



CAMBRIDGE ASSESSMENT

**All the right letters – just not necessarily in the right order.
Spelling errors in a sample of GCSE English scripts.**

**Gill Elliott
Nat Johnson**

Paper presented at the British Educational Research Association Conference,
Edinburgh, September 2008.

Research Division
Cambridge Assessment
1, Regent Street
Cambridge
United Kingdom
CB2 1GG

elliott.g@cambridgeassessment.org.uk
johnson.n@cambridgeassessment.org.uk

www.cambridgeassessment.org.uk

Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate, a department of the University of Cambridge. Cambridge Assessment is a not-for-profit organisation.

All the right letters – just not necessarily in the right order. Spelling errors in a sample of GCSE English scripts.

Abstract

For the past ten years, Cambridge Assessment has been running a series of investigations into features of GCSE English candidates' writing – the Aspects of Writing study (Massey et al, 1996, Massey et al 2005). The studies have sampled a fragment of writing taken from the narrative writing of thirty boys and thirty girls at every grade at GCSE. Features investigated have included the correct and incorrect use of various forms of punctuation, sophistication of vocabulary, non-standard English, sentence types and the frequency of spelling errors. This paper provides a more detailed analysis of the nature of the spelling errors identified in the sample of work obtained for the Aspects of Writing project from unit 3 (Literary Heritage and Imaginative Writing) of the 2004 OCR GCSE examination in English. Are there certain types of spelling error which occur more frequently than others? Do particular words occur over and over again? How many errors relate to well-known spelling rules, such as "I before E except after C"?

Literacy has enjoyed a high profile since 1994 and has been promoted in schools through the introduction of the National Literacy Strategy (NLS). It was unlikely that the 2004 GCSE cohort (the 'population' from whom our writing sample came) was fully exposed to the NLS, because many primary schools introduced the NLS from the bottom up, or at least did not implement it for this cohort (in their final year of primary education in the first year of the NLS) on the basis that it would get in the way of preparation for key stage 2 (KS2) national tests (Beverton & English, 2000). This notwithstanding, Beverton and English noted that, in contrast to previous years, grammar was being taught every day and that all teaching staff in the schools observed had a greater awareness of literacy as a subject in its own right. Therefore the performance of this cohort in spelling is likely to reflect some of the benefits of the NLS.

The study used a stratified random sample of writing taken from a narrative writing task. The only suitable question was found on a paper which formed an alternative to coursework; a question which asked candidates to imagine, rather than to inform, explain, describe, comment, argue or persuade. This option was taken by only 8.3% of candidates - but these amounted to over 5500 candidates from a wide range of schools. The sample was stratified by grade so the fact that this paper was a minority option should be incidental, as the calibre of a candidate achieving a particular grade should be comparable regardless of the route taken through the syllabus. Whilst the possibility existed that schools choosing the examination option might reflect systematic variations in curricular values, comparison of the examination option schools with the entry as a whole did not suggest that the former were unusually socially or educationally selective. The proportions of independent and selective schools as compared with comprehensives and others were the same for the sample as in the overall entry for this English specification.

Spelling errors were identified in the sampled writing by two researchers, working first separately, and then as a team. Each researcher first went through the printed versions of the script samples identifying and counting spelling errors. The two lists of errors and counts were then compared, again grade by grade, and any discrepancies identified and discussed. At any stage it was also possible to inspect the handwritten scripts to verify exactly what the candidate had written. The benefit of the doubt was given in any case where there was ambiguity, which usually arose as a consequence of either poor handwriting, or poor spacing technique. In some cases it was necessary to look elsewhere in the candidate's script for examples of particular letters or letter combinations, or to look at the spacing between other words to see whether the presence or absence of spacing appeared to be deliberate on the part of the candidate.

The study identified 345 spelling errors in 11,730 words written, and these were reported in Massey et al (2005), with a comparison by grade with samples of writing from 1980, 1993 and 1994. It was shown that a considerable decline in spelling in the early 1990s (compared with 1980) had been halted, and at the lower grades, improved.

