ORTHOGRAFIK INKONCISTENSY EPHEKTS IN FONEME DETEKTION?

Anne Cutler
MPI for Psycholinguistics, Nijmegen, The Netherlands
Rebecca Treiman
Dept. of Psychology, Wayne State University, Detroit, USA
and
Brit van Ooijen
Laboratoire de Sciences Cognitives et Psycholinguistique CNRS, Paris, France

ABSTRACT

The phoneme detection task is widely used in spoken word recognition research. Alphabetically literate participants, however, are more used to explicit representations of letters than of phonemes. The present study explored whether phoneme detection is sensitive to how target phonemes are, or may be, orthographically realised. Listeners detected the target sounds [b,m,t,f,s,k] in word-initial position in sequences of isolated English words. Response times were faster to the targets [b,m,t], which have consistent word-initial spelling, than to the targets [f,s,k], which are inconsistently spelled, but only when listeners' attention was drawn to spelling by the presence in the experiment of many irregularly spelled fillers. Within the inconsistent targets [f,s,k], there was no significant difference between responses to targets in words with majority and minority spellings. We conclude that performance in the phoneme detection task is not necessarily sensitive to orthographic effects, but that salient orthographic manipulation can induce such sensitivity.

1. INTRODUCTION

The spelling of *focus* in English renders the word's three consonants [f,k,s] as F, C and S, which indeed are the most common letters used to represent those sounds in English.

But English spelling notoriously allows different sound-to-letter mappings, and *focus* might well have been spelled *phokuce,* by analogy with *photo, yokel* and *lettuce*.

The complexity of the English spelling system causes problems for non-native writers, and indeed for many native users of English. Other languages closely related to English have much more regular spelling - Italian and Dutch, for instance, contain relatively few exceptions to the majority mappings of sounds to orthography. There is no reason to believe that these differences in how transparently the languages are written affect the use of spoken language in any important way. However, the study of spoken-language processing may be indirectly affected by orthographic regularity. This is because the experimental methods used to study auditory comprehension could in some cases encourage listeners to draw on their knowledge of orthography.

A task which is widely used in spoken word recognition research is phoneme detection. In this task listeners hear spoken input - words, nonwords or sentences - and press a response key as soon as they detect an occurrence of a pre-specified phonemic target. In the most common form of the task the target phoneme must occur at the beginning of a word or nonword, but in "generalised" phoneme detection an occurrence of the target in any position must be responded to (see Connine and Titone, 1996).

Researchers who construct materials for phoneme detection experiments usually try not to complicate their studies with orthographic effects, in that they avoid choosing target-bearing words which are irregularly spelled; for instance, researchers generally avoid asking subjects to respond to [n] in *knee*, or to [j] in *fuse*. Sometimes, however, the experimental situation demands that this practice be abandoned. Thus Van Ooijen, Cutler and Norris (1992) compared the detection of the vowel targets [i] and [u] with the semivowel targets [j] and [w], because the two sets of targets allowed a contrast between vowels and consonants with minimal acoustic difference. Most occurrences of [j] in British English words are found in words like *fuse*, and Van Ooijen et al. indeed used such items. They found that these items, however, produced very long response times (RTs) and a miss rate above 25%. They suggested that the absence of an orthographic representation of the target sound [j] in the spelling of *fuse* made the detection task particularly difficult. Similarly, Frauenfelder, Segui and Dijkstra (1990) explained an observation of longer phoneme detection RTs for the target [k] than for the targets [p] and [t] in French in terms of a larger repertoire of possible orthographic realisations for [k] than for the other two phonemes, in written French.

