CALL Scotland

Speech Recognition in National Assessments: Update August 2013

THE UNIVERSITY of EDINBURGH

CALL Scotland

July 2013

Speech Recognition Software in National Assessments: Update August 2013

Paul Nisbet

.

CALL Scotland

The Moray House Faculty of Education

The University of Edinburgh

Paterson's Land

Holyrood Road

Edinburgh EH8 8AQ

http://callscotland.org.uk

Contents

Background

Scribes have been used by candidates with disabilities in SQA assessments for many years and currently are still the most popular method of supporting candidates with difficulties with writing and recording (Nisbet, 2013).

However, there are concerns regarding a perception of reliance on scribes by schools and candidates, due to an apparent lack of independence on the part of the candidate, reported challenges associated with finding staff and separate accommodation, and the cost of paying for staff to scribe. ICT, by comparison, has been described by staff and students as providing a more independent method of support (Nisbet, 2005).

The use of ICT as an Assessment Arrangement has become much more popular over the past decade or so, particularly since Digital Question Papers were introduced in 2008. In contrast, the number of assessments where scribes are used has remained roughly constant: in 2008 SQA received 14,811 requests for a scribe and 4,741 requests for use of ICT and by 2012, requests for use of a scribe had fallen slightly, to 14,691 while use of ICT had more than doubled to 10,656. If we consider the percentage of requests that specifically include writing support, we find that scribes have fallen from 33% to 25%, while use of ICT has increased from 10.7% to 18.4% (Nisbet, 2013 p.8). While this does not prove that ICT is replacing scribes - approximately the same numbers of pupils are still using scribes - given the overall increase over the period in the number of candidates and requests, it is probable that many of these candidates would have used scribes had ICT not been an option.

SQA has recently issued specifications regarding reasonable adjustments in national qualifications in Scotland which clarify the role of scribes in assessments of writing. This states that "using human readers and scribes will not be reasonable adjustments where reading and writing abilities are being explicitly assessed" (SQA, 2013).

However, "In order to minimise the disadvantage faced by some disabled learners in attaining the National Units in Literacy, the use of word processors and other assistive technologies such as screen readers, spell checkers or speech-recognition software would be acceptable as reasonable adjustments." (SQA, 2013a).

In 2011, CALL reviewed the use of speech recognition in assessments and reported that the software could technically be used to access SQA digital question papers and assessments but that there was very little evidence of candidates actually using speech recognition in assessments or indeed for classwork (Nisbet, Aitken and Wilson, 2011).

This short report summarises the current situation with regard to use of speech recognition software by learners with disabilities.

Scribes and speech recognition: some considerations

Superficially, one might consider that the support offered to a candidate with a writing difficulty through use of a scribe would be similar to that provided through use of speech recognition software, but this is not the case. A scribe will normally be adept at understanding the candidate's speech regardless of accent or diction and will be able to clarify with the candidate any instances where the scribe is uncertain about the candidate's response. A scribe is likely to be capable of spelling vocabulary and technical terms accurately.

In contrast, while speech recognition software does not make spelling mistakes, the software is not 100% accurate and will make recognition errors – the words will not be misspelled, but they will be the wrong words. The difficulty here is that candidates with literacy difficulties may have difficulty recognising when errors are made and therefore the end result will contain mistakes. It is also much harder for the writer to identify misrecognitions when proof-reading than it is to identify mis-spellings, because the eye (and the computer's spell-checker) picks out unfamiliar mis-spelt words but does not see correctly spelt, but misused, words so easily. Secondly, a marker will often be able to decipher what a candidate means even when vocabulary is poorly spelt, but it is in fact much harder to work out the intended term if the completely wrong word has been typed by the speech recognition program.

Scribes should be able to understand the pupil's speech (and if not, can query and clarify the meaning while scribing). In contrast, speech recognition programs do not cope well with strong dialects or regional accents, or unclear speech, and so some candidates will find great difficulty in using the software. It is worth noting that candidates' level of understanding and ability to use speech recognition should already be understood by returning centres because alternative arrangements require any such arrangements to be in use by the candidate **before** the examination.

