The effects of voice recognition systems on writing styles: Implications for use in the education environment

Introduction

When commercially available voice recognition systems (speech to text translation) were introduced into Australia in the late 1980s, their prohibitively high cost and relatively poor ability to recognise speech made them effective only for users with disabilities. However, with recent advances in voice modelling techniques and a dramatic decrease in cost, voice recognition systems are reaching the point where their use is becoming feasible for all computer operators. Recent communications with the distributors of Dragon Systems "DragonDictate", the major voice recognition system in Australia, found that in the past 12 months, approximately 300 systems had been sold in the state of Queensland alone (HSM Consultants, pers. commun. 1996). About 80% of systems are now supplied to users who are not disabled and who choose to use voice recognition systems because it facilitates easy text entry. It is highly likely that the number of non-disabled users will rapidly increase as the cost of the technology continues to decrease and the quality of the voice recognition improves.

The possible use of voice recognition technology in the educational environment has been examined by Wetzel (1991a,b). In this study voice input was suggested to be of use for the transcription of thoughts and ideas from primary students, taking advantage of the far superior levels of verbal, compared to written fluency of students at this level. Voice recognition has also been suggested to aid in the teaching of pronunciation in early reading instruction (Kantrov, 1991). In the secondary school environment, voice recognition systems enable students to become more productive and self-sufficient and to create polished, accurately spelled documents in shorter periods of time (Wetzel, 1991a). Using voice recognition systems, students with severe disabilities who were previously unable to write, are now able to do so and therefore participate more independently in the regular classroom (Walker, 1990).

Incorporating voice recognition technology into schools

Wetzel (1991a) identified five areas that he believed needed to be addressed before voice recognition technology could be fully integrated into schools. These were the need for voice recognition systems to:

- Recognise speech from a variety of people, sexes and nationalities;
- Operate with only minimal need for training of users;
- Recognise continuous speech input without the need for pauses in dictation;
- More accurately recognise speech input; and
- Effectively screen out background noise.

In the six years since Wetzel’s study, voice recognition technology has improved significantly. However the need for speaker independent systems requiring only minimal user training still represents the major barrier to effective classroom usage of the technology. In the last versions of voice recognition systems, the pause time required between words has...
been reduced and is now somewhat less than 0.1 of a second. In terms of accuracy of recognition and the ability to screen out background noise, both are reasonably well catered for in current systems, and are constantly being improved as updated versions are released. With only minimal training, most users can approach 90-95% accuracy after short periods of continuous use. Built-in noise filters now give systems the ability to adjust to changes in background noise levels, making it possible to effectively use voice recognition systems in the school classroom.

These advances in voice recognition technology suggest that sometime in the future, the use of voice input may become common in our schools. Given this possibility, this study considered the following questions:

- How does the change to voice input affect writing styles? and
- Are the cognitive processes required to write using voice recognition systems different from those using pens and pencils or those using a conventional keyboard?

**Writing styles: Survey results and discussion**

In order to assess the effects of a change to voice input of text, this study surveyed users of the DragonDictate voice recognition system about their writing styles. Users of the DragonDictate system (Dragon Dictate 2 - Classic Edition, 1993-1995) were identified following communication with the system suppliers and sent a letter containing an explanatory note, a written survey and a reply paid envelope.

The survey focused on four main issues relating to the use of voice recognition systems:

- **User characteristics and speed of writing:**
- **Changes in the technical aspects of written English and in writing style:**
- **Editing using voice recognition:** and
- **Motivation**

The large majority of the survey questions were multiple choice in format with opportunities given for individual comments.

In the survey, no specifications were made about the specific genres or styles of writing to be considered by respondents, or whether writing was to be for work or for pleasure. Assessments of changes in the characteristics of writing are based on self-report data and written texts from individual respondents were not examined. Respondents were simply asked to compare different aspects of their writing styles prior to and while using voice recognition technology.

**“A lthough using voice recognition is slower than using a dictating machine, the document is ready for immediate revision and sending”**

Twenty-four users of the DragonDictate voice recognition system were sent the survey, of which 14 were returned. The limited number of respondents means that the results discussed here are preliminary only and therefore should be used only as a guide for future more detailed study.