Dijkstra, Roelofs and Fieuws (1995) conducted a direct test of whether phoneme detection is sensitive to how target phonemes are realised orthographically. Their experiment (which was in fact conducted at the same time as the present study, without either research group knowing of the other's work) examined the detection of [k], [t] and [s] in Dutch words and nonwords, using the generalised form of the phoneme detection task. Dutch is largely regularly spelled, but contains some phoneme-to-grapheme variation, particularly in loan words. Thus Dijkstra et al. compared detection of [k] in words with the majority spelling K (e.g. *kabouter* 'gnome'; *paprika*) versus words with the minority spelling C (e.g. *cabaret; replica*). They found little difference in response times to word-initial targets, but when the target occurred later in the word, after the word had become uniquely identifiable, majority spellings ([k] as K, for instance) produced faster RTs than minority spellings ([k] as C). Dijkstra et al. do not describe their experiment in great detail, and it is unclear how large the differences were for the separate target phonemes; one of their phoneme targets, [t], has the minority orthographic spelling D, but this spelling is confined to syllable-final position, and many of the words in question have the phoneme [d] in a morphological variant (thus the plural of *hoed* [hut] 'hat' is *hoeden* [hud@]).

The present study included an investigation of the same issue raised by Dijkstra et al., but the investigation was conducted in English. It is not at all clear whether results from Dutch (with its largely regular orthography, at least for native words) predict results to be found in English (in which the notoriety of the orthographic irregularity may result in greater listener sensitivity to orthographic effects). However, the present study also addressed a wider range of questions than Dijkstra et al's experiment. Finally, our study further differed from Dijkstra et al.'s study in that it used the standard word-initial form of the phoneme detection task.

The first question which we asked (Question 1) was the most general one, namely whether consistency of phoneme-grapheme mapping affects how easily consonants are detected. To address this question we compared three consonant targets which are orthographically fully consistent with three which are inconsistent. Every English word beginning with the singleton sounds [b,m,t] also begins with the letters B, M, T respectively (assuming that, for example, *pterodactyl* may be omitted from the relevant computation). The consonants [f,k,s], however, can be represented with different letters: *foal/phone, coy/key, seed/cedar*. The primary question in our study was therefore whether detection of [b,m,t] would prove systematically simpler than detection of [f,k,s].

Secondly, we also asked whether the likelihood of orthographic effects on phoneme detection is affected by the overall orthographic consistency of the context (Question 2). Perhaps English listeners do not usually pay attention to orthographic mappings in phoneme detection, but may do so if their attention is drawn to this factor by a large number of orthographically irregular words in the input (as may, for instance, have been the case with the materials used by Van Ooijen et al.). Thus we further independently manipulated whether the filler items (words without occurrence of a specified target, constituting the rest of the stimulus lists) were orthographically regular, or contained a large number of spelling irregularities.

Finally, we asked, as did Dijkstra et al., whether the actual orthographic realisation of a consonant influences ease of detection (Question 3). Thus for the targets [f,k,s] we compared detection RTs in words beginning with the majority spellings F, C, S with detection RTs in words beginning with the minority spellings PH, K, C. Spelling preferences were determined from the CELEX database (Baayen, Piepenbrock and Van Rijn, 1993), which showed a majority to minority ratio of 9.9:1 for [f], 15.7:1 for [k] and 17.1:1 for [s].

2. METHOD

Materials

90 English words of two or three syllables each were chosen to act as target-bearing words. For each of the six sounds [b,m,t,f,k,s] there were 12 words beginning with the majority spelling, i.e. beginning with the letters B, M, T, F, C, S: *bucket bandit barnacle basic basket burnish bacon budget baronet beverage balance bicycle; mouldy motive monkey magpie marriage mackerel medical mischievous mullet minstrel molecule manager; turban tangible tiny topical tavern toddler tunnel towel technical terrible turnip tipsy; fumble formula follicle ferry funnel final filament fodder fasten furtive function federal; corner carrot caravan cabin corduroy culprit couple cousin carton canopy consonant cabbage; supper single sickle sonnet seven socket sabotage supple serpent sovereign sapphire silhouette*. For the latter three sounds, i.e. [f,k,s], there were additionally six words each beginning with the minority spelling, i.e. PH, K, C: *pharmacy phantom physical phony pheasant photo; kidney kerosene kitchen kingdom kangaroo kennel; cedar centre celery cipher circuit citadel*.