Lastly, the operational and technical demands associated with using speech recognition are much higher than those required when using a scribe. In both cases, the candidate must learn how to compose satisfactory English and to dictate clearly and accurately, but a scribe is far more forgiving and flexible than a computer program.

Use of speech recognition in assessments by candidates with disabilities

Speech recognition in SQA examinations

As part of the 2011 report, we followed up nine enquiries received by SQA regarding the use of speech recognition software in examinations in 2009 and 2010, and we found that none of the candidates who were the subject of these enquiries actually used the software in the examination.

Literature review of Speech Recognition in assessments

The potential and use of speech recognition software by learners with writing difficulties has been discussed by many authors (Nelson & Parker, 2004; Nicol, Casey and MacFarlane, 2002; Nisbet, 2002; Phayer, 2011; Venkatagiri, 2002; Williams and Fairweather, 2000); but published research on use in assessment contexts appear to be very few in number (Laitusis *et al*, 2012). In fact, the only directly relevant paper that we could find was written in 2004 by Charles MacArthur and Albert Cavalier (MacArthur and Cavalier, 2004). MacArthur and Cavalier considered feasibility and validity of using speech recognition software for tests involving extended writing, in a study of 31 US high school students aged 14 and 15. There were 21 students with a learning disability affecting their reading and writing and 10 students with no learning disability. The study looked at the feasibility and validity of using speech recognition compared to a scribe and compared quality of written work produced by hand, with a scribe, and using speech recognition.

MacArthur and Cavalier found that students were able to learn to use the software with acceptable accuracy – which they defined as an error rate of under 10%. The present author regards an error rate of one in 10 to be probably unacceptable in an examination, but MacArthur and Cavalier point out that <u>initial</u> recognition accuracy is not as important as the accuracy of the students' dictated, revised and subsequently edited work. Looking at the end result, they report that only three students out of 29 (10%) produced papers that had more than 2% of words that could not be deciphered in context with confidence by those reading the text. This finding, together with the improvements in accuracy of speech recognition software that have taken place since 2004, suggests that the accuracy of speech recognition software would possibly not be a major barrier for many candidates with writing difficulties. However, it is important to note that the study involved North American students and that current generation speech recognition software is optimised for a North American accent. We might expect quite different results were the study to be repeated with students with a strong Doric, Shetland or Glaswegian accent, for example.

MacArthur and Cavalier report that the quality of essays dictated using speech recognition by the students with learning disabilities was higher than the work they produced by hand. This is as would be expected given the sample population.

More interestingly , they found that essays dictated to a scribe were of higher quality than those dictated using speech recognition, which raises two possible unfortunate concerns: either that candidates who use speech recognition might be disadvantaged (compared to the marks obtained if they had used a scribe); or, the use of a scribe may give the candidate an unfair advantage. The key question is whether use of a scribe, or use of speech recognition reflects the candidate's knowledge and expertise most accurately. Given the extensive and extended use of scribes in Scottish examinations evidence on this question is surprisingly sparse.

The MacArthur and Cavalier study also compared the use of speech recognition and scribes by students without learning difficulties and found that there was **no** difference in quality of essays produced by hand, with a scribe or with speech recognition. They argue that "differential impact on students with and without disabilities provides evidence that the accommodation removes a barrier based on disability" and therefore the use of a scribe or speech recognition are valid methods of support. But given that the quality of essays produced by students with difficulties who used a scribe was higher than those produced using speech recognition, the question remains about whether a pupil will be disadvantaged, or advantaged, if they have only the option of using speech recognition.

MacArthur and Cavalier regard such questions as "societal" rather than empirical. In Scotland, SQA have come to the same conclusion by undertaking the consultation on reasonable adjustments that included consideration by the user of scribes and technology.

SQA have determined that a scribe does constitute a reasonable adjustment in most assessments, but not where "writing abilities are being explicitly assessed"[1]. In assessments of writing for National Literacy Units, use of a scribe is not regarded a reasonable adjustment, whereas speech recognition is considered acceptable. The SQA rationale is that the technology is a completely independent method of writing which gives a more accurate measurement of attainment than use of a scribe.

