription into the emerging text.

### 3.2.1. The STT tool spells difficult words correctly, but not immediately

Consonant doubling, due to vowel length, is a challenge in Swedish (Naucclér, 1980). There are cases in the data where the STT tool provided the correct spelling after initially producing a semantic error that the writer detected and deleted. For example, one participant dictated *innan han fuskar* “before he cheats,” which the tool transcribed as *inga några buskar* “not some bushes.” The participant then started editing on the keyboard and typed *in* followed by a long pause (during which the participant was probably thinking about whether there should be one or two <n> in *innan* “before”). Next, the participant deleted the two letters *in* and re-dictated *innan han fuskar*, which the STT tool transcribed (and thus, spelled) correctly on its second attempt.

Irregular spelling—in this case that certain phonemes correspond to multiple grapheme combinations—is also problematic in Swedish, and there are examples where the STT tool both misinterprets and provides the correct spelling. For example, one participant dictated *stjåla* “to steal,” and the tool incorrectly transcribed *valla* “to herd”). The participant chose to use the strategy of re-dictating the word, and the tool produced a correct transcription on its second attempt.

### 3.2.2. The STT tool creates spelling errors

In Swedish, closed compounding is common and productive. However, the STT tool tends to write in two words what should be only one. For example, the tool incorrectly transcribed *fuskklapp* “cheat sheet” as two separate words: *fusk lapp*.

### 3.2.3. The STT tool creates uncertainty about spelling

There are examples where participants are “doubly punished” when interacting with the tool, and where signs of uncertainty regarding spelling can be noted. These are cases where a participant is uncertain about the spelling of a word and tries to use the STT tool to produce the correct spelling, but the STT tool, instead of providing the correct spelling, misinterprets the spoken input and produces an additional transcription error that the participant has to deal with. This increases the participant’s need to engage in problem-solving. In fact, since the STT tool does not always generate the correct spelling even on its second or subsequent attempts, the user may be reduced to typing, and hence to engaging in a spelling process of potentially high cognitive cost. One example: A participant with reading and writing difficulties dictated *det man vill bli* “what you want to become” and the tool wrote *det man vill be* “what you want to pray.” The participant then used the keyboard to edit *be* into *ble*, producing

an incorrect (although perhaps dialectally feasible) spelling of *bli* “become.”

One common editing strategy observed was to re-dictate only the single word that the STT tool had transcribed incorrectly. However, since this strategy gives the tool very little by way of context (see Kraft et al., 2022, on what enables an accurate and fluent transcription when children compose with STT in Swedish), there is a great risk of misinterpretation on the part of the tool and, as a consequence, of uncertainty about the correct spelling on the part of the writer. For example, one participant (without reading and writing difficulties) tried to dictate *sett* “seen,” but the tool transcribed *speciellt* “specially.” Next, the participant re-dictated *sett* and the tool this time transcribed it as its homophone *sätt* “manner.” The participant again re-dictated *sett*, and the tool produced *fett* “grease.” The presumably exasperated participant then switched to the keyboard and typed *sett*. However, interestingly, what happened next was that the participant started to revise the word and deleted the vowel, conceivably as a consequence of the STT tool having produced the homophone *sätt* on its second attempt. While it is perfectly possible that the participant could have hesitated about the spelling of the word in question even without the STT tool’s misinterpretation, it certainly did not help. Even so, the participant finally typed *sett* again, producing the correct spelling. In cases like this, the STT tool does not facilitate the spelling process, but instead risks increasing the writer’s cognitive load. Another example from the data involves a participant (with reading and writing difficulties) who dictated *rädd att* “afraid that,” which the tool transcribed as *rabatt* “discount,” forcing the participant to edit the word by means of the keyboard. However, the participant showed uncertainty when choosing whether to include one or two <d> in *rädd* “afraid,” as evidenced by a long pause and a space after typing the first <d>. The participant went on to delete that space and add a second <d>, ending up with a correctly spelled word, but the process of arriving at that spelling was without doubt cognitively demanding.

### 3.2.4. Incorporating unintended STT transcriptions (STT errors) in the emerging text

There are also instances in the data where a participant used transcription errors produced by the STT tool and incorporated them in the emerging text. For example, one participant with reading and writing difficulties dictated *skulle säga att man inte skulle fuska i prov* “would say that you are not supposed to cheat in a test” but the STT tool transcribed *skulle säga man skulle fuska på* “would say that you should cheat on,” whereupon the participant finished the now-incomplete sentence by adding *ett prov* “a test.” Hence the principal content was changed from *should not cheat* to *should cheat*, radically changing the meaning and global content of the whole text.

Another interesting consequence of STT errors is that they can actually be kept by participants in the text written so far, where the misinterpretations may provide a source for new ideas to be evaluated and possibly incorporated in the emerging text. One example is where a participant dictated *Han borde bett om hjälp istället* “He should have asked for help instead” and the tool produced *Han borg du vet om hjälp istället* “He castle you

know for help instead.” The participant went on to use the verb *veta* “know” inserted through the STT error, re-dictating *Han borde veta bättre än att tjuvkika* “He should know better than to peek.”

