An Exploratory Study on the Use of Speech-to-Text Technology as a Writing Modality for Pupils With Low Writing Achievement in Norwegian Lower Secondary Education

Marianne Engen Matre
University of Agder, Norway

Abstract

Six Norwegian lower secondary school pupils (ages 14–15 years) with low writing achievement participated in a stimulated recall study aimed at exploring how pupils write and experience writing with speech-to-text (STT) technology in an educational context. The study employed an exploratory design, collecting data from screen recordings and stimulated recall interviews. The screen recordings were captured while the adolescents wrote a reflective text in Norwegian, using STT and keyboard. Findings showed that the pupils were able to produce a reflective text using STT technology while experiencing both benefits and challenges due to the technology. Benefits included the opportunity to use words that they did not know how to spell and verbal skills to produce arguments in writing. Challenges were mainly related to transcription errors and technological inaccuracies. Findings suggest that technological issues need to be addressed and sufficient planning and instruction is necessary before STT can be a truly beneficial tool for adolescents with low writing achievement in secondary education.

Keywords: *writing; writing instruction; speech-to-text; special education; lower secondary education*

Introduction

Technology options for writing instruction have largely been limited to providing a choice between writing by hand and writing on a keyboard. Due to technological advances, speech-to-text (STT) technology, previously expensive and available only to pupils with documented needs, has now been integrated into writing software

Correspondence: Marianne Engen Matre, e-mail: marianne.e.matre@uia.no

from Google, Microsoft and Apple. Arcon et al. (2017) suggest that if pupils can dictate rather than transcribe, constraints pertaining to spelling and orthography would be reduced, and their texts would increase in quantity and quality. Indeed, research indicates that pupils with learning difficulties can produce higher-quality compositions when dictating texts to a scribe compared to writing by hand or typing (De La Paz & Graham, 1997).

In an exploratory study of usage patterns and perceptions of writing with STT, Ok et al. (2020) studied American pupils with high incidence disabilities in grades 4–8. Although the findings showed that pupils across all grades reported positive experiences of writing with STT, the younger pupils used it more frequently compared to older pupils. Moreover, pupils with spelling difficulties and strong oral skills tended to be more willing to use STT and used it more often. Pupils who were good spellers but had expressive language difficulties, such as speech impairments or accents, felt that STT did not aid them in writing. The teachers reported that the key benefits of using STT for pupils with high incidence disabilities were to overcome hurdles with writing tasks, to write more independently, generate more text, and provide more opportunities to write and improve pupils' confidence in writing (Ok et al., 2020). Very little research has been conducted on pupils with low writing achievement using STT as a writing modality in secondary education (Matre & Cameron, 2022). However, dissertations have been published on the use of STT, and other kinds of assistive technologies, targeting pupils with reading and writing difficulties in the Nordic countries (Kraft, 2023; Svendsen, 2016). In addition, there are promising results from research on the use of STT among pupils with learning difficulties in the United States (Ok et al., 2020; Quinlan, 2004) and Sweden (Kraft et al., 2019; Svensson et al., 2021). Yet, no studies have currently been published on the use of STT in the Norwegian educational context. Therefore, the aim of this study is to explore how pupils with low writing achievement write and experience writing with STT in a Norwegian educational context. The following research questions have guided the study:

1. How do pupils with low writing achievement approach the task of writing a reflective text using STT?
2. How do pupils with low writing achievement experience writing with STT?

Writing: Cognitive, social and technological dimensions

Aiming to describe the complexity of the cognitive processes involved in writing, MacArthur et al. (2016, p. 1) presented writing as a "complex social and cognitive process that requires shared understanding with readers about purposes and forms, knowledge of content, proficiency in language, and a range of skills and strategies, as well as motivation." That writing is considered both a social and a cognitive process has been realised through different areas of educational research, from sociocultural

studies emphasizing the communicative aspects of writing (Bazerman, 2008, 2016) to neurological or linguistic studies aiming to map and understand the underlying cognitive processes involved in writing and the development of writing-related skills (Hayes & Flower, 1980; MacArthur & Graham, 2016).

In educational research, theoretical frameworks have traditionally described either the cognitive processes underlying the development of writing skills or the social practices involved in writing. For example, the developmental model *The Simple View of Writing* by Berninger and Amtmann (2003) describes writing as three cognitive processes – transcription, self-regulation, and text generation – that are governed and constrained by working memory. Sociocultural models of writing move beyond cognitive processes and consider writing a mode of social action involving both readers and co-authors (Prior, 2006). This is particularly relevant in an educational context, as writing seldom is a solitary endeavour; rather, it is usually structured, guided and evaluated by teachers and peers. In recent theoretical models, both social and cognitive elements of writing have been included. In the revised Writer(s)-Within-Community (WWC) Model of Writing, Graham (2018, p. 258) proposes that writing is "shaped and bound by the characteristics, capacity, and variability of the communities in which it takes place and by the cognitive characteristics, capacity, and individual differences of those who produce it." The WWC model is twofold: one part describes the basic components of the writing community, while the second part shows the cognitive mechanisms involved in writing. Even though the model is presented in two separate structures, Graham (2018) underlines that the two are connected and that writing is an interaction between the writer and the writing community.

The revised WWC model (Graham, 2018) highlights the cognitive mechanisms involved in writing, from control mechanisms to long-term memory resources, the production process and modulators of the writing process. However, the model also emphasizes that cognitive mechanisms are not entirely individual traits. When writers compose, they take into consideration future readers and assessors of their texts. According to the revised WWC model, readers and assessors are referred to as the writing community that also influences text composition with elements such as the community's collective, institutional expectations, and the physical and social environment or goals. Thus, when analysing how adolescents approach and experience writing with STT in an educational context, it is important to consider both cognitive elements, such as transcription skills, attitudes towards writing and the ability to revise and reconceptualise, in addition to social factors, such as the abilities and opinions of the writing community and pupils' physical and social learning environment.