The dispensation to use technology does of course rest on the nature of the assessment and so use of technology is not acceptable in all assessments. For example, in Modern Languages and Gaelic (Learners), neither scribes nor speech recognition are reasonable adjustments because these assessments explicitly assess the candidate's technical knowledge of spelling and grammar, and dictating whole words would undermine the assessment objectives for the writing component. (SQA do allow a scribe, provided the candidate spells each word letter by letter; they also allow the candidate to record their response, play it back and then dictate letter by letter to a scribe.)
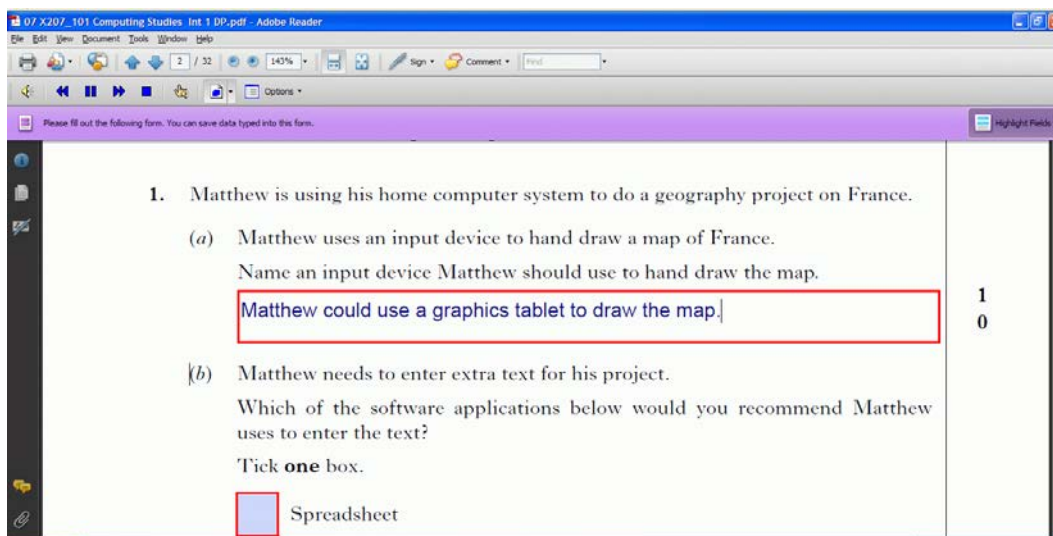
---

[1] SQA (2013) Specification 3 - Literacy Units, http://www.sqa.org.uk/sqa/64702.html

Short or extended answers?

The MacArthur and Cavalier study focused on extended writing and this is the most obvious application for speech recognition for learners with writing difficulties. However, speech recognition can also be used for assessments where short answers requiring one or two sentences are required. An example of this would be an SQA Digital Question Paper in question-and-answer format, with answer boxes inserted for pupils to type or dictate their answers (Figure 1). Using speech recognition for short answers is considerably easier for learners than using it for extended essays because the cognitive load associated with planning an essay and composing text internally is more or less non-existent. With a short answer question, the context and the task are immediate and the pupil can read it as many times as they wish (with text-to-speech, if necessary) before answering the question.

In addition, when answering a short answer question, the learner is more likely to be focusing on the dictated text and is more likely to identify and correct mis-recognised words.

Therefore, it is likely that learners who are new to speech recognition will achieve greater success by using the technology for short dictated exercises than for extended writing tasks.

Figure 1: SQA Digital Question Paper - question and answer format

iPads and mobile devices

One new development that has taken place over the past two years since the first CALL Scotland report on speech recognition, has been the explosion in use of iPads and mobile devices. Speech recognition apps for iPad are available from several suppliers including Nuance, the developers of Dragon NaturallySpeaking for Windows, and the accuracy of their Dragon Dictation app appears to be comparable with the Windows version. More significantly, 'Siri' speech recognition for dictation is now built-in as part of the iOS operating system for iPad Mini, 3 and 4 and informal feedback and reports suggests that its accuracy is at least as good as the speech recognition software available for conventional Windows or MacOS desktop or laptop computers. One advantage of these iPad apps, compared to Dragon on a Windows computer, is that they are free.

