Language and Speech

http://las.sagepub.com/

Strategic Deployment of Orthographic Knowledge in Phoneme Detection

Anne Cutler, Rebecca Treiman and Brit van Ooijen

Language and Speech 2010 53: 307
DOI: 10.1177/0023830910371445

The online version of this article can be found at:
http://las.sagepub.com/content/53/3/307

Published by:
SAGE
http://www.sagepublications.com

Strategic Deployment of Orthographic Knowledge in Phoneme Detection

Anne Cutler[1]**, Rebecca Treiman**[2]**, Brit van Ooijen**[3]

[1] *Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and MARCS Auditory Laboratories, University of Western Sydney, Australia*
[2] *Department of Psychology, Washington University in St Louis, St Louis, MO, USA*
[3] *Department of Human Resources, Maersk Oil, Doha, Qatar*

Key words

orthography

phonemes

spelling consistency

spoken-word recognition

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 realized. 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 spelling was rendered salient by the presence in the experiment of many irregularly spelled filler words. Within the inconsistent targets [f, s, k], there was no significant difference between responses to targets in words with more usual (*foam*, *seed*, *cattle*) versus less usual (*phone*, *cede*, *kettle*) spellings. Phoneme detection is thus not necessarily sensitive to orthographic effects; knowledge of spelling stored in the lexical representations of words does not automatically become available as word candidates are activated. However, salient orthographic manipulations in experimental input can induce such sensitivity. We attribute this to listeners' experience of the value of spelling in everyday situations that encourage phonemic decisions (such as learning new names).

Acknowledgments: This research was carried out at the Medical Research Council Applied Psychology Unit, Cambridge, UK. An earlier report, entitled "Orthografik inkoncistensy ephekts in foneme detektion?", was presented to the 5th International Congress on Spoken Language Processing (Sydney, 1998). Support is acknowledged from the National Science Foundation, USA (RT: grant SBR-9020956) and the University of Leiden (BvO). We thank Bram de Kruijff for computing the lexical statistics, and Ton Dijkstra, James McQueen and Dennis Norris for helpful discussions

Address for correspondence. Anne Cutler, Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands; e-mail: <anne.cutler@mpi.nl>

Language and Speech 0023-8309; Vol 53(3): 307–320; 371445; DOI:10.1177/0023830910371445
http://las.sagepub.com

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. However, English spelling notoriously admits multiple phoneme-to-grapheme mappings, and *focus* might also have been spelled *phokuce*, by analogy with *photo*, *yokel* and *lettuce*.

The English spelling system is far less regular than that of many other European languages (e.g., Italian or Dutch, both of which have largely predictable phoneme-to-grapheme correspondences). This causes problems for non-native writers, and indeed for many native users of English. But does a spelling system play a role in the processing of spoken as well as of written language? Findings from studies of spoken-word recognition have suggested that it might. In a variety of tasks, hearing a spoken word has been shown to make available to the listener the stored orthographic form of that word. Thus auditory lexical decision is easier for words with a single consistent spelling (e.g., *catch*) rather than many potential spellings (*deed*; cf., *bead*, *cede*, *keyed*; Ziegler & Ferrand, 1998); and it is easier for words which have previously been heard embedded in other words only when the word and its carrier match in both sound and spelling (*mess* in *message*, but not *deaf* in *definite* or *leg* in *legislate*; Jakimik, Cole, & Rudnicky, 1985; Slowiaczek, Soltano, Wieting, & Bishop, 2003). If rhyming words match orthographically as well as phonologically, then rhyme priming between them is stronger (*dirt–shirt* vs. *hurt–shirt*; Chéreau, Gaskell, & Dumay, 2007), and rhyme judgments about them are faster (*pie*, *tie* vs. *guy*, *tie*; Seidenberg & Tanenhaus, 1979). Orthography can even play a role when no stored lexical forms can be drawn upon; thus spoken nonwords prime real words if they could be spelled as the real words, but not if they could not ([stæl] primes *stall* but [kræl] does not prime *crawl*; Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008).

As Taft et al. (2008) pointed out, many experiments on orthographic effects in listening could be seen as allowing strategic responding; attending to spelling could help listeners perform the experimental task. To make conscious strategies less likely, Taft et al. used surrounding maskers in their auditory priming task, so that listeners were not aware of the primes they heard. A variable pattern resulted: orthographic effects were far stronger when listeners had to decide whether the input was a real word (lexical decision) than when they only had to repeat it aloud (shadowing).