According to the Norwegian Language Curriculum (Norwegian Directorate for Education and Training, 2020), primary school pupils are expected to be able to write texts with functional handwriting and using keyboard by the end of year 4 (age 8–9), and write fluently by hand and on keyboard by the end of year 7 (age 11–12).

Pupils with low writing achievement in Norwegian lower secondary education have not sufficiently mastered these goals and may not have acquired the mechanics of writing to a point of automaticity. Brandenburd et al. (2015) argue that pupils with low writing achievement often have impaired working memory related to central executive function and the phonological loop. Having reduced working memory has been found to influence both writing fluency and text quality (Hayes & Berninger, 2014). Studies on the use of assistive technology have shown that pupils with writing difficulties may benefit from writing with STT (MacArthur & Cavalier, 2004; Quinlan, 2004).

Quinlan (2004) found that STT significantly increased the length of less fluent writers' texts and decreased the number of surface errors in their narratives. Yet, the texts written using STT were not of a significantly higher quality than the texts written by hand. Similar outcomes have been observed among children without reading difficulties. For example, Hayes and Berninger (2009) found that primary school pupils in grades 2, 4, and 6 showed an increase in the number of ideas generated as well as an enhancement of the quantity and quality of texts produced when dictating to a scribe, compared to writing texts by hand or on a keyboard. However, the approach was not as effective for older pupils who had already developed solid handwriting and transcription skills (Hayes & Berninger, 2009). These findings may be related to Bereiter and Scardamalia's (1987) developmental model of writing which describes how writers mature from basic knowledge telling, to more advanced knowledge transformation. Writing as knowledge telling is characterized by idea retrieval and retelling, while knowledge transformation includes the interaction between planning, translating and reviewing ideas to make sure that the writer's ideas come across as the author intends (Bereiter & Scardamalia, 1987; Kellogg, 2008). One of the arguments for introducing pupils with low writing achievement to STT technology, has been that the technology can reduce barriers pertaining to spelling and encoding and allow the pupils to focus on planning and reviewing ideas, resulting in more advanced writing strategies, increased fluency and improved text quality (Arcon et al., 2017; De La Paz & Graham, 1997).

Methods

The study employed an exploratory design, collecting data from screen recordings, pupil texts and stimulated recall interviews. The six pupils recruited for this study were already participating in a related research project aiming to explore STT as an inclusive approach in lower secondary education (Matre, 2022). The pupils were introduced to STT technology in January 2020 by their teachers and practiced using STT with their classmates for approximately four hours per week for 10 weeks, until the 12th of March 2020. Due to the COVID-19 pandemic and subsequent period of home schooling, the stimulated recall sessions and interviews

had to be postponed until November and December 2020, eight months after the 10-week period. The pupils reported that they had used STT to a very little degree during the home school period.

Participants

Six pupils in grades 9 and 10 (M = 14.98 years) in a rural area in Norway were invited to write a text by dictating to a computer. The pupils were allowed to type and make revisions on the keyboard, yet they were encouraged to write primarily by speech. The participants performed in the lower levels of the compulsory national reading test for grade 8, scored in the 30th percentile or lower on a standardized Norwegian spelling test (Skaathun, 2013), and were considered writers with low writing achievement based on teacher nominations. National reading test scores are presented according to five levels of mastery, where levels 1–2 are mastery below average and levels 4–5 are above average. The skill domains underlying writing and reading are closely related (Fitzgerald & Shanahan, 2000; Wengelin & Arfé, 2017), thus a group of pupils performing at the lowest mastery levels of both a standardized writing and reading test are likely to display low writing achievement. Demographic information and the sample's results on the spelling test and mastery level on the national reading tests are presented in Table 1.

Table 1. Demographic information and results on reading and writing tests

Pupil	Gender	National reading test (level)	Spelling test (percentile)	Age (y;m)	Grade	Identified learning disability
1	F	1	30th	15;5	10	General learning disability
2	M	2	5th	14;1	9	Dyslexia
3	F	2	30th	14;11	10	Under assessment for dyslexia
4	F	2	30th	14;3	9	Under assessment for dyslexia
5	M	3	20th	15;7	10	No
6	M	2	30th	15;8	10	No

Data collection

Data collection was conducted in two parts. Part one consisted of a screen-recorded writing session, and part two comprised individual stimulated recall interviews. The use of more than one data collection method, also known as methodological triangulation (Noble & Heale, 2019) was employed to enrich and validate findings. The pupils were divided into two groups and situated in a small classroom with desks placed in each corner. Two stimulated recall interviews started immediately after the writing sessions, while the remaining four were conducted consecutively within two hours of the writing session. Both the writing sessions and interviews took place during school hours.

Screen Recordings

In part 1, the pupils were given five minutes to plan and 15 minutes to write a reflective text in Norwegian using STT and keyboard. The pupils were encouraged to write using STT but were allowed to use the touchpad (mouse) and keyboard. They were also provided with noise cancelling headphones and used the STT software integrated into Microsoft Office Word 2019, which enabled screen recording in Microsoft Office PowerPoint 2019. The STT software had been available in Norwegian in Microsoft Office for approximately one year at the time of data collection. According to Yu and Deng (2015, p. 1), STT relies on building models from big data collected from real usage scenarios to make a system robust. The learning algorithms underlying Microsoft's STT were trained on a universal language model and adapted to Norwegian using the Norwegian Language Bank's[1] dataset of speech and text. Compared to STT building on larger language corpora of English, Chinese or Spanish speech and text, the Norwegian dataset is significantly smaller and, thus, prone to produce more recognition errors.