All the speech recognition apps for iPad require a fast Internet connection (either 3G or wifi) which immediately raises security concerns in an examination context. However, the Guided Access feature introduced in iOS 6 enables the user to be restricted to only one app, while password protection to prevent access to other apps (e.g. the Internet Browser) provides additional security and so it seems reasonable for iPad speech recognition to be able to be used in assessments.

There is one significant difference between the Siri speech recognition that is built-in to the iPad operating system and the Dragon Dictation app, which is that the built-in Siri dictation can be used to dictate into *any* application whereas Dragon Dictation can only be used to dictate into a basic note pad (Figure 2). With Dragon, the learner can dictate text, correct it, and then either email, tweet or copy the text. In an assessment context then, the candidate would have to dictate the answer into the Dragon note pad, copy it, swap to the word processor app, or the app for completing the digital question paper, and paste the text, which is a relatively cumbersome process that could disadvantage many candidates.
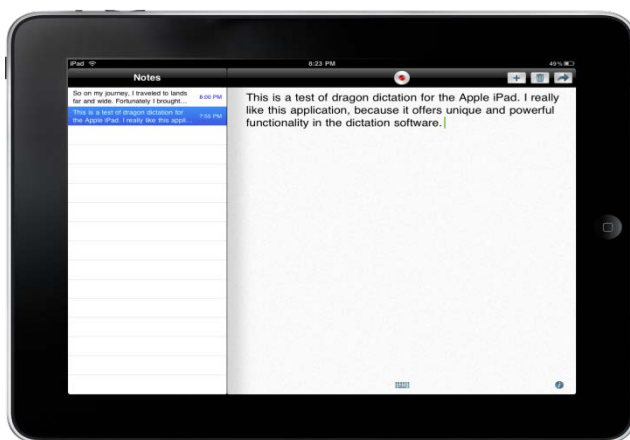

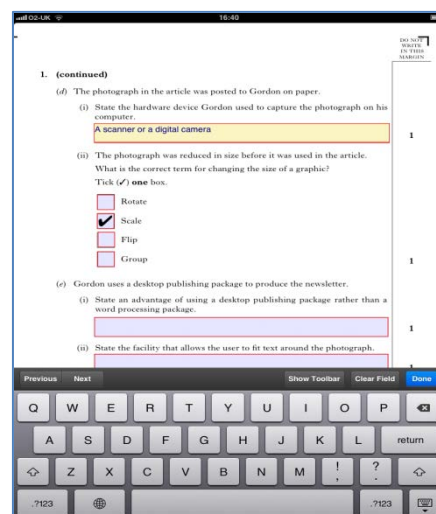**Figure 3: Dragon Dictation app - dictation into the Dragon note pad**

Figure 3: Typing and dictating into a Digital Question paper using Siri

In contrast, Siri can be used to dictate into both SQA digital question papers (opened in either the free Adobe Reader app, or our recommended app, PDF Expert – see Figure 3) and also into SQA digital answer booklets in Word format (opened in for example, the Pages or QuickOffice word processor apps).

This is much more convenient because the candidate does not need to swap between apps. Also, when using the digital answer booklets, the candidate has access to a full set of word processing and editing features and so has parity with candidates using a Windows or MacOS computer.

When dictating into digital question papers in question-and-answer format, the candidate can see both the question and his or her response on screen at the same time, taking advantage of the usability considerations discussed in the previous section.

At the present time, the built-in Siri speech recognition tool appears to be a good option for use in assessments, both for dictating into digital question papers and also answer booklets. (Note however that some alterations are required to be made to the Digital Answer Booklets in Word format to give full compatibility with Pages or QuickOffice.)

Apps for SQA assessments

App	Comment
Pages, £6.99	Pages is probably the most popular word processing app for iPad and provides a full set of word processing features. The text boxes on the front page of the SQA Digital Answer Booklets are not displayed accurately and so need replaced by linear text or tables.
QuickOffice Pro, £10.49	QuickOffice is another good word processing app for iPad. The text boxes on the front page of the SQA Digital Answer Booklets are not displayed at all and so need replaced by text or tables.
Adobe Reader, free	Answers can be typed into Digital Question Papers, and a small set of drawing and annotation tools is provided. Questions cannot be read out with text-to-speech.
PDF Expert, £6.99	Answers can be typed into Digital Question Papers. A full set of drawing and annotation tools is provided. Questions cannot be read out with text-to-speech.

