

Phonemic-graphemic correspondences in the mind

This is a new project initiated by Paul Tench, which has been supported and developed by Ben Clarke and Rhiannon Vaughan Griffiths as their Research Experience within the MA in Language and Communication Research. It is a project that was started from scratch in February 2007.

This study derives from four sources:

- Aitchison's *Words in the Mind*, in which the mental lexicon is described in terms of networks including the phonological, but, perhaps surprisingly, omitting the orthographical;
- second language research, for example van Berkel's paper in Cook & Bassetti *Second Language Writing Systems* (2004), in which the Dutch and English phonemic-graphemic correspondences are contrasted;
- mother tongue research, eg Kress's *Early Spelling*, in which the development of native speaker children's competence and creativity is presented;
- and PT's field studies in orthography for unwritten languages, in which he has sought to understand the processes by which adults who are literate in other languages used their orthographic awareness to construct an orthography for their own unwritten mother tongue.

The focus of enquiry was to ascertain whether an adult has a developed orthographical system that in some senses at least is parallel to a mature phonological competence, given the extensive discrepancy between pronunciation and spelling in current English. It seemed best to investigate this possibility by asking a group of adults to try and spell unfamiliar words, and see what correspondences between phonemes and letters each person displayed, and to what degree those adults held such correspondences in common. Thus, we would be able to see whether there was system in the use of the orthographic resources in the mind, and offer an explanation to supplement Aitchison's *Words in the Mind*.

Detailed hypotheses were prepared to cover young educated British adults' competence in all the phonemic-graphemic correspondences that are required for English, based on observations of the most common and regular ways of representing each phoneme. This was mainly PT's initiative, but it soon emerged that this was not to be a simple task. It became quickly apparent that the three collaborators did not share a simple equivalence of phonemes and letters, and that differences between us led us to realize that not only were there variations for many correspondences, but also that position in a word would be critical, and type of word also: there does seem to be a paradigm shift in the orthographic conventions for 'common' and technical words. The designation 'technical/scientific' was unproblematic, but we took some time to agree on a term to describe 'non-technical' words; the term 'common' was chosen as the least problematic.

Another paradigm shift appears to happen with the spelling of words deemed to be foreign; we did not have the resources to investigate this notion, but evidence for it did eventually appear in the analysis of the experimental data.

Eventually, the set of hypotheses included 9 general hypotheses, 19 specifically for consonants, 19 for vowels, 5 with additional grammatical information, and a further

19 for consonants and 18 for vowels in technical words. They were all tested in pilot and full experiments that involved the dictation of 50 'novel' words to 45 British students; the hypothesis that students would distinguish one system for 'common' words from a system for technical words was also tested. The data was analyzed for both the degree of uniformity among the subjects and the degree of conformity to the set of hypotheses.

We began our project with a discussion of the four sources of information noted above. We conducted a trawl of relevant publications through CU library holdings via Voyager; as far as we could see no publication addressed the point of enquiry we wished to address. Nevertheless, we agreed that we should each select one publication with a view to presenting it to the others, both to inform ourselves and sharpen up our own thinking. The selections were as follows:

PT: Luelsdorff, P, ed. (1987) *Orthography and Phonology*. Amsterdam: John Benjamins

RVG: Kress, G (2000) *Early Spelling: between Convention and Creativity*. London: Routledge

BC: Caramazza, A, ed (1997) *Access of Phonological and Orthographical Lexical Forms*. Hove: Psychology Press

We agreed to meet together weekly for an hour (at least) for ten weeks. In three successive weeks, first PT (to set the precedence), then RVG and BC presented their reviews. PT introduced the terminology that Sgall uses for a linguistic treatment of spelling; what emerged from RVG's presentation was the interesting parallel between the differences between Halliday's studies of the functions of language for children (see his *Explorations in the Functions of Language*) and for adults (see his *Introduction to Functional Grammar*), and Kress's study of children's spelling and the project we were engaged in; the value of BC's presentation was the psycholinguistic dimension of the examination of an aphasic's problems with spelling.

At each meeting, we sought to refine the hypotheses; this was done by testing each other with a set of nonsense words that each devised. It was during these sessions that we realized the extent of alternatives and variations of spellings according to position, eg /ei/ as <ai>, <a.e> and <ay>, and that the orthographic conventions for technical words differed significantly enough from 'common' words to warrant a division of labour. PT decided to ask RVG to concentrate on the common words and BC to concentrate on the technical words. Also during these sessions we constructed the experimental design for a pilot study, which was held in about week 4.