The writing task

The topic of the writing task was social media's influence on adolescents. The pupils were provided with the following prompt: "Do you think social media affects how adolescents behave? Reflect and argue your opinion." Reasoning and arguing in reflective texts are part of the Norwegian lower secondary school curriculum (Norwegian Directorate for Education and Training, 2020). The influence of social media was considered a topic well known to the pupils and suitable for reflection. It is also central to the latest Norwegian Core Curriculum implemented in the autumn of 2020, which emphasises health and life skills as one of three interdisciplinary topics (Norwegian Directorate for Education and Training, 2020).

Stimulated recall interviews

Stimulated recall is an approach where the researcher presents authentic stimuli to research participants to acquire thoughts and experiences on an original situation (Vesterinen et al., 2010). The authentic stimuli were screen recordings of adolescents writing a reflective text using STT and keyboard. Video stimulated recall has been frequently used to explore how pupils or teachers experience specific events in education (Lyle, 2002; van der Kleij, 2021). It is a data collection approach related to the verbal protocol approach, where the researcher encourages the subject to think-aloud during an activity to provide insight into cognitive processes. The verbal protocol approach has been applied in writing research (Hayes & Flower, 1981; Swain, 2006), and in research on writing technology for pupils with low writing achievement

---

[1] An open-source dataset provided by the Norwegian National Library: https://www.nb.no/sprakbanken/en/sprakbanken/

(Svendsen, 2016). In contrast to the verbal protocol approach, video stimulated recall allows the pupils to complete their task before they are encouraged to analyse and elaborate on their experiences.

To prompt recall of the situation, Lyle (2002) recommends that interviews should be conducted as soon as possible after recorded sessions. In this study, the pupils were presented with the recording within hours of writing and encouraged to describe their experiences. During the stimulated recall interviews the pupils were instructed to describe how they experienced writing with STT. Both the researcher and the pupils were able to pause the screen recording whenever they wanted to ask a question or comment. To prompt recall, the researcher asked open-ended questions, such as "What happened here?" or "Why did you stop here?" After watching the screen recording of the writing session, the pupils were asked questions about how they experienced using STT for that specific writing task. They were also asked to describe challenges or advantages of STT versus typing or writing by hand. The interviews lasted between 21 and 39 minutes.

Analyses

Two kinds of analyses were conducted: (1) analyses of screen recordings and (2) analyses of stimulated recall interview transcripts. To be able to explore how the pupils wrote with STT (research question 1), variables that describe text production (e.g., words produced with STT, words typed, words removed, words per minute and accuracy) were registered from the screen recordings. See Table 2 for an operationalization of the variables describing the pupil's text production with STT and keyboard. Only recordings of the 15 minutes of text production were analysed; thus, the five minutes of planning time were not included in the analyses. Measures from the final texts were also analysed, including variables such as final accuracy and final word count. Frequencies were registered by the author and a research assistant. To determine coding consistency, inter-rater reliability was calculated at 0.87 using Cohen's kappa (Carletta, 1996).

Table 2. Operationalization of measures of text production and interrater reliability

Measures	Operationalization	Inter-rater reliability (Cohen's kappa)
Words produced with STT	Number of words transcribed by the STT technology, including words that were deleted	0.92
Total number of words produced	Number of words produced, either with STT or keyboard, including words that were deleted	0.86
Typing-STT ratio	Number of words typed on a keyboard divided by the number of words produced with STT, including words that were deleted	0.70

Measures	Operationalization	Inter-rater reliability (Cohen's kappa)
Words produced with STT per minute	Number of words produced with STT divided by 15 (minutes)	0.88
Total words per minute	Total number of words produced divided by 15 (minutes)	0.86
Words removed	Number of words removed. If the pupil removed one or several letters but not the entire word, it was still counted as one word removed.	0.76
Switches between STT and keyboard	Number of times the pupil switches from keyboard to STT or from STT to keyboard	0.91
Accuracy in text produced with STT	Words produced with STT minus number of words incorrectly transcribed by the STT divided by number of words produced with STT	0.89
Final accuracy	Final word count minus number of errors in the submitted text divided by final word count	0.94
Final word count	Number of words included in the submitted version of the text	1.0

To be able to analyse the pupils' experiences of writing with STT (research question 2), screen recordings and stimulated recall interviews were transcribed and coded using NVivo 12. Three main categories were identified: (1) benefits, (2) challenges, and (3) emotional reactions. The categories emerged through analyses of prominent responses from the stimulated recall interviews and elements considered to influence text production (e.g., interruptions, switches between STT and keyboard or revision strategies) from the screen recordings.

Ethical considerations

The study follows guidelines provided by the Norwegian Centre for Research Data (NSD) and has received approval from the NSD to collect and store data. All participants gave written and oral consent to take part in the study. As the pupils were 14–15 years old, and their parents also provided consent for their participation. Participation was voluntary, and pseudonyms (Pupils 1–6) are used in place of pupils' names to provide anonymity.