Conclusion

It is likely that speech recognition software remains an underutilised method of support for pupils with disabilities and additional support needs in Scotland. The new specifications on reasonable adjustments in national qualifications for literacy that restrict the use of a human scribe but permit use of technology, coupled with improvements in speech recognition technology itself, the availability of free speech recognition built in to iOS devices, and the introduction of iPads in schools is likely to have at least some impact on this situation.

But the context is changing rapidly: introduction of new national qualifications next session; the new Glow; continuing development of Curriculum for Excellence; coupled with the continuing development and application of mobile technologies such as iPad, means that there is greater need than ever for clear guidance and advice on the use of speech recognition if learners are going to be able to make effective and productive use of the technology.

References

Laitusis, C., Buzick, H., Stone, E., Hansen, E. & Hakkinen, M. (2012) Literature Review of Testing Accommodations and Accessibility Tools for Students with Disabilities. Smarter Balanced Assessment Consortium. www.smarterbalanced.org/.../Smarter-Balanced-Students-with-Disabilities- Literature-Review.pdf

Nelson, A. and C. Parker (2004). Dyslexia and Voice Recognition Software – Really the Perfect Match? International Ergonomics Society Conference in Swansea University, Wales, 2004. http://www.intuitiveinternet.com/Ergsoc_04.pdf

Nicol, A., Casey, C., and MacFarlane, S. (2002) Children are ready for Speech Technology – but is the technology ready for them? http://www.chici.org/index.php/publications

Nisbet, P.D. (2002) Introducing Speech Recognition in Schools http://www.callscotland.org.uk/Common-Assets/ckfinder/userfiles/files/Final_Report.docx

Nisbet, P.D. (2013) SQA Digital Papers 2012. Report for Scottish Qualifications Authority. CALL Scotland. http://www.adapteddigitalexams.org.uk/Downloads/Reports/

Nisbet, P.D., Shearer, N. Balfour, F., Aitken, S. (2006) SQA Adapted Examination Papers in Digital Format: Feasibility Study 2005 – 2006: Final Report. Report to Scottish Qualifications Authority. October 2006. CALL Centre. http://www.adapteddigitalexams.org.uk/Downloads/Reports/

Nisbet, P,D., Aitken, S.A., Wilson, A. (2011) Speech Recognition Software in SQA Assessments. Report to Scottish Qualifications Authority. http://www.adapteddigitalexams.org.uk/Downloads/Reports/

Phayer, J. (2011) "Friends, Romans, Countrymen, lend me your ears…." Speech Recognition Software – its current trend and future potential for students with dyslexia in an educational setting. Closing the Gap April/May 2011. www.closingthegap.com

SQA (2013) 'Section 96 (7) Equality Act 2010: Reasonable Adjustments in National Qualifications in Scotland: Report on the Outcome of the Consultation Section 96(7) Equality Act 2010.' Scottish Qualifications Authority BE6470. http://www.sqa.org.uk/sqa/64698.html

SQA (2013a) 'Specification 3 - Literacy Units'. Scottish Qualifications Authority http://www.sqa.org.uk/sqa/64702.html

Venkatagiri, H.S. (2002) 'Speech Recognition Technology Applications in Communication Disorders', *American Journal of Speech Language Pathology*, 11, 323 – 332. http://ajslp.asha.org/cgi/content/abstract/11/4/323

Williams, S.M, Nix, D., and Fairweather, P. (2000) Using Speech Recognition Technology to Enhance Literacy Instruction for Emerging Readers. http://www.umich.edu/~icls/proceedings/pdf/Williams.pdf

CALL Scotland
The Moray House School of Education
The University of Edinburgh
Paterson's Land
Holyrood Road
Edinburgh EH8 8AQ

http://callscotland.org.uk