**User characteristics and speed of writing**

The length of time respondents had been using their system ranged from less than 3 months to over 5 years. Most users reported that they used voice input at least several times a week. The principal use of voice recognition systems was to compose text for business or personal communications. Other uses included transcription of dictaphone-prepared text, data entry and control computer functions.

Ten of fourteen respondents were also able to use a conventional keyboard to enter text, but chose to use voice recognition systems because they preferred talking to typing. Four of the respondents were people with disabilities who found using a conventional keyboard difficult, while one respondent was evaluating the system for use by students with disabilities.

Voice input was found to be a quicker method of text entry for nine of the fourteen, although this was dependent on the length of time the voice recognition system had been in use. The average speed of text entry using voice recognition was reported to be approximately 20 words per minute (w.p.m.), with more experienced users reporting the higher speeds. The minimum reported speed of text entry was less than 5 w.p.m. and the maximum was greater than 30 w.p.m. The slowest speed of text entry was reported by a respondent with a severe physical disability. For this person, a 5 w.p.m. speed of text entry represented a substantial improvement in her ability to communicate.

All respondents composed text directly onto the computer using voice recognition and seven of thirteen found this to be quicker than conventional writing. Two of those who reported a decrease in writing speed, previously used dictaphones to compose text. However, one of these users commented that

“Although using voice recognition is slower than using a dictating machine, the document is ready for immediate revision and sending”.

Other users who found writing using voice recognition to be slower than using a keyboard were those who were not experienced in the use of the program. The impressions and speed of these users may be affected by the need for them to learn to use a new program. Training the voice recognition system to recognise your voice and learning its basic commands may take a few weeks. However, once
these skills are acquired by the user, they enable writing to occur more quickly.

Writing that can be carried out more quickly and naturally presents considerable advantages for users of voice recognition systems in terms of the effective utilisation of their short-term memory. With traditional and often slower means of text entry, the correct completion of long complex sentences is difficult because individual parts of the sentence fade from short-term memory by the time the sentence is written (Daiute, 1983). Using voice recognition systems allows writers to get their thoughts and ideas onto paper more quickly and freely, thus allowing them to focus on the more complex mental activities of composing and revising.

For younger school students, the faster speed of writing using voice recognition, allows writing to take place without the need for their short-term memory to be burdened with tasks such as spelling and letter formation (see also Wetzel, 1991b). This increased speed, together with a larger spoken rather than written vocabulary may have a positive impact on both quality of student writing and on their enjoyment of the writing process.

Changes in the technical aspects of written English and in writing style

Respondents were asked to consider how the use of voice recognition systems had affected the technical aspects of their writing, such as sentence construction, grammar and spelling and whether composing using voice input had influenced the style of their writing.

The responses to these survey items are summarised in Tables 1 and 2.

Considering first the technical aspects of written English (Table 1), seven of fourteen respondents considered that use of a voice recognition system had improved their spelling, while four of fourteen believed it had had a positive effect on their sentence construction, grammar and use of vocabulary in their written work. Importantly, only one respondent indicated that these factors had deteriorated. The remainder of the respondents believed that these factors had not changed. An interesting comment from one respondent regarding vocabulary and voice recognition was: "Being able to type (using voice input) longer words more quickly makes me more likely to use them than before."

Table 1 Effect of composing using voice recognition systems on the technical aspects of written English (number of respondents n = 14).

<table>
<thead>
<tr>
<th></th>
<th>Improved</th>
<th>Deteriorated</th>
<th>No Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence Construction</td>
<td>4</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Grammar</td>
<td>4</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Spelling</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Use of Vocabulary</td>
<td>4</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 2 Influence of using voice recognition systems on writing styles (number of respondents n = 13). Note: one respondent did not complete this section of the survey.