Results

Screen recording results
Figure 1 describes how the pupils wrote with STT, from lowest to highest number of words produced with STT and the relation to final word count, words typed, words removed and the number of switches between keyboard and STT.
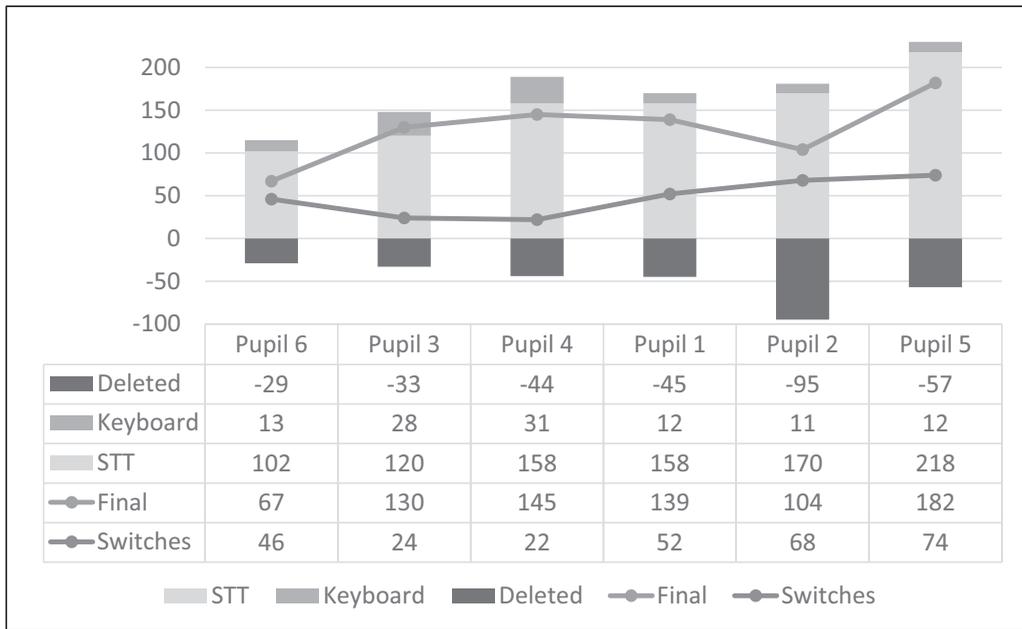
	Pupil 6	Pupil 3	Pupil 4	Pupil 1	Pupil 2	Pupil 5
Deleted	-29	-33	-44	-45	-95	-57
Keyboard	13	28	31	12	11	12
STT	102	120	158	158	170	218
Final	67	130	145	139	104	182
Switches	46	24	22	52	68	74

Figure 1. Words produced with stt from lowest to highest and their relation to final word count, words typed, words removed and switches between keyboard and STT.

All the pupils typed words on the keyboard but produced more words by STT than by typing. The pupils used different approaches to writing with STT; for example, some pupils had fewer switches between keyboard and STT, while others had more frequent transitions. The ratio between words typed and words dictated ranged from 1:5 (Pupil 3) to 1:20 (Pupil 5). There was also variation in fluency and text length between the pupils. Fluency, which was measured in words per minute, ranged from 6.8 (Pupil 6) to 14.5 (Pupil 5) words per minute using STT and from 7.6 (Pupil 6) to 15.3 (Pupil 6) words per minute when typing on a keyboard. Pupil 5 produced the longest text and the highest number of words with STT. Pupil 2, who had dyslexia, produced the second-highest number of words with STT but submitted one of the shortest texts. This is explained by the number of words removed, as Pupil 2 deleted 64% of his text. Except for Pupil 2, there appears to be a tendency for the pupils who produced the most text with STT to submit the longest texts.

Accuracy ranged from 70% (Pupils 3 and 6) to 85% (Pupil 1) when comparing what the pupils said to what the STT transcribed. The accuracy of the final texts ranged from 92% to 97%, except for Pupil 6, who had 75% accuracy. Pupil 6 explained during interviews that he did not revise the accuracy errors on purpose, as he wanted the researchers to see the mistakes the STT technology had made. Figure 1 shows that the pupils were required to make several revisions and, therefore, also switched between keyboard and STT to produce a more accurate text.
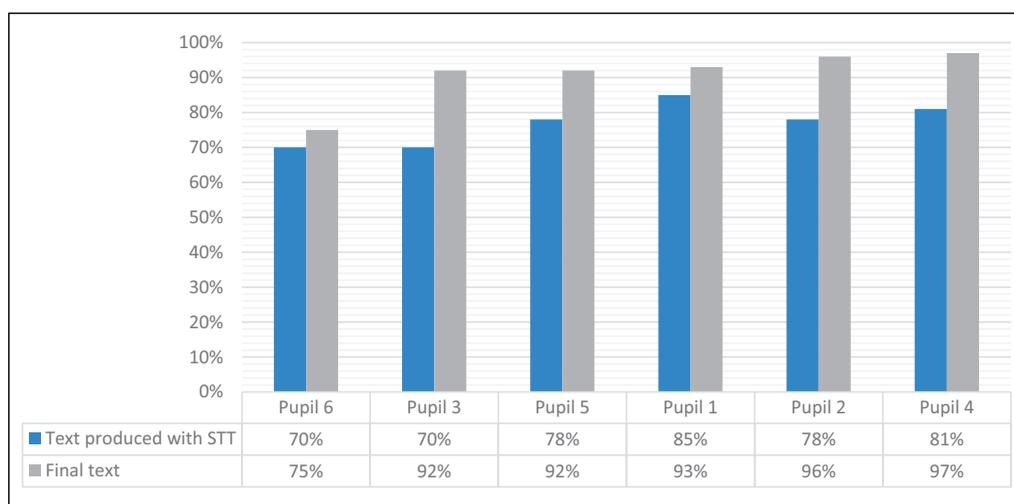
	Pupil 6	Pupil 3	Pupil 5	Pupil 1	Pupil 2	Pupil 4
■ Text produced with STT	70%	70%	78%	85%	78%	81%
■ Final text	75%	92%	92%	93%	96%	97%

Figure 2. Accuracy of text produced with stt and accuracy in final text, sorted from least to most accurate final text.

An overview of the types of errors recorded is presented in Table 3. The most frequently occurring error (69.1% of categorised errors) was labelled as a transcription error. An error was considered a transcription error if the STT technology produced a similar-sounding word, a misspelled word, a word from another language or if it added or removed a word spoken by the pupil. According to the analyses of screen recordings, all pupils experienced transcription errors, ranging from 23–36 errors for each pupil.

Table 3. Number and percent of errors observed and the range and number of pupils who experienced each type of error.