To conclude, the STT tool can help with difficult spelling issues such as consonant doubling and irregular spelling when children compose, but it can also create uncertainty and produce repeated semantic errors that the children have to consider and correct.

### 3.3. Characteristics of the final text product

Figure 2 shows the measures investigated when it comes to the characteristics of the final text product: the proportion of spelling errors left, the proportion of STT errors left, the sum of spelling and STT errors left, and overall text quality, by group and composing condition (since STT errors occur only in the STT condition, there is no interaction plot but a boxplot showing group differences). Note that spelling errors thus occur even in the STT condition.

#### 3.3.1. Errors left in the final text product

Regarding the proportion of spelling errors left in the final text product, the analysis revealed that there was a statistically significant interaction between the effects of group and condition [ $F_{(1,52)} = 8.88, p = 0.004, \eta_p^2 = 0.15$ ], meaning that the combined effect of group and condition had a significant effect on the proportion of misspellings left in the text. Simple main-effect analysis showed that both group [ $F_{(1,52)} = 16.40, p \leq 0.001, \eta_p^2 = 0.24$ ] and condition [ $F_{(1,52)} = 37.90, p \leq 0.001, \eta_p^2 = 0.42$ ] had a statistically significant effect on the proportion of spelling errors left in the text; see Figure 2A.

Regarding the proportion of STT errors left in the final text product, the Mann–Whitney *U*-test showed a significant difference between the groups ( $p = 0.01$ ), meaning that the *Spell* group left proportionally more STT errors in their final texts than the *Ref* group; see Figure 2B.

To enable comparison of the overall prevalence of product errors across conditions and groups, the spelling and STT errors left in the final text product were summed up. Analysis revealed that there was no statistically significant interaction between the effects of group and condition [ $F_{(1,52)} = 3.99, p = 0.051$ ]. Simple main-effect analysis, however, showed that both group [ $F_{(1,52)} = 19.43, p \leq 0.001, \eta_p^2 = 0.27$ ] and condition [ $F_{(1,52)} = 14.92, p \leq 0.001, \eta_p^2 = 0.22$ ] had a statistically significant effect on the proportion of errors left in the final text product. This result shows that both groups had significantly fewer errors in the final text product in the STT condition, and that the *Ref* group had significantly fewer errors than the *Spell* group. As is clear from the descriptive statistics in Table 3, the difference in the proportion of errors between conditions was greater for the *Spell* group; see Figure 2C for an interaction plot.

Out of interest, an additional correlation analysis (beyond the research questions) was performed with regard to the proportion of surface revisions in the two composing conditions, to investigate whether those participants who made proportionally more surface revisions in one condition also did so in the other. No significant

correlation was found ( $r = -0.14, p = 0.47$ ), meaning that the two writing conditions yielded different revising behaviors ( $r = -0.28, p = 0.30$  for *Spell*;  $r = 0.20, p = 0.54$  for *Ref*).

#### 3.3.2. Text quality

Regarding text quality, the analysis revealed that there was no statistically significant interaction between the effects of group and condition ( $p = 0.58$ ). Simple main-effect analysis showed that group had a statistically significant effect on text quality [ $F_{(1,52)} = 11.49, p = 0.001, \eta_p^2 = 0.18$ ] while condition had no statistically significant effect on text quality ( $p = 0.95$ ). This result shows that the *Ref* group produced texts of higher assessed quality regardless of writing condition; see Figure 2D.