A further 600 words, of one to four syllables, were chosen for use as filler and practice items. 192 of these were in some way irregularly spelled (e.g. *kneel, guardian, wrinkle, tongue, cough, pyjamas*). The rest were orthographically regular (e.g. *nip, garnish, wiggle, lung, honk, paradise*).

Four materials sets were constructed, each containing 96 lists, varying in length from two to six words. 72 of these were experimental lists, containing a target-bearing word in third, fourth or fifth position in the list. The remaining 24 lists contained no occurrence of the specified target, or contained a word beginning with the target in first, second or final position. For each of the six phoneme targets [b,m,t,f,k,s] there were 12 experimental lists and four filler lists (two without occurrence of the target).

The four sets of materials differed in (a) the target-bearing words used for the three inconsistent phoneme targets [f,k,s], and (b) the nature of the filler items. In two sets the experimental lists contained only target words with majority spellings *(fumble, corner, supper*), while in the other two sets half of the targets for each of [f,k,s] began with minority spellings (e.g. *pharmacy, kennel, cedar*). Crossed with this factor, in two of the sets all 384 filler words were orthographically regular (e.g. *nip, garnish, wiggle*) while in the other two 192 filler words were regular and 192 were orthographically irregular (e.g. *kneel, guardian, wrinkle*).

The lists were recorded onto Digital Audio tape by a male native speaker of British English. Each list's target sound was specified immediately prior to the list. Six short practice lists were also recorded. Cross-splicing between lists ensured that the same recording was used for each occurrence of any experimental target-bearing word.

Subjects and Procedure

Forty-eight native speakers of British English, students of St. Catherine's College, Cambridge University, and members of the MRC Applied Psychology Unit subject panel, took part in the experiment in return for a small payment. All had normal hearing.

Subjects were tested individually in a quiet room, and were instructed to listen within each list for a word beginning with the target sound, and to press the response key as quickly as possible once the target sound had been detected. The instructions emphasised that only target sounds in word-initial position counted, but did not draw attention to spelling. The six practice lists were presented first, followed by one of the sets of 96 lists; twelve listeners heard each of the four sets. The lists were presented binaurally via closed Sennheiser HD-520 headphones. Coincident with the onset of each target word, a frequency pulse on the non-speech channel of the tape, inaudible to the subjects, started the clock of a Zenith microcomputer running an experimental control program; the timing was stopped by the subject's keypress. The RTs were collected and stored by the computer.

3. RESULTS

RTs below 100 ms or above 1500 ms were discarded. Discarded responses plus missed responses amounted together to only 4.17% of the data, an inadequate amount to warrant further analysis of miss rates. The grand mean RT was 392 ms. Mean RTs were calculated for each subject and item, and these were subjected to separate analyses of variance with subjects and with items as random factors. Table 1 shows the mean RTs to the consistent target phonemes [b,m,t] versus the inconsistent target phonemes [f,k,s] as a function of filler condition and target spelling condition.

	Spelling of [f,k,s] targets			
	majority only		majority/minority	
Phonemes	b,m,t	f,k,s	b,m,t	f,k,s
100% regular fillers	381	377	359	371
50% irregular fillers	357	411	416	466

Table 1: Mean RTs (ms) across subjects and items for consistent [b,m,t] versus inconsistent target phonemes [f,k,s], as a function of filler condition (all orthographically regular versus 50% irregular), and of target spelling for [f,k,s] (majority only, i.e. F, C, S, versus majority/minority, i.e. F, C, S and PH, K, C).