	Number of times error observed	Percent of total errors (n = 259)	Number of pupils (n = 6)	Range across pupils (min–max)
Transcription errors	179	69.1%	6	23–36
Produces text without intending to dictate	28	10.8%	4	0–12
Erroneous capitalization	19	7.3%	6	1–9
Speech not registered by STT	13	5.0%	5	0–5
Dictates while STT is switched off	11	4.2%	5	0–5
STT registers speech from other pupils	9	3.5%	3	0–4

Some of the pupils produced text with STT without the intention of dictating; they were reading their text out loud, thinking out loud or commenting on something while STT was activated. There were occurrences where sounds in the task environment, such as sighing, heavy breathing or other pupils dictating, were picked up by STT and transcribed into text. The most frequent transcription of heavy breathing and sighing was [hm…]. Three pupils experienced that STT picked up and transcribed

something that other pupils were saying. Five pupils spoke with the intention to write, yet the STT technology did not respond. At times, this was caused by a technical error that was solved when the pupils turned STT off and on again. On other occasions, pupils attempted to dictate; however, STT had been automatically turned off and, therefore, did not provide any transcriptions. This happened because the software default setting is for dictation to turn off if a pause lasts more than 20 seconds. Six pupils experienced that STT did not produce any text, even though they spoke and STT was activated. Pupils 2 and 5 experienced this five times, while the other four pupils experienced it once.

Table 4 presents different subgroups of transcription errors observed and the range and number of pupils who experienced each type of error.

Table 4. Number and percent of transcription errors observed and the range and number of pupils who experienced each type of error.

	Number of times error observed	Percent of total transcription errors (n = 179)	Number of pupils (n = 6)	Range across pupils (min–max)
Semantic errors	114	63.1%	6	14–23
Homophone errors	15	8.9%	6	2–3
STT adds words that were not dictated	24	13.4%	6	3–6
STT suggests text in another language	11	6.1%	3	0–5
Spelling errors	9	5.0%	4	0–3
STT does not transcribe words that were dictated	6	3.3%	3	0–3

Semantic errors were the most frequently occurring type of transcription error. These kinds of errors were usually words that sounded somewhat like the word the pupil had pronounced. Sixty-three percent of the transcription errors were semantic errors. STT also produced homophone errors, a word that has a different meaning yet the exact same pronunciation. The most frequently occurring homophone error was the transcription of the Norwegian conjunction 'and' (transcribed 'og'), which is a homophone of the Norwegian infinitive marker (transcribed 'å'). Both 'å' and 'og' are pronounced /ɔ/ when unstressed; thus, accurate spelling presupposes semantic knowledge. This is similar to the homophones 'to' and 'too' in English. The STT software suggested 'å' for 'og' and vice versa. As these errors were not spelling errors but semantic errors, the homophones were not marked by spellcheck. All pupils made two or three homophone errors while writing with STT. Pupils 5 and 6 revised all homophone errors, while the other pupils revised one or two errors, leaving one error in the final text. Another example of transcription errors was the emergence of similar-sounding English words or phrases that were transcribed in English, even though the pupils spoke in Norwegian, and Norwegian was set as dictation language. Four pupils experienced that STT transcribed English or German words. While this

happened, on average, 3.25 times per pupil (ranging from zero to five times), most pupils noticed these errors and deleted them during revision.

Interview results

Perceived benefits

Regarding the main benefits of using STT, the pupils reported that it was exciting to try something new; STT helped with spelling; and it was easier to elaborate when producing text orally. Some pupils said that they experienced text production with STT as faster than typing on a keyboard. One example of spelling assistance was provided by Pupil 2, who wanted to write about a documentary on social media: "So, I was thinking. How do I spell 'documentary'? And then I said it out loud and it just appeared. It [STT] was useful for those difficult words that I don't really know how to spell."

Several pupils stated that it was easier to argue and elaborate when using dictation. One pupil stated, "It is easier to explain something when you speak compared to when you have to put it down in writing." (Pupil 1). Other pupils described similar experiences, emphasising that they were allowed to think aloud and focus on the content that they were trying to convey without having to consider spelling and syntax. Pupil 6 stated that a benefit of writing with STT was that he was able to produce more text in a shorter amount of time. He stated, "I felt like I was thinking faster. Or… I was thinking at the same speed, but when I wrote with STT, more text appeared." The pupils described that it was easy to write with STT when the technology was accurate. All six pupils were able to produce a reflective text by speech, even though they reported that they had not used STT regularly since the initial training period.

Perceived challenges

The main challenges relate to the use of dialects, transcription errors and disruptions due to revisions. Pupils 4 and 2 stated that they had to alter their pronunciation to a more standard dialect to be successful when dictating. Pupil 4 expressed that it was embarrassing to speak out loud when she had to enunciate each word, while Pupil 2 said, "I lost focus when I had to say everything clearly and correctly. I noticed that I had to change my dialect to make it able to pick up what I meant. I had to pronounce everything with a posh dialect and that was very distracting."

Pupil 6 explained that one of the challenges with STT was that he had to plan and speak at the same time. He said, "When I dictate, the words come straight out of my mouth before I get the chance to think them through." As he watched himself revise his text, Pupil 6 explained that he had to delete more text when writing with STT because he did not have the opportunity to formulate sentences and "test them out" while he was speaking. Others described similar experiences. For example, they had to delete text that they had dictated because they considered it "too oral." Two pupils had identified learning disabilities – Pupil 1 had a general learning disability, and

Pupil 2 had dyslexia. Comparing Pupils 1 and 2 to the other pupils, the pupils with identified learning disabilities became notably more frustrated during the writing process when they encountered accuracy errors. Pupil 2 stated that he experienced the writing process as less efficient because a lot of time was spent on revision. He explained, "I got distracted when editing, and then I could not remember what I originally planned to write."