Such a variable pattern is itself an indication of strategic responding, albeit not necessarily a consciously chosen strategy. Similar variability has been found in the spelling-consistency effect of Ziegler and Ferrand (1998), described above; although this effect appears not only in lexical decision but also in other categorization judgments about words (Peereman, Dufour, & Burt, 2009), it is not found in shadowing unless a lexical decision is simultaneously required (Ventura, Morais, Pattamadilok, & Kolinsky, 2004). The orthographic congruency of real-word rhyme primes as first reported by Seidenberg and Tanenhaus (1979) also fails to affect responses in shadowing (Pattamadilok, Kolinsky, Ventura, Radeau, & Morais, 2007). Results of cross-modal masked priming studies have likewise been claimed on the one hand to support autonomy of the systems underlying written versus spoken word processing

(Kouider & Dupoux, 2001), on the other to indicate mandatory interaction between these systems (Grainger, Diependaele, Spinelli, Ferrand, & Farioli, 2003).

Nearly all studies in this line of research have involved decisions about words. There is little debate about where orthographic knowledge is stored; it is associated with the lexical representations of words. The principal question at issue has then been: under what circumstances does this aspect of lexical knowledge play a role in lexical processing: only when it is clearly relevant (i.e., in reading), or also when it is ancillary to the task at hand (i.e., in speech)? Note that orthography is not unique in being a non-essential aspect of stored lexical knowledge which may play a role in speech processing; the same is true, for example, of morphological structure (McQueen & Cutler, 1998). As is obvious from the above review, the weight of the evidence now makes it clear that orthographic knowledge can become available as listeners hear spoken words. The central issue, with orthography as with all such types of knowledge, is whether it must become available.

Variable patterns of results across differing experimental situations are not well captured by a model in which a particular knowledge source has mandatory influence; they are better accounted for in a framework allowing for strategic recourse to knowledge which is in principle independent. In the Merge model of speech processing (Norris, McQueen, & Cutler, 2000), for instance, information from different sources is drawn upon as and when it is useful. Varying task demands render varying types of information strategically relevant, and these, extracted from their multiple sources, can be integrated at the earliest feasible stage to inform the decision required by the experimental task. The weight assigned to each information source will also vary as a function of the demands of the task. As a result, differences in the task induce different patterns of results.

Mandatory activation of information has important theoretical consequences, in particular for distinguishing between models which incorporate top-down flow of information from lexical to prelexical levels of processing during speech recognition (e.g., McClelland & Elman, 1986) and models in which information flow is strictly bottom-up (e.g., Norris & McQueen, 2008; Norris et al., 2000). Top-down connections make lexical information automatically available to prelexical processes, while a bottom-up architecture requires a motivated call upon such information. If orthographic information were shown to be unavoidably available at a prelexical processing stage during speech recognition, the former type of model would be supported. If it were shown to be subject to strategic recourse, the latter type of model would provide the more appealing explanation.

In the case of orthographic effects in speech processing, the evidence concerning mandatory activation is not fully clear. For one thing, the use of lexical tasks has restricted the arena to the processing level at which knowledge of spelling, stored in lexical representations, is indubitably present. Moreover, although strategic responding has been frequently raised as an issue, this has mostly been in the context of trying to rule it out, rather than explicitly testing for it by predicting variable patterns as a function of deliberate task manipulations.

In the present study, therefore, we take the debate a step further by attempting to induce a bias in favor of the use of orthographic knowledge in speech processing.

We do this by manipulating the experimental conditions but not the response items themselves; that is, we predict that identical response items, which allow room for the use of orthographic knowledge, will only show effects of such knowledge when the experimental conditions make it salient. Further, we use a prelexical task, that is, one in which the decisions are not about words but about the phonemes that make them up; this is the phoneme detection task, in which the use of lexically stored knowledge is known to vary as a function of response bias (Cutler, Mehler, Norris, & Segui, 1987).

In phoneme detection, 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, although in "generalized" phoneme detection an occurrence of the target in any position must be responded to (see Connine & Titone, 1996). Researchers have usually tried to keep phoneme detection studies free of orthographic effects by avoiding targets with irregular spelling; so subjects are generally not asked to respond to [n] in *knee*, or to [j] in *fuse*, for example. Sometimes, though, the experimental question requires that this practice be abandoned. Thus in order to contrast vowels and consonants with minimal acoustic difference, Van Ooijen, Cutler, and Norris (1992) compared detection of the vowel targets [i] and [u] versus the semivowel targets [j] and [w]. Most occurrences of [j] in British English words are found in words like *fuse*, and Van Ooijen et al. indeed used such items. The response times (RTs) to these items, however, were very long, and the miss rate was above 25%. Frauenfelder, Segui, and Dijkstra (1990) also 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 realizations for [k] than for the other two phonemes in written French. These findings indicate that orthography could affect phoneme detection.