Since then, we have conducted a detailed analysis of the 345 misspelled words to see if there is evidence of particular types of error. Each misspelling has been categorised, and five broad types of error identified. These are sound-based errors, rules-based errors, errors of commission, omission and transposition, writing errors and multiple errors. This paper will present a detailed examination of the misspellings and the process of developing the categorisation system used. A number of words – *woman*, *were*, *where*, *watch(ing)*, *too* and the homophones *there/their* and *knew/new* are identified as being the most frequently misspelled words. Implications for the findings upon teaching and literacy policy are discussed.

Background

The way in which children learn to spell is linked closely to learning to read, and with other elements of learning to write. Westwood (2008) reviewed the literature from 1995 to 2007 pertaining to the strategies used to teach children to read in English in Australia and Great Britain and Wanzek et al (2006) published a review of a large number of intervention studies carried out between 1995 and 2003.

A number of authors have looked at stages by which a child learns to spell. Ehri (1994) identified a 'logographic' stage, whereby a child deduces meaning from the appearance of the words. Later stages include the ability to match letters to speech sounds (Henderson, 1990) and use these to decode words (read) or to generate their own words (spell). Moats (1995) suggests that a phonetic spelling stage is then attained, with children following a "one letter spells one sound" strategy. This is the point at which spelling can deviate from conventional 'correct' spellings, especially in English where sound rules do not necessarily match letter rules. At this point the successful speller must memorise specific rules such as grammatical endings, and different words which sound the same but are spelt differently. A study carried out between 1995 and 1998 by the Centre for Language in Primary Education (O'Sullivan and Thomas, 2000) collected data from London primary schools and investigated the teaching and learning of spelling throughout the primary years. Amongst other findings the study reported that it is helpful for teachers to study the mistakes made by individual spellers, in order to assess whether the mistakes they are making are phonetic or visual.

In the UK there have been two main methods of teaching a child to read – synthetic phonics, where children are taught letter sounds before being introduced to whole words (Auger & Briggs, 1992), and analytic phonics, where whole words are introduced from the start. Johnston and Watson (2003, 2004, 2005) have suggested that the reading and spelling skills developed by children taught to read using synthetic phonics are very good.

A number of frameworks already exist which incorporate categories of spelling error. QCA (1999) mentions errors due to unstressed vowels, long 'e', omission of single letters, confusion of consonants and homophones. Homophones are also a feature studied by Hepburn (1991) along with doubling and singling of consonants, articulation and errors related to inflectional and derivational morphemes. Finally, Mudd (1994) discusses reasonable phonic alternatives – in other words plausible alternative spellings.

Method

The sample of writing from which the spelling errors were identified consisted of the fourth sentence¹ of question 1 (an extended narrative piece of writing) as written by the candidate, and was taken from the scripts of thirty boys and thirty girls at each grade. Where there were

¹ Everything which appeared between the third and fourth full stop.

insufficient suitable scripts available additional sentences were taken from available scripts. The sentences sampled were keyed into Word™ by a temporary member of staff, preserving all errors of punctuation and spelling. Careful checking was undertaken to ensure that the keying, including errors, had been accurately undertaken. Counts of the numbers of words were then obtained from Word™ software.

Table 1 shows the number of words which were sampled at each grade.

Table 1: Number of words sampled at each grade.

Grade	A*	A	B	C	D	E	F	G
Number of words	1238	1082	1303	1208	1567	1734	1739	1859

Spelling errors were identified by two members of staff, working first separately, and then as a team. Each person first went through the printed versions of the script samples, grade by grade, identifying and counting spelling errors. The two lists of errors and counts were then compared, again grade by grade, and any discrepancies identified and discussed. At any stage it was also possible to inspect the handwritten scripts to verify the exact marks placed on the paper by the candidate. The benefit of the doubt was given in any case where there was ambiguity, which usually arose as a consequence of either poor handwriting, or poor spacing technique. In some cases it was necessary to look elsewhere in the candidate's script for examples of particular letters or letter combinations, or to look at the spacing between other words to see whether the presence or absence of spacing appeared to be deliberate on the part of the candidate.