Emotional reactions

Some of the pupils stated that they found it embarrassing to dictate because other pupils could listen to what they were saying. Pupil 4 noted that if she had been in a room with only her closest friends, it would not have been embarrassing to use STT. She added that the worst part was that she was sharing her text as it was being produced, not the final version. The pupils reacted differently to the challenges that emerged with STT. Some met the challenges without noticeable reactions, while others were frustrated, and some found it amusing. Pupils 1 and 2 expressed frustration when they experienced transcription errors or technical difficulties. When asked to describe this experience, Pupil 2 explained that it was annoying because he knew he could just type, and he lost focus trying to write by speech. Pupils 1 and 4 giggled when the STT suggested something entirely different from what they had intended to write. For example, the English abbreviation 'omg' (oh my god) was suggested when one of the pupils dictated a phrase with similar sounding phonemes in Norwegian.

Discussion

The aim of this study was to explore how pupils with low writing achievement approach and experience writing with STT. Findings from screen recordings and stimulated recall interviews showed that pupils with low writing achievement were able to produce reflective texts with STT in Norwegian. However, they experienced both benefits and challenges caused by the technology. The main benefits relate to spelling, being able to elaborate arguments orally and the excitement of trying something new. The challenges mainly pertain to disruption due to transcription errors, the need to revise the text by hand and the embarrassment of speaking out loud.

Text production with STT

The pupils produced texts with STT, and described that they were relieved of some challenges related to spelling when composing with STT. These findings are in line with studies by Ok et al. (2020) and Nordström et al. (2019), who found that STT was especially helpful for pupils who struggled with spelling. However, pupils in the current study could not rely on STT to be 100% accurate and provide correct orthography and syntax in Norwegian. Thus, this study does not entirely support the hypothesis by De La Paz and Graham (1997) stating that STT allows pupils to

spend less effort on lower-order skills and enables them to devote more attention to higher-order skills, such as planning content, creating a good structure, and text coherence. It is important to emphasize that this study was conducted with recently integrated STT technology, and transcription errors are likely to be reduced as the technology is further adapted to Norwegian users. When Norwegian language users provide input in different dialects and correct transcription errors, the corpora will grow, and the learning algorithms will be able to produce more accurate output. For example, transcription errors in which STT produces text in languages other than the preferred setting are likely to disappear as Norwegian text corpora increase. This is likely to cause pupils less disruption and reduce the need for proofreading and editing.

Challenges such as STT being automatically turned off and recognition errors are also likely to be reduced as the technology improves. Algorithms underlying STT are already trained to distinguish sounds in the task environment, such as coughing and heavy breathing, from speech input intended for transcription (Yu & Deng, 2016). However, these algorithms must be trained on larger language corpora of specific languages to improve accuracy. As the accuracy of STT improves, the writing experience of pupils with low writing achievement is also likely to improve. Regardless of improved technology, pupils with low writing achievement may still experience a lack of control when writing with STT. The pupils in this study described how STT produced text at a higher pace compared to writing by hand and typing on a keyboard. Research on handwriting and typing shows that compositional fluency correlates with text quality (Troia et al., 2020), yet the experience of higher fluency was not considered a benefit by some of the pupils with low writing achievement in this study. This issue is not likely to be resolved as technology improves.

A challenge with homophone errors is that they are correctly spelled words found in the dictionary; thus, the spell check does not mark them as errors. Some of the pupils did not recognise the homophone errors and did not correct them. Reading, planning, text evaluation and revision are central elements in the writing process and expert writers spend more time editing before considering a text a final product (MacArthur, 2016). Pupils with low writing achievement are also likely to struggle with reading (Wengelin & Arfé, 2017), and even though STT may reduce constraints relating to encoding, as long as the technology is not entirely accurate, pupils still have to decode their texts to check for errors in transcribed text.

Experiences

The pupils described that it was exciting to try a different approach to writing and easier to elaborate and produce arguments orally compared to when they were typing or writing by hand. However, their emotional reactions of embarrassment, frustration, and amusement epitomise how STT also caused disruption to the writing process. The environmental factors of writing, as described in the WWC model

(Graham, 2018), appeared to influence how pupils experienced writing with STT. They had to expose their opinions, and although they were all composing at the same time, they clearly considered their peers' reactions to what they were writing. Thus, providing a safe learning environment is especially important when pupils in lower secondary education produce texts with STT. It should be noted that writing in educational contexts may be distinctive from writing in other contexts. As the pupils stated that it would have been less embarrassing to produce text with STT with only close friends, it is recommended to introduce and practice using the technology during low-stakes activities in a safe learning environment where the pupils feel secure speaking out loud and are less concerned with producing texts of high quality.

The two pupils with identified learning disabilities expressed frustration when they encountered accuracy errors. The pupil with dyslexia stated that it was frustrating to use STT because he could just type what he wanted to write instead. However, according to the measures of the pupil's spelling ability (see Table 1), he was also likely to make spelling mistakes when typing. The fact that some pupils experience more frustration when writing with STT is supported by previous research (e.g., Ok et al., 2021). Personality traits and emotions are central elements in writing, according to the WWC model (Graham, 2018). Thus, it is worth noting that even as STT technology improves, pupils with low self-efficacy towards writing are likely to have less patience and perseverance when testing out new writing approaches. They may also have more to lose if STT does not provide sufficient assistance compared to their peers, who, to a larger degree, master writing by hand or typing. Thus, as is important with all assistive technologies, teaching professionals should consider and evaluate individual benefits and constraints (Edyburn, 2006) when introducing pupils to STT.

Implications for practice and future research

For pupils who have the prerequisites to understand how STT works, it is useful to explain in simple terms why the technology will benefit from clear and continuous dictation.

Continuous dictation is more likely to improve than word-to-word dictation, as speech input algorithms perform better when they receive more input and can learn from the context (Yu & Deng, 2016). Consequently, when a STT writer pauses between words, the software is less able to use the surrounding words as contextual indicators and prediction accuracy is reduced. Pupils should therefore be advised to dictate continuously, even as STT technology improves.