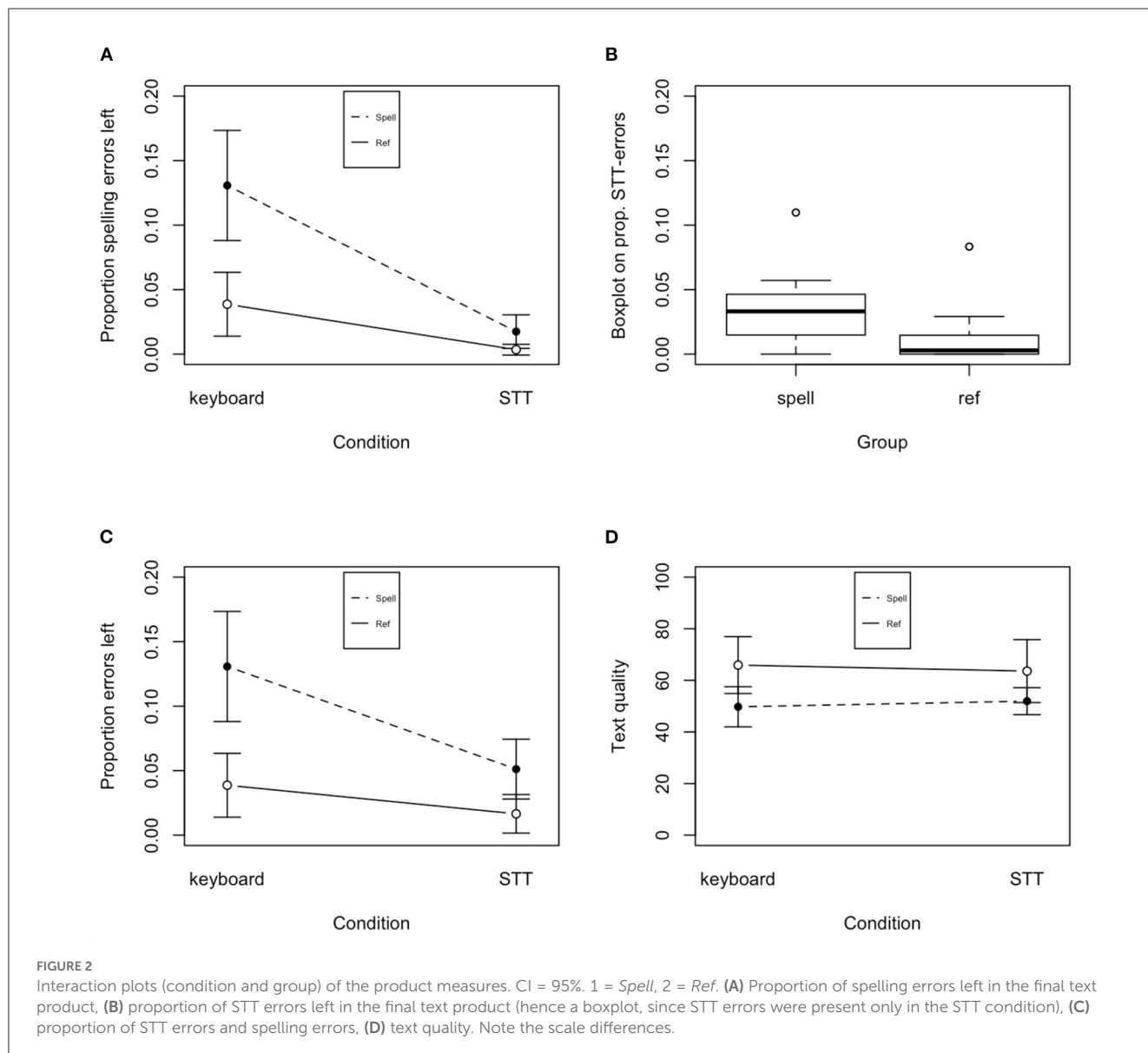
## 4. Discussion

This study is unique in studying composition processes in children using STT. Its general aim was to investigate whether there were any process differences in terms of revisions at various levels in children with and without reading and writing difficulties (due to underlying decoding and spelling difficulties) composing by means of STT and a keyboard, respectively, as well as whether the final text product differed by group and condition in terms of the proportion of spelling and STT errors left in the final text and in terms of text quality. Since spelling is especially hard for children with reading and writing difficulties, a more specific aim was to explore how children manage revisions related to spelling during composition with STT, in order to identify obstacles as well as opportunities associated with the use of STT as a means of composition.

As regards the management of errors (spelling errors and STT errors) during composing by means of STT, the analysis showed that the STT tool could both facilitate and aggravate this process. STT could facilitate spelling in Swedish when it came to both consonant doubling and irregular spelling, but when the tool misinterpreted the writers' speech, they were forced to perform an additional problem-solving process and sometimes ended up having no other option than to spell the words themselves using the keyboard after all. Moreover, the STT tool showed an inadequate ability to transcribe closed compounds correctly. One overall conclusion to be drawn from the above is that children who are to compose using an STT tool must be taught what transcription strategies are most effective (one example being to avoid dictating words one by one; see Kraft et al., 2022) and must be given knowledge about the shortcomings of the STT tool.

Overall, the proportion of meaning-related revisions was small, and there was no difference either between writing conditions or between groups. The rarity of meaning-related revisions was an expected result, since these have previously been shown to be rare in this age group (for handwriting) (Chanquoy, 2001; Limpo et al., 2014), although group differences have in fact been reported between adults with and without reading and writing difficulties composing by means of a keyboard (Wengelin, 2002). Since those studies differed in the composing condition explored, there is a possibility that differences in processing demands could have affected the results to some extent. However, since the results of the





present study (for both STT and keyboard) did not show any group differences in terms of revising at the level of meaning, it seems likely that the rarity of meaning-related revisions mainly reflects the young age of the children investigated. According to capacity theory, it should be expected that the children without reading and writing difficulties would have progressed farther in their development of meaning-related revision, since the automatization of transcription skill will have developed faster in this group, freeing up cognitive capacity to enable the development of higher-level processes such as revision. However, for the children in the present study, and as previously reported for the age group in question, revisions of meaning are not yet common, and the use of STT did not change this. In fact, the finding that none of the groups made meaning-related revisions to a high extent can be seen to complement previous research about revising behavior in children of the same age (Limpo et al., 2014) by adding descriptive data on revising behavior in two additional modalities: keyboard and