<table>
<thead>
<tr>
<th></th>
<th>Observed Change</th>
<th>No Observed Change</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>More conversational</td>
<td>4</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>More descriptive</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>More fluent / understandable</td>
<td>3</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

In questions relating to changes in writing style (Table 2), six of thirteen respondents (note: one respondent did not complete this section of the survey) thought that their writing had become more descriptive, while four and three users respectively considered it to be more conversational, fluent and understandable.

Only one of the users considered that his writing was now more repetitious or less precise. An example of this feeling was a comment from one user who said: "It makes me think more clearly about what words and sentences I want to use."

The information contained in Tables 1 and 2 indicates that in general, respondents felt that the use of a voice recognition system had a positive impact on the technical aspects of their writing and might also lead to their writing style becoming more readable. These results are consistent with those of Daiute (1983), who examined student writing styles using word processors (without voice recognition systems) and observed that faster typing meant a faster flow of ideas and also released a natural linguistic style to replace a more stilted one. Daiute also reported that the faster production of text kept memory free for activities such as developing the content structure of a written piece.

Changes in writing styles with time of use

A detailed examination of the data relating to changes in the technical aspects of written English and writing styles indicates that the total number of such changes reported by individual respondents is related to the length of time they have been using voice recognition systems. Respondents who have used voice recognition systems for longer periods were more likely to report changes in writing styles. The Spearman non-parametric rank-order correlation coefficient of this relationship \( F = .57 \) (\( P = .03 \)), indicates that this relationship is significant at the 95% level.

An explanation for this correlation relationship is that users become more experienced in composing using voice recognition systems, they are able to
devote more of their short-term memory to
tasks of mentally planning and composing
text. Inexperienced users, who are still
learning how to operate voice recognition
systems effectively, may need to
concentrate devote more of their short-
term memory to remembering operating
commands and less to the writing process.
A second factor that may influence writing
is that with increasing use, the recognition
ability of speaker-dependent voice systems
constantly improves. With
better recognition, the user is
able to compose with less
editing interruptions and
therefore the flow of writing
will improve.

**Editing using voice recognition**

When questioned on
the need for editing
documents composed using
voice input, nine of thirteen
believed that such documents
required either the same or
less editing than those which
were produced using a
conventional keyboard. Only
four respondents actually
carried out this editing using
voice recognition commands,
and three of these were people with
disabilities who were unable to use a
keyboard. All other respondents carried
out editing of documents using mouse
controls and conventional keyboards.

The consensus amongst almost all
users was that editing commands were too
time consuming to use. The comments
made by one user regarding making text
corrections using voice recognition
commands, probably sums up the feelings
of many:

“It is too annoying and intrudes on the
flow of dictation. You tend to lose your
train of thought”.

Studies by Gould (1981) estimated
that to produce a simple one page letter
using a computer-based text editor, the
average user made approximately 69
editing and formatting commands and a
further 65 scrolling commands. Carrying
out these operations using a keyboard and
mouse can be relatively easy and quick.
However to do similar tasks using voice
recognition systems is time consuming
and often requires tedious repetition of
voice commands. For instance, to move
the cursor up 11 lines and to the right 2
spaces using a mouse requires a simple
movement and a click. On a voice
recognition system, the following series of
voice commands would need to be given:

“move up 5; move up 5; move up 1;
moving up 2”

On newer versions of voice
recognition programs, including (Dragon
cursor can be moved around the screen
using a voice operated mouse emulator,
but even this process requires a number of
voice commands. The time consuming
and repetitious nature of editing using
voice commands may explain the
preference of survey respondents for using
a conventional keyboard and mouse.

Along similar lines, the reluctance
to use voice commands to carry out text
editing during writing may be related to
the demands this places on short-term
memory. Having to remember and use
voice recognition commands to edit text,
places additional pressure on memory that
may be better utilised elsewhere.

**Motivation and other comments**

The final question in the survey
related to the impact of using voice
recognition systems on the users’
motivation to write. The responses
indicated that of the people surveyed, half
felt an increased motivation to write,
although in some cases this was as a result
of the novelty value of the voice
recognition system. The other users
reported no change in motivation to write.
Two of the users who reported an
increased level of motivation were those
with a disability, for whom writing using a
conventional keyboard was extremely slow
and cumbersome.