To address Question 1, we compared mean RTs to the consistent targets [b,m,t] with those to the inconsistent targets [f,k,s]. Across all conditions, mean RT to [b,m,t] was 378 ms, to [f,k,s] 406 ms. Because subjects with majority/minority spellings were responding in part to a different set of target words than subjects with only majority spellings, the two spelling groups were analysed separately. The difference between [b,m,t] and [f,k,s] was marginally significant for the majority spelling group (F1 [1,22] = 6.77, $p < .02$, F2 [1,66] = 3.24, $p < .08$) and significant for the majority/minority spelling group (F1 [1,22] = 12.03, $p < .005$, F2 [1,66] = 7.82, $p < .01$). To address Question 2 we looked at the strength of this consistency effect as a function of the regularity of orthography within the filler items. For both groups the consistency comparison interacted significantly with the filler regularity comparison: F1 [1,22] = 8.85, $p < .01$, F2 [1,66] = 22.12, $p < .001$ for the majority group, F1 [1,22] = 4.48, $p < .05$, F2 [1,66] = 4.4, $p < .05$ for the majority/minority group. (The main effect of filler regularity was not significant.) T-tests across subjects and across items showed that the consistency comparison was never significant for either group when all fillers were orthographically regular, but was significant for both groups when fillers included irregular spellings (t1 [11] = 3.24, $p < .01$, t2 [35] = 3.09, $p < .005$ for the majority group, t1 [11] = 3.03, $p < .02$, t2 [35] = 2.88, $p < .01$ for the majority/minority group). In other words, the answer to questions (1) and (2) is that phoneme-grapheme consistency has no necessary effect, but an effect can be induced by context which draws listeners' attention to phoneme-grapheme mapping.

To address Question 3, we carried out a separate analysis of RTs to the inconsistent targets [f,s,k] for the half of our subjects who had responded to targets with both majority (F, C, S) and minority (PH, K, C) spellings (see Table 2). There was no significant difference between words spelled with majority (mean RT 429 ms) versus minority spellings (409 ms), and this spelling factor did not interact either with the phoneme target ([f] versus [k] versus [s]) or with the filler regularity comparison. Thus to question (3) the answer is that the actual orthographic realisation of a consonant target does not determine RT.

Target phoneme	[f]		[k]		[s]	
Spelling	F	PH	C	K	S	C
100% regular fillers	400	360	382	349	368	367
50% irregular fillers	491	473	444	464	488	440

Table 2: Mean RTs (ms) across subjects and items for inconsistent target phonemes [f,k,s] in the majority/minority spelling condition as a function of target spelling (F, C, S versus PH, K, C) and filler condition (all orthographically regular versus 50% irregular).

4. DISCUSSION

The results of this study indicate that performance in the phoneme detection task is not necessarily sensitive to orthographic effects. At least in the standard, word-initial, form of the task, listeners can detect an inconsistently spelled consonant as rapidly in words in which it is represented by its minority spelling (*pharmacy, kennel, cedar*) as in words in which it receives its majority spelling (*fumble, corner, supper*). This is in accord with the finding of Dijkstra et al. (1995) with the generalised phoneme detection task and Dutch materials; in their study, too, responses to word-initial phoneme targets showed minimal orthographic majority/minority effects.

Furthermore, listeners can detect consonants which are always consistently spelled (i.e. [b,m,t]) and consonants which are inconsistently spelled (i.e. [f,k,s]) equally rapidly, as long as the experimental situation does not call attention to the orthographic dimension. This suggests that the standard practice of researchers conducting phoneme detection experiments, namely avoidance of obvious orthographic irregularities, should exercise the desired effect of averting unwanted confounding factors.

However, it is possible by the use of salient orthographic manipulation to induce sensitivity to orthographic consistency. Thus when half of the experimental items were irregularly spelled, listeners responded more rapidly to the consistent target phonemes [b,m,t], which are always realised in word-initial position as B, M, T, than to the inconsistent phonemes [f,k,s], which have varying realisations. This finding accords with other evidence that performance in the phoneme detection task varies in sensitivity to higher-level factors as a function of changes in the experimental situation. Thus Cutler, Mehler, Norris and Segui (1987), for example, found that RTs to phoneme targets were faster in words than in nonwords when the filler items in the experiment were phonologically varied, but with the same target-bearing items no word-nonword difference was found when the filler items in the experiment were phonologically uniform.