STT, and in an additional language: Swedish. A further point to be emphasized is that the present study investigated revisions in real time as they emerge during composition. In other words, it explored functional writing, which is an approach that has been called for in previous studies of children's revisions (Limpo et al., 2014). Further, it should be noted that variation in terms of meaning-related revision was greater in the *Spell* group than in the *Ref* group, especially for composing by keyboard. This underscores the importance for future research to investigate the development of meaning-related revisions in developing writers, and of linking revising behavior both to outcome measures such as text quality and to individual abilities that contribute to revising behavior. It must be stressed that meaning-related revisions do not always contribute to better text quality and that formulation processes and the generation of ideas depend both on individual language abilities and on long-term memory pertaining, for example, to genre knowledge and to experience with and exposure to written

text. Hence the relationship between the amount of meaning-related revising and text quality is not necessarily linear. Detecting and analyzing content errors in the text written so far, deciding whether or not to revise, and performing the actual revisions are all demanding processes that do not necessarily improve quality. However, with increasing experience, these processes will most likely develop their efficiency and effectiveness, causing the revision of content to become an increasingly goal-oriented process that will make ever greater contributions to text quality. As [Limpo et al. \(2014\)](#) noted at the group level, the amount of meaning-related revisions did not predict text quality for children in grades 4-6. The considerable variation seen in this study in terms of the amount of revisions that change meaning can in fact reflect developmental aspects of children's meaning-related revisions. Previous research has shown that transcription processes are more important in developing writers ([Connelly et al., 2007](#); [Kim and Park, 2019](#)), while higher-level processes such as the revision of content will develop once the former processes have been sufficiently automatized, at which point writing can be used as a tool to transform knowledge. For this reason, longitudinal research is needed to fully understand the development of revising behavior in children with reading and writing difficulties who use an STT tool rather than a keyboard or a pen for composing in order to facilitate the technical aspects of the composition process. In the present study, the participants were writing in a single session, which further could have contributed to the low number of meaning-related revisions, since it is possible that their teachers have taught them to write a first draft and to revise later. In addition, professional writers composing with STT tend to postpone revisions of content more often than revisions of STT-errors, which are more likely to be solved immediately ([Leijten et al., 2010](#)). Furthermore, the task was low-stakes and the incentive to change and improve the content of the text may have been low.

Since the development of revising skill is dependent on a complex interaction of strategic writing knowledge and performance, along with topic knowledge (see [Torrance et al., 2007](#), for an expanded discussion), the effect of using a facilitatory tool (for transcription) will not necessarily show up as a gain in higher-level processes such as revision simply as a result of the removal of the burden of transcription. At the very least, there is likely to be a need to teach lower-level transcription skills (adapted for the use of STT) alongside higher-level composition-related instruction (see also [Berninger et al., 2002](#)). In this context, the quality of writing instructions should be addressed. In fact, evidence-based instructions in writing strategies are essential for children's writing development ([Graham and Harris, 2017](#); [Graham, 2019](#)).

In contrast to meaning-related revisions, surface revisions were proportionally more common—in both groups—when composing with STT, because of the numerous STT errors. The need to deal with those frequent errors makes composing with STT a cognitively demanding process, and this could also help to explain the finding that the proportion of meaning-related revisions did not increase even though the burden of spelling had decreased. Hence the present study highlights the need to consider the accuracy of the STT tool prior to implementation, because failure to do so entails a risk that the correction of STT errors will be a burdensome task, possibly on a par with traditional spelling. That there was no

difference in text quality between conditions could be interpreted as a further argument suggesting that no cognitive capacity was actually freed up in composing with STT. However, it is important to remember that the participants had no prior experience of composing text with STT. Hence, once again, there is a great need to investigate the effect of using STT as a facilitatory tool over time, and to train writers in the use of the STT tool prior to having them use it for composing. It should also be noted that, regardless of composing condition, the *Ref* group produced texts of higher assessed quality. The ability to produce high-quality texts is dependent on genre knowledge and prior text experience, and it is well known that children with reading and writing difficulties due to underlying decoding and spelling difficulties generally read and write less than children without such difficulties, meaning that they receive less text experience and so obtain less genre knowledge. Since the present study did not control for reading and writing habits, this could be part of the explanation for its findings. A further possible interpretation is that, because of the numerous STT errors, the STT condition did not in fact free up any additional cognitive capacity for higher-level processes. However, to fully understand the implications of the present findings, STT composition must be investigated over time, preferably combined with writing instruction in fields such as revising and/or planning.

As regards the proportion of spelling revisions, the *Spell* group did not revise spelling proportionally more than the *Ref* group. Hence their final text products inevitably contained proportionally more spelling errors. There are several possible explanations for this. First, the *Spell* writers might not have detected their spelling mistakes. This could be due to the experimental situation in Scriptlog, which differs from what those writers are probably used to from their everyday writing in that it does not highlight incorrect spellings with a red underscore. Second, the *Spell* writers might have abstained from trying to correct their errors, either as a strategic choice due to an awareness of their shortcomings in revising spelling errors (research has shown that children with reading and writing difficulties succeed in less than half of their spelling revisions; [Wengelin et al., in preparation](#)) or because they had previously been told by their teachers to focus less on spelling and more on content.