Results

Overall numbers of spelling errors

The study identified 345 errors in 11,730 words written. Therefore, 97.1% of words were correctly spelled.

Figure 1 shows the overall numbers of spelling errors by grade. As expected, the number of errors increase by descending grade. Given that spelling errors are one of the (admittedly many) criteria for judging English writing, it would be unexpected if they did not. Figure 2 shows the same data as a percentage of the total number of words, thus adjusting the bars for the number of words written in total (candidates at different grades wrote different numbers of words, and as every word written presents an opportunity for a spelling error, variability in the total number of words might influence the pattern of results). In fact the adjusted graph remains very similar to the raw data.

Figure 1: Number of spelling errors by grade

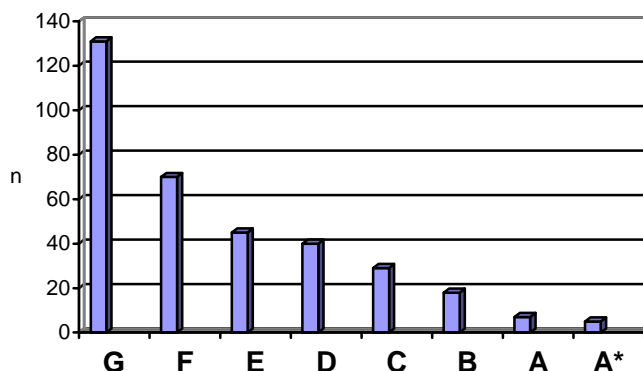
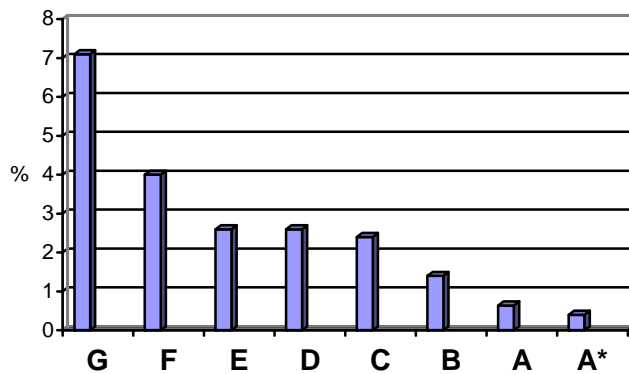


Figure 2: Rate of spelling errors by grade



This paper provides detailed analysis of all the errors to see if there is evidence for particular types of error. Appendix 1 gives the entire list of words which were spelled wrongly, arranged in alphabetical order.

Frequently occurring misspelt words

A few words occur more frequently than others. Words which appear in the list more than twice are listed in table 2, along with the frequency of their occurrence, a list of each misspelling and a list of the grades at which the misspellings occur (The misspellings and corresponding grades are given in the same order, to enable the reader to identify which particular misspelling occurs at which grade).

Table 2: Frequently occurring misspelt words.

Word	Freq.	Misspellings	Grade
before	3	<i>befor, befor, be for</i>	GGG
finally	4	<i>finaly, finily, finaly, finaly</i>	BGGG
here	3	<i>he, hear, hire,</i>	EFG
knew	5	<i>new, new, new, new, new</i>	DEEFF
their	3	<i>ther, there, thire,</i>	FFF
there	10	<i>ther, their, their, their, the, their, their, their, their, ther</i>	BCEEFFGGGG
they	6	<i>thay, thay, thay, thay, thay, thay</i>	GGGGGG
too	4	<i>to, to, to, to</i>	DEFF
towards	3	<i>to-wards, to words, to words</i>	EGG
until	3	<i>untill, untill, untill</i>	CCF
watch(ing)	4	<i>wach, waching, waching, waching</i>	EFGG
went	3	<i>when, when, whent</i>	GGG
were	5	<i>where, where, where, where, where</i>	CDDDG
where	4	<i>were, were, were, were</i>	DDDF
woman	11	<i>women, women, women, women, women, women, women, women, women, women, woneman, women,</i>	BCDDEEEFFG
you	3	<i>u, yo, yoy</i>	GGG

Seven of these words – *here, their, there, too, were, where, you* – appear in published Key Stage 1 lists and *before, knew, until, watch, woman* all appear in Key Stage 2 lists.