Dijkstra, Roelofs, and Fieuws (1995) conducted a direct test of whether phoneme detection is sensitive to how target phonemes are realized orthographically, using the generalized form of the phoneme detection task with Dutch words and nonwords containing [k], [t] and [s]. Dutch is largely regularly spelled, but allows some phoneme-to-grapheme variation, particularly in loan words. Thus Dijkstra et al. compared detection of [k] in words with the usual spelling K (e.g., *kabouter* 'gnome'; *paprika*) versus words with the less usual spelling C (e.g., *cabaret*; *replica*). Although their study again involved no explicit strategic manipulation, they too found a variable pattern of results: little RT difference for word-initial targets, but faster RTs to usual spellings ([k] as K, for instance) than to less usual spellings ([k] as C) when the target occurred later in the word, after the word had become uniquely identifiable.

The design of the present study can be summarized in terms of three questions. In order to investigate whether an effect can be made to come and go as a function of strategic manipulations, we first need to create the appropriate conditions for such an effect. Our initial comparison therefore concerns whether consistency of phoneme–grapheme mapping in English affects how easily consonants are detected. This issue is addressed by comparing three consonant targets that are orthographically fully consistent with three that are inconsistent. The singleton sounds [b, m, t] at the onset of English words are always spelled with the letters B, M, T respectively (assuming that, e.g., *pterodactyl* may be omitted from the relevant computation). The consonants [f, k, s] in the same position, however, may be represented by different letters even in similar phonetic contexts: *foam/phone*, *cattle/kettle*, *seed/cede*. Our Question 1 was thus whether detection of [b, m, t] is simpler than detection of [f, k, s].

The strategic manipulation was overlaid upon this comparison, in that we then asked whether orthographic effects in phoneme detection, if they are on offer, could be rendered more or less likely by the overall orthographic regularity of the context (Question 2). The likelihood of English listeners paying attention to orthographic mappings should increase if the input contains a large number of orthographically irregular words. Thus we 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 also asked, as did Dijkstra et al., whether the actual orthographic realization of a consonant influences ease of detection (Question 3). Thus for the three inconsistent targets [f, k, s] we contrasted words beginning with the more usual ("majority") spellings F, C, S with words beginning with the less usual ("minority") spellings PH, K, C in the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993) and the phoneme–grapheme correspondence lists of Gontijo, Gontijo, and Shillcock (2003); this revealed majority–minority ratios of 9.9:1 for [f], 15.7:1 for [k] and 17.1:1 for [s].

2 Method

2.1 Materials

Ninety 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, that is, beginning with the letters B, M, T, F, C, S, and for the latter three sounds, that is, [f, k, s], there were additionally six words each beginning with the minority spelling, that is, PH, K, C. The target word sets were balanced on two factors known to play a role in detection tasks: word length in phonemes, and frequency (using British English frequency counts: Johansson & Hofland, 1989). A further 602 words, of one to four syllables, were chosen as filler and practice items. Of these, 192 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*). All target-bearing words, the two sets of filler words, and the remaining 26 words used in the practice lists are listed in the Appendix.

Four materials sets were constructed, each containing 96 lists, varying in length from two to six words; 72 of these lists contained 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 lists containing one of the listed target-bearing words, and four filler lists (two without occurrence of the target).

As all word sets featured all three consistent and all three inconsistent targets, all sets provided relevant data for Question 1. The four sets of materials differed, however, in (a) the nature of the filler items, and (b) the target-bearing words used for the three inconsistent phoneme targets [f, k, s]. In Sets 1 and 2, all 384 filler words were orthographically regular (e.g., *nip*, *garnish*, *wiggle*) while in Sets 3 and 4, 192 filler words were regular and 192 were orthographically irregular (e.g., *guardian*, *kneel*, *wrinkle*). Comparison of the responses given all-regular fillers versus some irregular fillers addresses Question 2. Crossed with this factor, in two of the sets (Sets 1, 3) the experimental lists contained only target words with majority spellings (e.g., *fumble*, *corner*, *supper*), while in the other two sets (Sets 2, 4) half of the targets for each of [f, k, s] had minority spellings (e.g., *pharmacy*, *kennel*, *cedar*). This difference in orthographic realization of the targets allows Question 3 to be addressed.