One risk associated with composing by means of STT that has been mentioned above is that STT errors might burden working memory to the same extent as spelling errors. The results of the present study showed that surface errors were indeed significantly more common in the STT condition, precisely because of the numerous STT errors. By contrast, the overall transcription errors in the final text product (spelling errors and STT errors combined) were significantly fewer in composing with STT for both groups, and the difference was even greater for the *Spell* group. In other words, it would seem that the STT condition encourages children to monitor the text written so far and—more importantly—that their monitoring is highly successful. Previous studies have shown that children with reading and writing difficulties (15-year-olds composing by keyboard) read their text written so far to the same extent as their peers without such difficulties, but that they read more slowly ([Wengelin et al., 2014](#)). The results of the present study corroborate the finding that children with reading and writing difficulties read their own text, since the present participants

managed to deal with the STT errors to a large extent (that is, they both detected errors by reading and then corrected them to a similar extent as their peers without reading and writing difficulties). In this context, it is worth mentioning that it would be valuable to investigate motivational factors associated with the fact that the errors corrected had been made by a tool rather than by the writers themselves, since the awareness of shortcomings in one's writing ability can have negative effects on motivation and overall perceptions about writing (Waldmann et al., 2022), an issue that should not be underestimated. For this reason, future research should highlight motivational factors associated with composing by means of STT and with reading and writing difficulties.

## 4.1. Limitations

The time-consuming nature of the annotation of STT-process data makes it hard to build large amounts of research data, and general conclusions should be interpreted with caution because of the small number of participants. However, the mixed-methods approach, combining inferential statistics with more in-depth analysis regarding the obstacles and solutions observed during the STT composition process (concerning STT errors and spelling errors), is a strength of the present study, which remains valid and from which it is possible to draw instructional conclusions. Further, its results can be used to build hypotheses for further exploration.

The low inter-rater reliability for annotating meaning-related revisions should be discussed. Even though the raters in the present study reached similar values as the raters in Conijn et al. (2022), this emphasizes how hard it is to capture revisions of meaning. The criteria in question may be too broad, and future studies could potentially benefit from adding further information (and examples) to the present criteria to be able to reach a higher inter-rater reliability.

The results suggested that the participants with reading and writing difficulties might not have detected (or might have chosen not to correct) some of their spelling mistakes, since they left proportionally more spelling errors in their final text product than the participants without such difficulties when composing with keyboard. However, for composing with STT, the results showed that they both detected and corrected errors to a high extent.

## 4.2. Conclusion and implications for teaching

The general conclusion to be drawn from the present study is that when children with reading and writing difficulties, due to underlying decoding and spelling difficulties, compose with STT, they leave fewer errors in their final text product, even though they need to engage more in the correction of surface errors because of the large number of STT errors. Further, the difference between the conditions in the proportion of errors left in the text was greater for the children with reading and writing difficulties. In other words, the *Spell* group gained more from composing with STT than the *Ref* group did. Despite the numerous STT errors,

and the need to correct them, neither the proportion of meaning-related revisions nor text quality decreased in composing with STT. Taken together, these results suggest, albeit not emphatically, that STT may be appropriate as a facilitatory tool for children with reading and writing difficulties. However, the participants in the present study had no prior experience with STT, and children need to have learned appropriate transcription strategies for composing with STT before they can use this method effectively by avoiding an excess of STT errors and the attendant need to engage in problem-solving during the process.

Further, the results showed that there was no difference in text quality between conditions for the children with reading and writing difficulties (nor for the reference group, for that matter). In part, this could be explained by the fact that the children had to direct their focus, and hence their cognitive capacity, toward the local word level, because they had to detect and correct errors. This, in turn, could have hindered higher-level composition processes. However, since text quality did not decrease although the participants had no prior experience writing with STT, these results are actually quite promising and should prompt further studies on STT as a facilitatory writing tool.

When children grow older, the demand on a more elaborate text structure will increase (see Beers and Nagy, 2009). To produce a long sentence using more complex syntactic structures might be hard to achieve in speech, and the capacity to do so may depend on underlying cognitive and linguistic abilities. This process may further be hindered if the tool accuracy is low. On the other hand, it could also be that composing by STT may facilitate the production of these more complex structures, since the writer has no need to keep their eyes on the keyboard, but instead can focus on the forthcoming text on the screen, and the text could then possibly be used as a strategy to keep track on the planned text kept in working memory. Therefore, future research should couple STT process data with eye-tracking to investigate developmental possibilities or obstacles related to monitoring, error correction and text production in children composing with STT.

Finally, the *Ref* group generally produced texts of higher quality across writing conditions. This might reflect their greater prior text exposure and their reading and writing habits in general (see also Sumner and Connelly, 2020, for a similar discussion) and therefore highlights the need to investigate early implementation of facilitatory tools such as STT for children with reading and writing difficulties, in order to explore its capacity for bridging the gap in reading and writing development that has been found in previous research (Stanovich, 2009).