Although *women* for *woman* is the single most frequently occurring mistake with ten instances (and occurs at every grade from B downwards), the *their/there* homophone is a close second, with eight occurrences, seven of which are *there* for *their*.

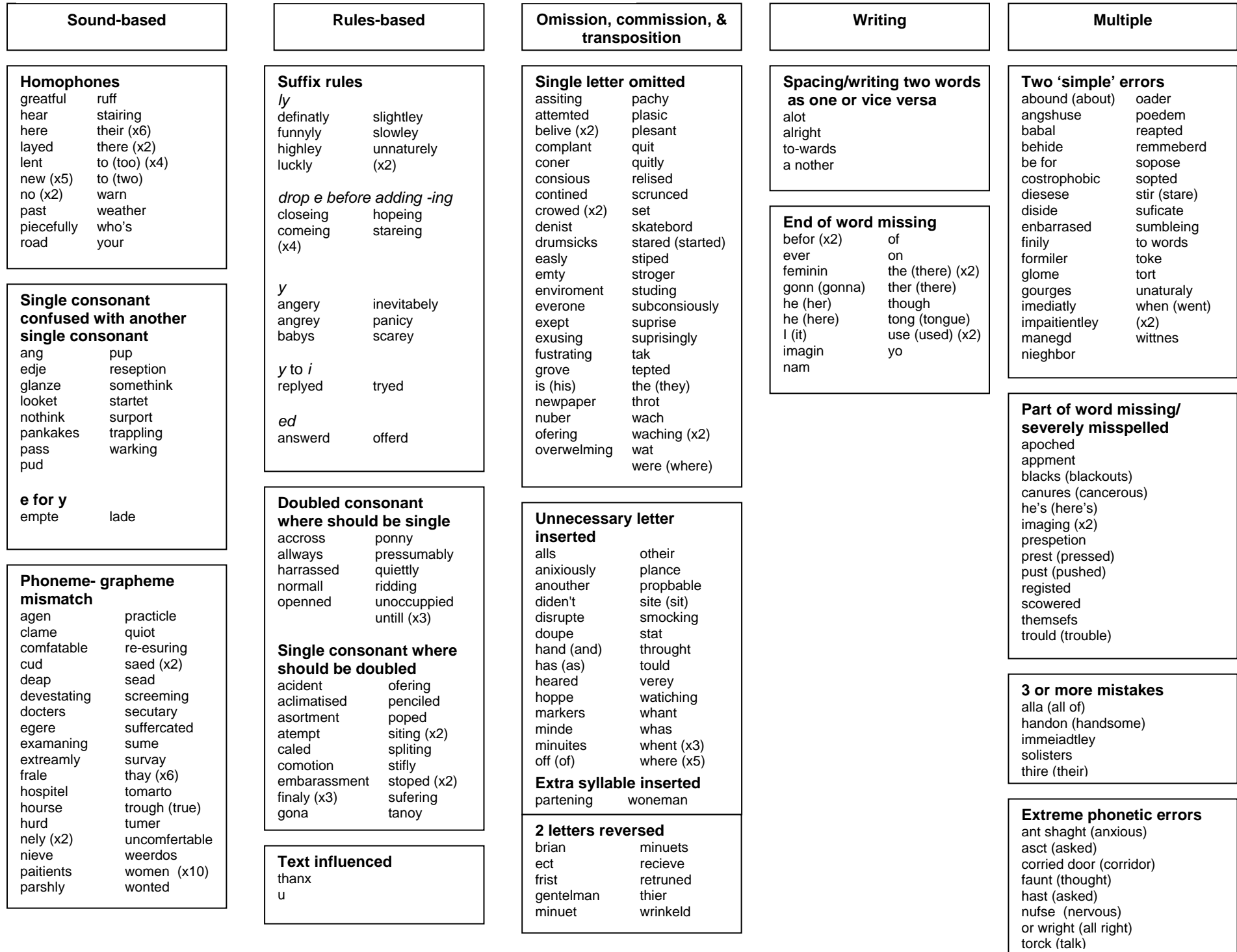
Misspellings by type.