A comment made by one voice
recognition user related to
problems encountered with
prolonged use:

“I learnt quickly to treat
voice typing like conventional
typing - don’t use it for hours
without stopping for a
break …”

**Implications for educators and the
need for further research**

It is important to note that all
respondents to this study were
adults, who used voice
recognition systems to compose
text either in the workplace or at
home. Given that voice recognition
systems have been commercially available
in Australia only for around eight years, it
is highly probable that these respondents
had already developed their written writing
skills prior to using this technology. That
is, respondents had already acquired skills
in composing written text to suit particular
purposes, for example an essay, a business
letter, a short memo or a personal
communication. The ability to compose
written text in these different styles is a
learned skill, begun in school and
developed over many years in the
workforce. It is interesting to consider the
effect of using voice recognition systems
on the process of learning to compose
written text.

Learning to compose written text by
speaking (that is using voice recognition
systems), may present problems for inexpe-
rienced writers because of differences
between spoken and written language.
Horowitz and Samuels (1987), in their
summary of literature relating to spoken and written communication, comment that speaking is generally associated with conversation produced, processed and evaluated in the context of a face-to-face exchange and is therefore grounded in interpersonal relationships. Writing can lack this personalised context depending on genre and therefore needs to provide the context in words. Spoken language has been found to contain longer sentences, more repetition and more elaboration than written text (Horowitz & Newman, 1964). In comparison to this, written language had shorter sentences, longer words, more attributives and a more varied vocabulary (Drieman, 1962). There is considerable debate in the literature about whether these variations represent fundamental differences between writing and speaking or whether they are a function of the particular discourse style or genre chosen for writing. However, this discussion is beyond the immediate scope of this study (see Horowitz & Samuels, 1987; Sperling, 1996).

The implication of the previous discussion is that there is a danger that students learning to compose written text using speech (voice recognition), may be more prone to compose written text as if they were speaking. Young children as developing writers are heavily dependent on oral strategies when they attempt to write, but with experience, modify their writing to reflect the required discourse style (Sperling, 1996). Using voice recognition systems to compose written text may make this transition harder. Teachers therefore, will need to assist students in developing skills in speaking in the forms required by different written discourse styles.

In addition to this, students of the future will need to be taught the technical aspects of using voice recognition systems and also the thinking skills that will enable them to make the most effective use of voice input for writing. Vygotsky, in Halpern and Ligget (1984), noted that the positive aspect of voice input which aids short-term memory and written fluency, also has a negative side

“The speed of oral composition is unfavourable to a complex process of formulation - it does not leave time for deliberation and choice.”

If voice input is to be used in the school classroom, students need to be taught the memory and organisational skills to allow them to monitor the progress of oral writing. Students will also need to develop the ability to switch input modes from pen, to keyboard, or to voice (Jones, Frankish & Hapeshi, 1992).

Curriculum designers of the future may also have to consider the relevance of teaching conventional keyboarding skills and instead, may have to design courses around teaching skills in the use of voice recognition systems. Teachers of the future may need to adapt their teaching style to accommodate verbal writing into the classroom situation. As with any new technology, teachers will need to familiarise themselves with voice recognition systems and be given the opportunity to explore its special features and possible uses.

Conclusions

The advent of accurate and relatively easy-to-use voice recognition technology suggests that in the future, such systems could become commonplace in our classrooms. While such technology has many possible uses in the field of education, the implications of its use on writing skills should be carefully considered.

The results of this study indicate that voice input of text may well have a positive effect on the technical aspects of writing and on writing styles. These results however, should be treated with caution because of the small sample size. Future studies should examine effects of voice input on different types of writing with both experienced and inexperienced writers. Examination of actual texts produced using voice input are required to corroborate the self-report data from this survey.

This study indicated that voice recognition systems may be a valuable tool in classrooms of the future. To maximise the potential of this technology, students will need to develop skills in learning to compose written text using voice input and also the memory and organisational skills required to plan, monitor and revise documents produced using voice writing.

REFERENCES