However, it would be ignorant to assume that a facilitatory tool will be enough to develop these children's writing. Adequate instruction in higher-level processes is also needed to develop sufficient writing strategies (see Graham, 2019, for an overview). Future research should therefore investigate instruction that addresses strategies for STT transcription, highlights the shortcomings of the tool in the target language, and also focuses specifically on higher-level aspects of composition such as planning or revising, in order to gain further knowledge about the feasibility of using STT as a means of composition for children who struggle with writing due to underlying decoding and spelling difficulties, including about its possible effects over time.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by Central Ethical Review Board, Gothenburg. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## Author contributions

The author confirms sole responsibility for the following: study conception and design, data collection, analysis, interpretation of results, and manuscript preparation.

## Funding

This research was funded by the Marcus and Amalia Wallenberg Foundation (Ref. No. 2014–0122).

## Acknowledgments

The author would like to acknowledge Celia Wik Mergulhão, Petter Åström, and Ingrid Henriksson, Institute

of Neuroscience and Physiology, Department of Health and Rehabilitation, Gothenburg University for their contribution of text quality ratings, and Åsa Wengelin, Department of Swedish, Multilingualism, Language Technology, Gothenburg University, for her contribution of text quality ratings and constructive reading of the text. The author would also like to thank Maria Levlin, Department of Language Studies, Umeå University, for valuable comments on the manuscript, and Johan Segerbäck for language editing.

## Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Alamargot, D., Chesnet, D., Dansac, C., and Ros, C. (2006). Eye and pen: a new device for studying reading during writing. *Behav. Res. Methods* 38, 287–299. doi: 10.3758/BF03192780
- Alamargot, D., and Fayol, M. (2009). "Modelling the development of written composition," in *The SAGE Handbook of Writing Development*, eds R. Beard, J. Riley, D. Myhill, and M. Nystrand (London: Sage Publications), 23–47.
- Aro, M. (2006). "Learning to read: the effect of orthography," in *Handbook of Orthography and Literacy*, eds R. Malatesha Joshi and P. Aaron (London: Routledge), 545–564.
- Beers, S. F., and Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: which measures? Which genre? *Read. Writ.* 22, 185–200. doi: 10.1007/s11145-007-9107-5
- Berman, R., and Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: speech and writing. *Written Lang. Literacy* 5, 1–43. doi: 10.1075/wll.5.1.02ber
- Berninger, V. W., and Swanson, H. L. (1994). "Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing," in *Advances in Cognition and Educational Practice, Volume 2: Children's Writing: Towards a Process Theory of the Development of Skilled Writing*, Volume 2, eds J. S. Carlson and E. C. Butterfield (Greenwich, CT: Jai Press Inc.), 57–81.
- Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., et al. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *J. Educ. Psychol.* 94, 291. doi: 10.1037/0022-0663.94.2.291
- Chanquoy, L. (2001). How to make it easier for children to revise their writing: a study of text revision from 3rd to 5th grades. *Br. J. Educ. Psychol.* 71, 15–41. doi: 10.1348/000709901158370
- Chenoweth, N. A., and Hayes, J. R. (2001). Fluency in writing: generating text in 11 and 12. *Written Commun.* 18, 80–98. doi: 10.1177/0741088301018010004
- Conijn, R., Speltz, E. D., Zaanen, M. V., Waes, L. V., and Chukharev-Hudilainen, E. (2022). A product-and process-oriented tagset for revisions in writing. *Written Commun.* 39, 97–128. doi: 10.1177/07410883211052104
- Connelly, V., Gee, D., and Walsh, E. (2007). A comparison of keyboarded and handwritten compositions and the relationship with transcription speed. *Br. J. Educ. Psychol.* 77, 479–492. doi: 10.1348/000709906X116768
- De La Paz, S., Swanson, P. N., and Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *J. Educ. Psychol.* 90, 448. doi: 10.1037/0022-0663.90.3.448
- Delattre, M., Bonin, P., and Barry, C. (2006). Written spelling to dictation: sound-to-spelling regularity affects both writing latencies and durations. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1330. doi: 10.1037/0278-7393.32.6.1330
- ELAN (2019). *Version 5.6 [Computer software]*. Nijmegen: Max Planck Institute for Psycholinguistics, the Language Archive. Available online at: <https://tla.mpi.nl/tools/tla-tools/elan/> (accessed October 9, 2019).
- Elwér, Å., Fridolfsson, I., Samuelsson, S., and Wiklund, C. (2011). *Läst: Test i läsning och Stavning*. Stockholm: Hogrefe/Psykologiförlaget.
- Faigley, L., and Witte, S. (1981). Analyzing revision. *College Compos. Commun.* 32, 400–414. doi: 10.2307/356602
- Fitzgerald, J. (1987). Research on revision in writing. *Rev. Educ. Res.* 57, 481–506. doi: 10.3102/00346543057004481
- Flower, L., and Hayes, J. R. (1981). A cognitive process theory of writing. *College Compos. Commun.* 32, 365–387. doi: 10.2307/356600
- Frid, J., Johansson, V., Johansson, R., and Wengelin, Å. (2014). "Developing a keystroke logging program into a writing experiment environment," in *Writing Across Borders* (Paris).
- Galbraith, D. (2009). "Writing about what we know: Generating ideas in writing," in *The SAGE Handbook of Writing Development*, eds R. Beard, J. Riley, D. Myhill, and M. Nystrand (London: Sage Publications), 48–64.