The misspellings presented by candidates have been grouped into broad related categories of error. Categories were derived via a process of grouping together similar error patterns, and are shown in figure 3. As far as possible the 'types' of error were kept as simple as possible, in the spirit of the original Aspects of Writing (the generic name given to the series of reports produced by Cambridge Assessment, and its predecessor, UCLES) research. This resulted in the following categories:

- Sound-based error – homophones, incorrect consonant, e for y, vowel sound error, morpheme error.
- Rules based error – doubling/singling, text-speak
- Omission, commission and transposition – single or paired letters added, omitted or transposed
- Writing error- spacing, end of word missing.
- Multiple errors.

Where a misspelling might fall into several categories (i.e. *accross*, which is both a doubling error and the insertion of an additional letter) the most obvious/most precise error type was allocated; in this case, doubling).

Figure 3: Misspellings by type.



Discussion of error types

Sound-based

Homophones form the first category of error types. 34 of the 345 errors (9.8%) were of this type. The *there/their* and *know/no, /knew/new* confusion accounted for nearly half of these. These two errors have already been discussed in the section on frequently occurring misspelt words.

Fifteen errors consisted of the transposition of a single wrong consonant. Many of these were phonetically plausible spellings – however there were instances of a ‘k’ at the end of *-ing*, instead of the ‘g’, and of ‘t’ replacing ‘d’ in *-ed* endings. These were potentially due to articulation error, resulting in spelling error. Two errors involved the transposition of a vowel for a consonant – in both cases ‘e’ for ‘y’.

Fifty two errors related to the vowel sound. Again (or *agen* according to one such candidate), most of these were phonetically plausible spellings. Nonetheless, many of these words are to be found on the lists of spellings at KS1 and KS2 - e.g. *hospital, heard, some, doctor, they*.

Rules based

Doubling/singling errors

There were 13 doubling errors and 22 singling errors, together accounting for 10% of all errors. Only one of the errors (*acclimatised*) was an example of an affix error.

Suffix errors

There were 24 suffix errors (7% of the total), of which a very high proportion involved adding *-y* or *-ly* to a word or involved the *y* to *i* rule (changing a *y* to an *i* before adding *-ed* (e.g. *replied*).

Two errors were ‘text’ (mobile phone/computer text messaging) influenced. Once again these are phonetically plausible alternatives to conventional spelling and are intentionally used in defiance of ‘conventional’ spelling rules during text messaging. The very small number of these errors was remarked upon in the original report, and it is pleasing to see that candidates seem by and large to be aware that they must not use such devices in a written English examination, however much they are used in social contexts.

Omission/commission/ of single letter and **transposition.**

Forty nine errors consisted of the omission of a single letter, whilst thirty four were the insertion of a single letter. In some cases these were clearly the result of idiosyncratic spellings – notably silent letters. In other cases, the error perhaps owes more to carelessness.

Only ten errors were a straight reversal of two letters, and just one of these related to the ‘*ie/ei*’ rule.

Writing errors

Two types of error have been categorised as ‘writing’ errors. These are errors of spacing – writing two words as one or vice versa, and missing the last letter from a word. In several instances there is evidence from the scripts that candidates did know the correct spelling in the case of the latter category, but had left off the final letter in haste.

Multiple errors

These errors form the most serious type of mistake, and have most effect upon the appearance of the word. First are those misspellings which seem to be made up from two separate errors. For example:

impaitientley consists of two separate inserted letters; **impa(i)tientl(e)y**

and
nieghbor consists of a transposition and an omitted letter (in UK spelling); **n(ie)ghbo(u)r**

Second are those errors where a whole part of a word is either missing or severely misspelled. The third category within this group contains those few words with three or more individual mistakes, and it was one of the misspellings – *immeiadtley* – which prompted the title of this paper – all the right letters, just not necessarily in the right order. Finally there are a group of words which bear little physical resemblance to their correct spellings, yet have clear phonetic links with them. These are referred to as extreme phonetic errors. It is possible that this latter category may be related to the very specific types of error made by sufferers of dyslexia, but further discussion of this is beyond the scope of the present paper.