The lists were recorded to digital audio tape by a male native speaker of British English. Each list's target sound was spoken immediately prior to the list (e.g., "[bə] *as in boy*"). Six short practice lists, for which other target sounds ([g, l, v, w]) were specified, were recorded at the same time. Cross-splicing between lists ensured that the same recording was used for each occurrence of any experimental target-bearing word.

2.2 Participants

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

2.3 Procedure

Participants were tested individually in a quiet room. They were first given written instructions to listen, within each list they heard, for a word beginning with the sound specified as target for that list, and to press the response key as quickly as possible once the target sound had been detected. The instructions emphasized that only target sounds in word-initial position counted. Attention was not drawn to spelling; the example lists in the instructions used the target sound [v] represented by the letter V. The lists were then presented binaurally via closed Sennheiser HD-520 headphones. The six practice lists were presented first, followed by one of the four sets of 96 lists; 12 listeners heard each set. Coincident with each target word onset, a frequency pulse, inaudible to the subjects, started the clock of a 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; this was insufficient to warrant further analysis of miss rates. The grand mean RT was 392 ms. Mean RTs were calculated for each subject and item, and separate analyses of variance were carried out with subjects and with items as random factors. Table 1 shows the mean RTs to

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)

	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

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.

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 analyzed separately. Overall, the difference between [b, m, t] and [f, k, s] was significant across subjects but not items for the majority spelling group, $F1(1, 22) = 6.77$, $p < .02$, $F2(1, 66) = 3.24$, $p < .08$, and significant across both for the majority/minority spelling group, $F1(1, 22) = 12.03$, $p < .005$, $F2(1, 66) = 7.82$, $p < .01$. The results thus show some effect of phoneme–grapheme consistency, but this effect does not appear to be fully robust.

To address Question 2, our strategic question, we computed 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 filler regularity: $F1(1, 22) = 8.85$, $p < .01$, $F2(1, 66) = 22.12$, $p < .001$ for the majority group, and $F1(1, 22) = 4.48$, $p < .05$, $F2(1, 66) = 4.4$, $p < .05$ for the majority/minority group. (The main effect of filler regularity was not significant.) This interaction was further explored with *t*-tests across subjects and across items. When all fillers were orthographically regular, the consistency comparison was insignificant for both groups. When fillers included irregular spellings, however, the consistency effect proved significant, and robust across subject and items, and again this held for each group, $t1(11) = 3.24$, $p < .01$, $t2(35) = 3.09$, $p < .005$ for the majority group, and $t1(11) = 3.03$, $p < .02$, $t2(35) = 2.88$, $p < .01$ for the majority/minority group. Thus the full answer to Questions 1 and 2 is that phoneme–grapheme consistency has no necessary effect on phoneme detection responses, but an effect can be induced by context in which phoneme-to-grapheme mapping is highly variable.

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

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)

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

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 realization of a consonant target does not determine RT.

4 General discussion

There is no mandatory effect of orthography on the prelexical processing of speech as indexed by phoneme detection; but orthographic effects at this early processing level can be induced by making spelling particularly salient. Our results clearly show that listeners can detect an inconsistently spelled word-initial 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*), and that detection of phonemes allowing inconsistent spellings ([f, k, s]) is in principle no harder than detection of phonemes that are always consistently spelled ([b, m, t]). Saliency of spelling, though, makes detecting members of the inconsistent set harder.

These findings break new ground in showing that strategic use of orthographic knowledge is so easily manipulated; however, they are also in accord with indications from the previous literature. They support the general surmise of Taft et al. (2008) that much of the evidence in favor of orthographic effects in spoken-word processing probably reflects strategic responding, and add explicitly induced task-related variability to the variability across experimental situations described in the introduction both for the spelling-consistency effect in spoken-word processing and the orthographic contribution to rhyme priming. With respect to the present phoneme detection task, they also match in English the results of Dijkstra et al. (1995) with the generalized form of the same task and Dutch materials; although that study is often cited (e.g., by Kouider & Dupoux, 2001, or Peereman et al., 2009) as showing adverse effects of spelling mapping in general, Dijkstra et al. in fact found significant effects only in targets which occurred after the word could be uniquely identified. Just as in the present experiment, no significant orthographic effects appeared in responses to their word-initial phoneme targets.