- Graham, S. (2019). Changing how writing is taught. *Rev. Res. Educ.* 43, 277–303. doi: 10.3102/0091732X18821125
- Graham, S., and Harris, K. R. (2017). “Evidence-based writing practices: a meta-analysis of existing meta-analyses,” in *Design Principles for Teaching Effective Writing* (Leiden: Brill), 13–37.
- Haug, K. N., and Klein, P. D. (2018). The effect of speech-to-text technology on learning a writing strategy. *Read. Writ. Q.* 34, 47–62. doi: 10.1080/10573569.2017.1326014
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Commun.* 29, 369–388. doi: 10.1177/0741088312451260
- Hayes, J. R., and Berninger, V. W. (2014). “Cognitive processes in writing: a framework,” in *Writing Development in Children With Hearing Loss, Dyslexia, or Oral Language Problems: Implications for Assessment and Instruction*, eds B. Arfe, J. Dockrell, and V. W. Berninger (Oxford: Oxford University Press), 3–15.
- Higgins, E. L., and Raskind, M. H. (1995). Compensatory effectiveness of speech recognition on the written composition performance of postsecondary students with learning disabilities. *Learn. Disabil. Q.* 18, 159–174. doi: 10.2307/1511202
- Järpsten, B., and Taube, K. (2010). *DLS: för skolår 4-6.Handledning*. Stockholm: Hogrefe Psykologiförlaget.
- Kim, Y.-S. G., and Park, S.-H. (2019). Unpacking pathways using the direct and indirect effects model of writing (diew) and the contributions of higher order cognitive skills to writing. *Read. Writ. Q.* 32, 1319–1343. doi: 10.1007/s11145-018-9913-y
- Kraft, S., Rønneberg, V., Thurffjell, F., Rack, J., and Wengelin, Å. (2022). *Exploring Transcription Processes When Children With and Without Reading and Writing Difficulties Produce Written Text Using Speech Recognition*. Manuscript submitted for publication.
- Kraft, S., Thurffjell, F., Rack, J., and Wengelin, Å. (2019). Lexikala analyser av muntlig, tangentbordsskriven och dikterad text producerad av barn med stavningssvårigheter. *Nordic J. Literacy Res.* 5, 102–122. doi: 10.23865/njlr.v5.1511
- Leijten, M. (2007). *Writing and speech recognition: Observing error correction strategies of professional writers* (Dissertation). Netherlands Graduate School of Linguistics.
- Leijten, M., Janssen, D., and van Waes, L. (2010). Error correction strategies of professional speech recognition users: Three profiles. *Comput. Hum. Behav.* 26, 964–975. doi: 10.1016/j.chb.2010.02.010
- Leijten, M., and Van Waes, L. (2005). Writing with speech recognition: the adaptation process of professional writers with and without dictating experience. *Interact Comput.* 17, 736–772. doi: 10.1016/j.intcom.2005.01.005
- Leijten, M., and Van Waes, L. (2013). Keystroke logging in writing research: using inputlog to analyze and visualize writing processes. *Written Commun.* 30, 358–392. doi: 10.1177/0741088313491692
- Limpo, T., Alves, R. A., and Fidalgo, R. (2014). Children’s high-level writing skills: development of planning and revising and their contribution to writing quality. *Br. J. Educ. Psychol.* 84, 177–193. doi: 10.1111/bjep.12020
- Lu, X., Li, S., and Fujimoto, M. (2020). “Automatic speech recognition,” in *Speech-to-Speech Translation*, eds H. Kawai, E. Sumita, and Y. Kidawara (Singapore: Springer).
- MacArthur, C. A. (2009). Reflections on research on writing and technology for struggling writers. *Learn. Disabil. Res. Pract.* 24, 93–103. doi: 10.1111/j.1540-5826.2009.00283.x
- MacArthur, C. A., and Cavalier, A. R. (2004). Dictation and speech recognition technology as test accommodations. *Except. Child* 71, 43–58. doi: 10.1177/001440290407100103
- MacArthur, C. A., and Graham, S. (1987). Learning disabled students’ composing under three methods of text production: Handwriting, word processing, and dictation. *J. Spec. Educ.* 21, 22–42. doi: 10.1177/002246698702100304
- McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educ. Psychol. Rev.* 8, 299–325. doi: 10.1007/BF01464076
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educ. Psychol.* 35, 13–23. doi: 10.1207/S15326985EP3501\_3
- McCutchen, D., Covill, A., Hoyne, S. H., and Mildes, K. (1994). Individual differences in writing: implications of translating fluency. *J. Educ. Psychol.* 86, 256. doi: 10.1037/0022-0663.86.2.256
- McCutchen, D., Francis, M., and Kerr, S. (1997). Revising for meaning: Effects of knowledge and strategy. *J. Educ. Psychol.* 89, 667. doi: 10.1037/0022-0663.89.4.667