Furthermore, as algorithms learn from feedback, pupils can be encouraged to correct transcription errors to improve the technology. STT alone cannot solve every challenge related to text production for pupils with low writing achievement. Even if accuracy improves and STT becomes an efficient transcription aid, it is equally important to provide pupils with writing strategies and capabilities to approach other

aspects of writing, such as ideation, planning, reviewing, and revising while using STT. It is also important to provide instruction on how to write using STT. Compared to handwriting, to which a lot of time and practice are devoted in elementary education (Fancher et al., 2018), research on typing shows that few teachers provide keyboarding instructions (Poole & Preciado, 2016) because pupils are thought to acquire typing skills without formal instruction (Grabowski, 2008). This study indicates that it is important to provide instruction on how to utilize STT, as pupils with low writing achievement experience difficulties when using this writing modality.

Due to school closures to prevent the spread of the COVID-19 virus, the pupils had only a few weeks to practice writing with STT. This was not enough time to master a new writing approach. It should be noted that STT is likely to be more accurate in widespread languages such as English, French, Spanish or Chinese (McCrocklin et al., 2019; Yu & Deng, 2016). Thus, using STT and a keyboard during writing exercises in foreign language learning, may be beneficial for pupils with low writing achievement. Further research on STT should include longitudinal studies to explore the benefits and challenges of using STT in Norwegian and foreign languages over time for pupils with low writing achievement.

Limitations of the study

The pupils in this study were introduced to STT for approximately four hours per week for a 10-week period. However, this study was conducted eight months after this introduction period, as schools were closed due to the COVID-19 pandemic. During the eight months between the introduction period and the stimulated recall writing session, the pupils reported that they had not used STT on a regular basis. Some of the challenges reported in the results, such as STT turning off automatically, may be explained by a lack of practice and familiarity with the technology. The software provided had only been available in Norwegian in Microsoft Office for approximately one year at the time of data collection. The high number and variety of recognition errors may be caused by software building on a relatively small language corpus.

When analysing the stimulated recall interviews, it became apparent that the pupils, to a larger degree, described the content of their texts and challenges with the technology rather than their experience of writing with a different modality. Future research on adolescents with low writing achievement should consider that the stimulated recall procedure sets high demands on the informants' metalinguistic knowledge and ability to reflect on their own writing process. It should also be noted that this study contains a small sample of six pupils using rural dialects to produce a reflective text with STT in the southern part of Norway. The sample size and design of the study do not provide an opportunity to generalise findings to larger populations of pupils with low writing achievement. Yet, some of the findings may be applicable to pupils using STT in other Nordic languages as well, as the STT technology is likely to be

as inaccurate in other smaller languages compared to widespread languages such as English or Chinese.

Concluding remarks

This study indicates that at present STT is not yet the ideal writing technology for pupils in Norwegian lower secondary education with low writing achievement. The pupils experienced transcription errors and technological inaccuracies, and elements of using STT that may have been considered a beneficial also caused the participants some distress. For example, the process of speaking instead of typing – which is faster and was expected to lead to more fluent production – was not considered a benefit by all. However, there are some recognizable benefits, and STT may become an efficient writing modality as Norwegian speech recognition technology develops. The findings underline the importance of providing sufficient instruction and a safe learning environment if pupils with low writing achievement are to exploit the technology's potential in an educational context.

Author biography

Marianne Engen Matre is a Ph.D. candidate and Assistant Professor at the Department of Education, University of Agder. Her Ph.D. project explores the use of speech-to-text technology as a writing modality in lower secondary education.

References

Arcon, N., Klein, P. D., & Dombroski, J. D. (2017). Effects of dictation, speech to text, and handwriting on the written composition of elementary school English language learners. *Reading & Writing Quarterly*, *33*(6), 533–548. https://doi.org/10.1080/10573569.2016.1253513

Bazerman, C. (Ed.). (2008). *Handbook of research on writing*. Erlbaum.

Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 11–23). Guilford Press.

Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). The Guilford Press.

Brandenburg, J., Klesczewski, J., Fischbach, A., Schuchardt, K., Büttner, G., & Hasselhorn, M. (2015). Working memory in children with learning disabilities in reading versus spelling: Searching for overlapping and specific cognitive factors. *Journal of Learning Disabilities, 48*(6), 622–634. https://doi.org/10.1177/0022219414521665

Carletta, J. (1996). Assessing agreement on classification statistics: The kappa statistic. *Computational Linguistics, 22*(2), 249–254.

De La Paz, S., & Graham, S. (1997). Effects of dictation and advanced planning instruction on the composing of students with writing and learning problems. *Journal of Educational Psychology, 89*(2), 203. https://doi.org/10.1037/0022-0663.89.2.203

Edyburn, D. L. (2006). Assistive technology and mild disabilities. *Special Education Technology Practice*, *8*(4), 18–28.

Fancher, L. A., Priestley-Hopkins, D. A., & Jeffries, L. M. (2018). Handwriting acquisition and intervention: A systematic review. *Journal of Occupational Therapy, Schools, & Early Intervention*, *11*(4), 454–473. https://doi.org/10.1080/19411243.2018.1534634

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, *35*(1), 39–50. https://doi.org/10.1207/S15326985EP3501_5

Grabowski, J. (2008). The internal structure of university student's keyboard skills. *Journal of writing research*, *1*(1), 27–52. https://doi.org/10.17239/jowr-2008.01.01.2

Graham, S. (2018). A revised writer (s)-within-community model of writing. *Educational Psychologist*, *53*(4), 258–279. https://doi.org/10.1080/00461520.2018.1481406

Hayes, J. R., & Berninger, V. W. (2009). 13 relationships between idea generation and transcription. *Traditions of Writing Research*, 166.

Hayes, J. R., & Berninger, V. W. (2014). Cognitive processes in writing: A framework. In B. Arfé, J. Dockrell, & V. Berninger (Eds.), *Writing development in children with hearing loss, dyslexia, or oral language problems: Implications for assessment and instruction* (pp. 3–15). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199827282.003.0001

Hayes, J. R., & Flower, L. (1981). *Uncovering cognitive processes in writing: An introduction to protocol analysis*. ERIC Clearinghouse.