Discussion

This paper has attempted to categorise spelling errors made by students in their GCSE English examination in 2004 into various categories. The purpose of the research was to establish whether certain spelling errors – or certain categories of error – are particularly common, and how they relate to spelling conventions, as taught within schools.

The study has identified five categories of spelling error which further subdivide into sixteen subsections. The categories were derived from the errors observed, rather than from existing categories, so there may be other groups of spelling error which have not been discussed here, simply because they were not encountered. In general most misspellings fall into the first three categories – sound-based error, rules based error and errors of omission, commission and transposition. The first two of these categories contain many misspellings that are undoubtedly very familiar to teachers. However, there are no particular sub-categories that are particularly prone to more errors in our sample than others. English is a language which has more than its fair share of idiosyncratic spellings and complex spelling rules. Not surprisingly, many of these errors are connected with those. However, within the category of a single additional letter, there were a number of examples of an unnecessary silent 'h' – *where (were)*, *whant*, *whas*, which are worthy of comment. The category of omission, commission and transposition is more difficult to interpret. It is quite possible that many of these errors occurred as a result of the examination conditions under which candidates were writing, combined with, perhaps, a lack of effective proof-reading of their final piece. The sub-category of writing errors, where the ends of words are missing, could in some cases be due to the same issues. However the spacing of two words as one, or vice versa, is almost certainly due to candidates' perceptions of those words. Finally, the category of multiple errors produces words which look least like conventional spellings. Interestingly, two simple errors can produce a word that is almost unrecognisable, and it is important to be able to decode these errors for what they are, rather than simply seeing a very distorted word.

Fifteen individual words were identified as occurring with relatively high frequency. In particular, two of these were seen far more often than others. They were the *there/their* homophone, which has been known to be problematic since time immemorial, and *women* for *woman* (not vice versa). *Knew/new* and *know/how* also occurred with relative frequency, but again, this is unlikely to surprise the teaching profession.

A major limitation to the data presented here is the fact that there is no control over which words candidates choose to use. Therefore the study is not a 'fixed' spelling test, and cannot be generalised in the same way as reports of spelling tests. A word spelt wrongly just once does not mean that 479 students can spell it, simply that they did not necessarily try. It would be possible to investigate correctly spelt words to give the other side of the picture, but that would be an enormous task.

There is clearly no single over-riding type of error which is made by the group of GCSE students from whom we have sampled. Those errors that are made are varied, and although it is disconcerting to note the number of most frequently occurring errors which are taught at Key

Stage 1, it is, on the other hand, heartening to see how few (relatively speaking) errors are made, when you consider the number of words written overall, especially given that the text was written under examination conditions with no access to dictionaries.

References

Augur, J. and Briggs, S. (1992) *Hickey Multi-Sensory Language Course*. London: Whurr Publishers Ltd.

Ehri, L. C. (1994) *Development of the ability to read words: update*. In R. Ruddell, M. Ruddell and H. Singer (eds). *Theoretical Models and Process of Reading*. Newark, Del. International Reading association.

Henderson, E. (1990) *Teaching Spelling*. Houghton Mifflin, Boston.

Hepburn, J. (1991) Spelling categories and strategies. *Reading* April 1991.

Qualifications and Curriculum Authority (1999) *Technical Accuracy in written English: Research findings*. QCA. London.

Johnston, R.S and Watson, J. (2003) *Accelerating Reading and Spelling with Synthetic Phonics: A five year follow up . Insight 4*. Scottish Executive Education Department: Edinburgh.