- Nauclér, K. (1980). *Perspectives on misspellings : a phonetic, phonological and psycholinguistic study* (dissertation). Travaux de l’Institut de linguistique de Lund, 15. Liber Läromedel/Gleerup, Lund.
- No More Marking (2021). *Online-Based Tool for Comparative Judgement*. Available online at: from [www.nomoremarking.com](http://www.nomoremarking.com) (accessed on May 02, 2021).
- Ok, M. W., Rao, K., Pennington, J., and Ulloa, P. R. (2022). Speech recognition technology for writing: usage patterns and perceptions of students with high incidence disabilities. *J. Special Educ. Technol.* 37, 191–202. doi: 10.1177/0162643420979929
- Plumb, C., Butterfield, E. C., Hacker, D. J., and Dunlosky, J. (1994). Error correction in text. *Read. Writ. Q.* 6, 347–360. doi: 10.1007/BF01028848
- Pollitt, A. (2012). Comparative judgement for assessment. *Int. J. Technol. Design Educ.* 22, 157–170. doi: 10.1007/s10798-011-9189-x
- Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: Improving fluency. *J. Educ. Psychol.* 96, 337. doi: 10.1037/0022-0663.96.2.337
- R., Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Scardamalia, M., and Bereiter, C. (1987). “Knowledge telling and knowledge transforming” in *Advances in Applied Psycholinguistics, Volume 2: Reading, Writing, and Language Learning, volume 2*, ed S. Rosenberg (Cambridge: Cambridge University Press Archive), 142–175.
- Seymour, P. H., Aro, M., Erskine, J. M., and Collaboration with COST Action A8 Network (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859
- Stanovich, K. E. (2009). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *J. Educ.* 189, 23–55. doi: 10.1177/0022057409189001-204
- Stevenson, M., Schoonen, R., and de Glopper, K. (2006). Revising in two languages: a multi-dimensional comparison of online writing revisions in l1 and fl. *J. Second Lang. Writ.* 15, 201–233. doi: 10.1016/j.jslw.2006.06.002
- Sumner, E., and Connelly, V. (2020). Writing and revision strategies of students with and without dyslexia. *J. Learn. Disabil.* 53, 189–198. doi: 10.1177/0022219419899090
- Sumner, E., Connelly, V., and Barnett, A. L. (2013). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Read. Writ. Q.* 26, 991–1008. doi: 10.1007/s11145-012-9403-6
- TechSmith (2018). *Camtasia. (Version 3.1.6, Copyright 2006–2018)*. Okemos, MI: TechSmith Corporation (accessed May 2018).
- Torrance, M., Fidalgo, R., and García, J.-N. (2007). The teachability and effectiveness of cognitive self-regulation in sixth-grade writers. *Learn. Instruct.* 17, 265–285. doi: 10.1016/j.learninstruc.2007.02.003
- Treiman, R. (2018). Statistical learning and spelling. *Lang. Speech. Hear. Serv. Sch.* 49, 644–652. doi: 10.1044/2018\_LSHSS-STLT1-17-0122
- Verhavert, S., Bouwer, R., Donche, V., and De Maeyer, S. (2019). A meta-analysis on the reliability of comparative judgement. *Assessment Educ.* 26, 541–562. doi: 10.1080/0969594X.2019.1602027
- Waldmann, C., Ranjesh, R., Malmström, A., Lindgren, E., and Levlin, M. (2022). “Ungdomars skrivpraktiker på fritiden,” in *Språk i skola, på fritid och i arbetsliv : aktuella arenor för Svensk Forskning Inom tillämpad Språkvetenskap: Language in School, Leisure time, and Working Life: Current Arenas for Swedish Research in Applied linguistics, ASLA: s Skrifserie, Vol. 29* (Växjö), 187–212.
- Wengelin, Å. (2002). *Text production in adults with reading and writing difficulties* (dissertation). Department of Linguistics, Gothenburg University, Gothenburg.
- Wengelin, Å. (2007). The word-level focus in text production by adults with reading and writing difficulties. *Stud. Writ.* 20, 67–82. doi: 10.1108/S1572-6304(2007)0000020006
- Wengelin, Å. (2016). “Hur stavar du lille vän,” in *Språk och Norm. ASLA:s skrifserie, Vol. 26* (Uppsala), 169–178.
- Wengelin, Å., Johansson, R., and Johansson, V. (2014). “Expressive writing in Swedish 15-year-olds with reading and writing difficulties,” in *Writing Development in Children With Hearing Loss, Dyslexia, or Oral Language Problems: Implications for Assessment and Instruction*, eds B. Arfe, J. Dockrell, and V. Berninger (Oxford: Oxford University Press), 242–269.
- Wengelin, Å., Leijten, M., and Van Waes, L. (2010). Studying reading during writing: new perspectives in research. *Read. Writ. Q.* 23, 735–742. doi: 10.1007/s11145-009-9187-5