Varying use of different sources of information in phoneme detection as a function of task demands is easily captured in the most recent model accounting for performance in this task, the Merge model (McQueen, Norris and Cutler, in press; Norris, McQueen and Cutler, submitted). In this model, information from autonomous bottom-up processing is integrated at a dedicated decision-making stage with information from the lexicon (and potentially from other sources). Varying task demands can cause different weighting of the information arriving from the various sources. Assuming that orthographic information comprises part of the lexically stored knowledge about words, this information can be made available at the decision-making stage once lexical candidates are activated and compete for recognition.

In general, spelling may be accorded very low weight in phonemic decision-making. Salience of orthographic realisation in the experiment as a whole (i.e. in the present case, a large number of irregularly spelled filler words) can, however, cause the decision-making mechanism to take spelling into account in its decision. Taking spelling information into account in the detection of consistently spelled targets (in the present case, in words beginning with the phonemes [b,m,t]) will have little effect on RTs, since all activated candidates will have the same orthographic realisation of the target. Consulting spelling information for inconsistently spelled targets (in the present case, the phonemes [f,k,s]) will, however, complicate decisions, since some activated candidates will be spelled one way and some another. This will be equally true whether the words in question actually have majority or minority spellings (thus accounting for the finding that RTs were not significantly different across these two groups of items).

Finally, consider Dijkstra et al.'s finding that RTs to majority- and minority-spelled phonemes did differ at later points in the word, once only a single viable activated candidate word remained. We suggest that listeners may have been able to use the orthographic information in the fully activated lexical representations to generate phoneme candidates prior to the actual occurrence of the target in the input. That is, *papr-* and *repl-* suffice to activate *paprika* and *replica* and to generate expectation of a [k]. This generation process may then be harder in the case of minority spellings. Whether this final finding from Dutch (with its relatively regular orthography) would carry over to English (with its highly irregular orthography) might form a future research phokuce.

5. ACKNOWLEDGEMENTS

This study was carried out at the Medical Research Council Applied Psychology Unit, Cambridge, England. The second author acknowledges support from the National Science Foundation, USA (grant SBR-9020956). The third author was supported by a grant from the University of Leiden. Thanks to Bram de Kruijff for computing the lexical statistics and to Ton Dijkstra, James McQueen and Dennis Norris for useful discussions. Email addresses of the authors: anne.cutler@mpi.nl, treiman@math.wayne.edu and brit@lscp.ehess.fr.

6. REFERENCES

1. Baayen, R.H., Piepenbrock, R., and van Rijn, H. *The CELEX lexical database (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA, 1993.

2. Connine, C.M., and Titone, D. “Phoneme monitoring”, *Language and Cognitive Processes* 11: 635-645, 1996.

3. Cutler, A., Mehler, J., Norris, D.G., and Segui, J. “Phoneme identification and the lexicon”, *Cognitive Psychology* 19: 141-177, 1987.

4. Dijkstra, T., Roelofs, A., and Fieuws, S. “Orthographic effects on phoneme monitoring”, *Canadian Journal of Experimental Psychology* 49: 264-271, 1995.

5. Frauenfelder, U.H., Segui, J., and Dijkstra, T. “Lexical effects in phonemic processing: Faciltatory or inhibitory?”, *Journal of Experimental Psychology: Human Perception and Performance* 16: 77-91, 1990.

6. McQueen, J.M., Norris, D.G., and Cutler, A. “Lexical influence in phonetic decision-making: Evidence from subcategorical mismatches”, *Journal of Experimental Psychology: Human Perception and Performance*, in press.

7. Norris, D.G., McQueen, J.M., and Cutler, A. “Merging phonetic and lexical information in phonetic decision-making”, *Behavioral and Brain Sciences*, submitted.

8. Van Ooijen, B., Cutler, A., and Norris, D.G. “Detection of vowels and consonants with minimal acoustic variation”, *Speech Communication* 11: 101-108, 1992.