The pilot study proved to be an invaluable resource for the design of the experiment proper. It helped

- to define more precisely the procedure for conducting the experiment; RVG took notes of the conduct of the pilot, and agreed to formulate the precise wording for the delivery of instructions at the experiment proper;
- to construct items that could realistically be regarded as 'English', eg to eliminate sequences that were not typical of English, like long vowels (other than /ɑ:/) before final /-sk/;
- to eliminate items that were too similar to taboo words for comfort;
- to eliminate items that could be spelt as existing words;

- to expect ‘legitimate’ alternative spellings;
- to exclude items as unnecessary for further investigation because of total agreement, eg /lʌs/ as <luss>;
- to give experience for the timing of the list of items; it gave us confidence that a list of 50 items would not in fact be too onerous – it took less than 20 minutes to complete;
- to confirm a decision to use two speakers, one for the ‘common’ words (PT) and one for the technical words (BC), but without indicating to the subjects why; it helped us to talk about “Paul’s words” and “Ben’s words” without reference to there being any difference in the nature of the words themselves;
- to confirm our suspicions that some grammatical information is occasionally necessary, eg final /d/ as a past participle marker, or not;
- to require information on any recognized linguistic disorder in any subject, eg dyslexia;
- to give two novice researchers actual experience of conducting an experiment of this kind;
- and to decide to audio record the dictation of items as a checking mechanism for consistency.

The experiment proper took place in week 6; PT constructed a revised set of ‘common’ words and BC adjusted his technical words in accordance with our discussions following the pilot. The experiment had the advantage of taking place in a scheduled class for Year 2 students, who were identified in terms of age, gender, regional background and any linguistic disorder, but suffered the disadvantage of a venue adjacent to a drilling work. In the qualitative data that we asked subjects to provide, much comment was made about the interference of noise, but the data was not unduly affected by the disturbance, we felt, except perhaps in one respect: some subjects occasionally wrote items that conformed to real words, as if they expected those words to have been intended – despite the introduction in which it was made plain that none of the words were real. For instance, /kli:/ was interpreted as *clean* by six (of 45), as if having heard /kli:/, they felt obliged to overcompensate for the noise interference. The value of the audio recording was confirmed in this case, as no final /n/ was detected by a third party on our behalf.

A sheet was prepared for the subjects to write down their spellings, consisting simply of two columns numbered 1-25, and 26-50, and space for comments. This would help in locating items across the 45 subjects. We took care to preserve an archive copy but photocopy the sheets for our own use. PT left RVG and BC for a month to analyse the data; RVG to analyze the 35 ‘common’ words, and BC the 15 technical words. This was a roughly equal division of labour, as the technical words were much longer and contained as many items for analysis as were found in the ‘common’ words. We agreed on Excel spreadsheets for documenting the data.

After the month, we met to review the data and agree the writing up protocols. RVG and BC each had to write their own account of both the process and the product of their Research Experience. The results show that the hypotheses should be tightened up in wording in some cases, but be radically changed in others. PT will undertake this, and also arrange the hypotheses more systematically in the light of the data, eg to present each consonant phoneme in at least five different word positions:

- 1 initial;
- 2 medial, and 3 final, following either a long vowel or a short vowel that is perceived as being spelt as a bigrapheme (eg /ʌ/ as <oo>) or a weak vowel, and preceding a consonant;
- 4 medial, and 5 final, following a short vowel that is perceived as being spelt as a protographeme;

and vowels, simply in non-final position, and final open and closed syllables.

The hypotheses for technical words can be made more specific.

The grammatical information that appears necessary for spelling novel words needs expansion.

Due care will need to be taken over accent differences among subjects in future investigations; this was not a noticeable feature in the present data, but it is realized that it could be with a different range of items, eg /ɑ:/ in certain environments as <ar> for Northerners and <a> for Southerners . Due care will need also to be taken to avoid unduly foreign sounding items, unless that aspect is specifically being targeted.

The current analyses provide more than enough material for the presentation of a paper at a BAAL conference.

In conclusion, it is asserted that just as it is possible to investigate and describe the phonological systems that a person operates, it is also possible to investigate and describe the orthographical systems, including the paradigm shift in spelling what are perceived to be technical items. The study therefore completes Aitchison's *Words in the Mind*, contributes a much more comprehensive set of phonemic-graphemic correspondences than van Berkel's for second language research, extends Kress's study into young adulthood, matching Halliday's *Learning how to Mean*, and provides evidence of literate competence in any culture.

P Tench
 June 2007
 (1872 words)