Johnston, R.S and Watson, J. (2004) *Accelerating the development of reading, spelling and phonemic awareness*. *Reading and Writing*, 7 (4), 327-357

Johnston, R.S and Watson, J. (2005) *A seven year study of the effects of synthetic phonics teaching on reading and spelling achievement . Insight 17*. Scottish Executive Education Department: Edinburgh.

Massey and Elliott (1996) *Aspects of Writing in 16+ English examinations between 1980 and 1994*. Occasional Research Paper 1. University of Cambridge local Examination Syndicate.

Massey, A. J., Elliott, G. L. & Johnson, N.K. (2005) *Variations in aspects of writing in 16+ English examinations between 1980 and 2004: Vocabulary, Spelling, Punctuation, Sentence Structure, Non-Standard English*. *Research Matters: Special Issue*, November 2005.

Moats, L. C. (1995) *Spelling: Development, Disability and Instruction*. York press. Baltimore.

Mudd, N. (1994) *Effective Spelling.: A practical guide for teachers*. Hodder & Stoughton. London.

O'Sullivan, O. and Thomas, A. (2000) *Understanding Spelling*. Routledge. London.

Wanzek, J., Vaughn, S., Wexler, J., Swanson, E.A., Edmonds, M and Kim, A.H. (2006) *A synthesis of spelling and reading interventions and their effects on the spelling outcomes of students with LD* *Journal of Learning Disabilities*, Vol 39(6), pp 528-543.

Westwood, P. (2008) *Revisiting issues in spelling instruction: A literature review 1995-2007*. *Special Education Perspectives*, Vol. 17 (1), pp. 33-48

Appendix 1: Alphabetic list of words which were spelled wrongly in a sample of GCSE English writing.

<p>A</p> <p>about accident acclimatised across again all of all always a lot all right and angry another anxious anxiously answered approached appointment as asked assortment assisting attempt attempted</p> <p>B</p> <p>babble babies before behind believe blackouts brain</p> <p>C</p> <p>called cancerous claim closing coming comfortable commotion complaint corner conscious continued corridor claustrophobic crowded could</p> <p>D</p> <p>deep decide definitely dentist devastating didn't disease disrupt doctors dope drumsticks</p>	<p>E</p> <p>eager easily edge embarrassed embarrassment empty environment etc every everyone examining except excusing extremely</p> <p>F</p> <p>face familiar feminine found finally first frail frustrating funnily</p> <p>G</p> <p>gentleman glance gloomy gonna gorgeous grateful groove</p> <p>H</p> <p>had handsome harassed her hear heard here here's highly his hope hoping hospital horse</p> <p>I</p> <p>imagine imagining immediately impatiently inevitably it</p> <p>J</p> <p>K</p> <p>knew know</p>	<p>L</p> <p>lady laid leant looked luckily</p> <p>M</p> <p>makers managed mind minute minutes</p> <p>N</p> <p>naive name nearly neighbour nervous newspaper normal nothing number</p> <p>O</p> <p>odour of off offered offering one opened other overwhelming</p> <p>P</p> <p>pancakes panicky partially parting patchy patients passed past peacefully pencilled pleasant place plastic podium pony popped practical prescription pressed presumably probably pub pushed</p>	<p>Q</p> <p>quiet quietly quite</p> <p>R</p> <p>realised reassuring reception receive registered remembered repeated replied returned riding rode rough</p> <p>S</p> <p>said sat scary scoured screaming scrunched seat secretary sit sitting skateboard slightly slowly smoking solicitors some something splitting spotted stare staring started stiffly stopped striped stronger studying stumbling subconsciously suffering suffocate suffocated support suppose surprise surprisingly survey</p>	<p>T</p> <p>talk tannoy tempted thanks their themselves there they thought throat too told took tomato tongue towards trampling tried trouble true tumour two</p> <p>U</p> <p>uncomfortable unnaturally unoccupied until used</p> <p>V</p> <p>very</p> <p>W</p> <p>walking want wanted was watch watching weirdoes went were what whether whose where witness woman worn wrinkled</p> <p>X</p> <p>Y</p> <p>you you're</p>
---	---	--	--	---