Kraft, S. (2023). Speaking the text. When children with reading and writing difficulties compose with speech-to-text. University of Gothenburg. http://hdl.handle.net/2077/76015

Kraft, S., Thurfjell, F., Rack, J., & Wengelin, Å. (2019). Lexikala analyser av muntlig, tangentbordsskriven och dikterad text producerad av barn med stavningssvårigheter. *Nordic Journal of Literacy Research*, *5*(3) https://doi.org/10.23865/njlr.v5.1511

Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, *29*(6), 861–878.

MacArthur, C. A. (2016). Instruction in evaluation and revision. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 272–287). Guilford Press.

MacArthur, C. A., & Cavalier, A. R. (2004). Dictation and speech recognition technology as test accommodations. *Exceptional Children*, *71*(1), 43–58. https://doi.org/10.1177/001440290407100103

MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.). *Handbook of writing research* (pp. 24–40). Guilford Press.

Matre, M. E. (2022). Speech-to-text technology as an inclusive approach: Lower secondary teachers' experiences. *Nordisk tidsskrift for pedagogikk og kritikk*, *8*.

Matre, M. E., & Cameron, D. L. (2022). A scoping review on the use of speech-to-text technology for adolescents with learning difficulties in secondary education. *Disability and Rehabilitation: Assistive Technology*, 1–14. https://doi.org/10.1080/17483107.2022.2149865

McCrocklin, S. & Humaidan, A., & Edalatishams, I. (2018). Asr dictation program accuracy: Have current programs improved? *Pronunciation in Second Language Learning and Teaching Proceedings, 10*(1).

Noble, H., & Heale, R. (2019). Triangulation in research, with examples. *Evidence-Based Nursing*, *22*(3), 67–68.

Nordström, T., Nilsson, S., Gustafson, S., & Svensson, I. (2019). Assistive technology applications for students with reading difficulties: Special education teachers' experiences and perceptions. *Disability and Rehabilitation: Assistive Technology, 14*(8) 798–808. https://doi.org/10.1080/17483107.2018.1499142

Norwegian Directorate for Education and Training. (2020). *Norwegian subject curriculum – Competence aims after year 10*. https://www.udir.no/lk20/nor01-06/kompetansemaal-og-vurdering/kv111?lang=eng

Norwegian Directorate for Education and Training. (2020). *Core curriculum – Health and life skills*. https://www.udir.no/lk20/overordnet-del/prinsipper-for-laring-utvikling-og-danning/tverrfaglige-temaer/folkehelse-og-livsmestring/?lang=eng&TilknyttedeKompetansemaal=true

Ok, M. W., Rao, K., Pennington, J., & Ulloa, P. R. (2022). Speech recognition technology for writing: Usage patterns and perceptions of students with high incidence disabilities. *Journal of Special Education Technology*, *37*(2), 191–202. https://doi.org/10.1177/0162643420979929

Poole, D. M., & Preciado, M. K. (2016). Touch typing instruction: Elementary teachers' beliefs and practices. *Computers & Education*, *102*, 1–14. https://doi.org/10.1016/j.compedu.2016.06.008

Prior, P. (2006). A sociocultural theory of writing. *Handbook of Writing Research*, 54–66.

Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: Improving fluency. *Journal of Educational Psychology*, *96*(2), 337. https://doi.org/10.1037/0022-0663.96.2.337

Skaathun, A. (2013). *The reading test by the Norwegian reading centre* [Lesesenterets staveprøve]. University of Stavanger.

Skar, G. B., & Aasen, A. J. (2021). School writing in Norway: Fifteen years with writing as key competence. In J. Jeffery, & J. M. Parr (Eds.), *International perspectives on writing curricula and development* (ch. 3). Routledge. https://doi.org/10.4324/9781003051404-10

Svendsen, H. B. (2016). *Technology-based reading and writing strategies in Danish general education* [Doctoral dissertation]. Aarhus University.

Svensson, I., Nordström, T., Lindeblad, E., Gustafson, S., Björn, M., Sand, C., Almgren, G. B., & Nilsson, S. (2021). Effects of assistive technology for students with reading and writing disabilities. *Disability and Rehabilitation: Assistive Technology*, *16*(2), 196–208. https://doi.org/10.1080/17483107.2019.1646821

Swain, M. (2006). Verbal protocols. *Inference and Generalizability in Applied Linguistics: Multiple perspectives*, *12*, 97.

Troia, G. A., Brehmer, J. S., Glause, K., Reichmuth, H. L., & Lawrence, F. (2020). Direct and indirect effects of literacy skills and writing fluency on writing quality across three genres. *Education Sciences*, *10*(11), 297. https://doi.org/10.3390/educsci10110297

University of Stavanger. (2023). *Writing technology for pupils with dyslexia*. https://www.uis.no/nb/forskning/skriveteknologi-for-elever-med-dysleksi

van der Kleij, F. (2020). Teacher and student perceptions of oral classroom feedback practices: A video-stimulated recall study. *The Australian Educational Researcher, 50*, 353–370. https://doi.org/10.1007/s13384-021-00478-0

Vesterinen, O., Toom, A., & Patrikainen, S. (2010). The stimulated recall method and ICTs in research on the reasoning of teachers. *International Journal of Research & Method in Education*, *33*(2), 183–197. https://doi.org/10.1080/1743727X.2010.484605

Wengelin, Å., & Arfé, B. (2017). The complementary relationships between reading and writing in children with and without writing difficulties. In B. Miller, P. McCardle, & V. Connelly (Eds.), *Writing development in struggling learners* (pp. 29–50). Brill. https://doi.org/10.1163/9789004346369_004

Yu, D., & Deng, L. (2016). *Automatic speech recognition. A deep learning approach*. Springer.