Another low-level phonemic processing task that has displayed effects of orthography is phoneme goodness rating; asked to rate the goodness of a token of [s] on a continuum varying from [z] through [s] to [ʃ], listeners tested by Cutler, Davis, and Kim (2009) produced higher mean ratings for the best tokens of [s] when these were heard at the end of a word spelled with S (e.g., *bless*) than when the same tokens were heard at the end of a word spelled with C (e.g., *voice*). No such difference appeared when the tokens were rated in nonsense words inheriting the exceptionless orthographic probabilities of the same lexical neighborhoods (*pless*, *bloice*). The spelling effect here is likely to have reflected strategic responding; in making hundreds of goodness decisions about minimally different tokens, listeners found spelling a useful aid, and this delivered two letters differing in how consistently they represent the sound [s]. The sound [s] is nearly always written S, so the letter S is a better [s]. This effect did not arise by automatic lexical influence, because in that case it should have affected the judgments for nonwords with exceptionlessly spelled real-word rhymes.

Phoneme detection is by now a well-understood task. One obvious characteristic of the task is that it is not a necessary stage in everyday speech processing; we do not recognize words by performing sequential detection of prescribed phoneme targets which make them up. Explicit decisions about phonemes sometimes need to be made in ordinary speech situations (when we fear we may have misheard, or when we learn a new name, for example). However, spoken-word recognition is normally a continuous process, in which probabilistic information about words is constantly evaluated and adjusted, and phonemes play a role by virtue of being the minimal unit distinguishing one word from another rather than by constituting an independent stage in the processing of speech (see McQueen, 2007; Norris & McQueen, 2008). The goal of this process is optimally efficient veridical recognition of incoming spoken messages, whereby both speed and accuracy are of the essence—so that on the one hand, any information which can be drawn upon to speed the processing of the input will indeed be drawn upon, but on the other hand, allowing incoming input to be over-ridden by higher-level information is disfavored because of the resultant risk of hallucination. On this view, the processing of nonwords is not different from the processing of real words; nonword input, too, will simply activate the words that it most closely resembles (with the result that when someone makes a speech error such as a phoneme substitution, we usually understand what they meant to say anyway).

Phoneme detection taps into speech processing not by forming a necessary part of it but by drawing on the continuous output of its multiple stages. In the Merge model (Norris et al., 2000), a dedicated decision-making operation integrates information from bottom-up processing of the incoming speech signal with information from activated lexical representations (and potentially from other sources). Varying task demands can cause different weighting of the information arriving from the various sources. This idea captures the fact that changes in the experimental situation cause variation in the sensitivity of phoneme detection performance to higher-level factors. Thus Cutler et al. (1987) found that listeners detected phoneme targets more rapidly in words than in nonwords when the filler items in the experiment were phonologically varied, but exactly the same target-bearing items showed no word–nonword difference when filler items were phonologically uniform. In other studies, phoneme detection in nonwords proved to be affected by how similar the items were to real words (Connine, Titone, Deelman, & Blasko, 1997; Mirman, McClelland, & Holt, 2005).

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 some cases, targets can then be responded to even if they are not actually present in the signal (e.g., because they have been deleted in casual speech: Kemps, Ernestus, Schreuder, & Baayen, 2004; or assimilated to an adjacent phoneme: Gaskell & Marslen-Wilson, 1998; Hallé, Chéreau, & Segui, 2000). If the target indeed occurs in a real word, and that word was already uniquely identifiable, then orthographic information may directly contribute in the detection response. Dijkstra et al.'s (1995) listeners, whose latencies to detect majority- and minority-spelled phonemes differed at late word positions where only a single viable word candidate remained, may have drawn on fully activated lexical representations in this way. That is, once *papr*- (or *repl*-) had been heard, this sufficed to activate *paprika* (or *replica*), whereupon the lexically stored orthographic information could generate expectation of a [k] before the target actually occurred in the input. This generation process may then be rendered harder by the presence of minority spellings, leading to the RT difference they found.

Obviously, activated words are needed for this generation process to operate, so it will never apply to targets in word-initial position, and indeed both we and Dijkstra et al. found no majority–minority differences in that position. Word-initial phonemes give a view of processing unaffected by the lexical characteristics of the target items that bear them. Regardless of the relative difficulty of generating targets from lexical representations, a given phoneme is no harder to process in a word in which it is spelled in a less usual way than in a word in which it is spelled in the most usual way.

Orthographic information, as our results showed, will in general only be made available in this experimental situation if it appears relevant. Salient manipulation of spelling (lots of filler words like *kneel*, *guardian* and *wrinkle*) apparently causes listeners to assign greater weight to spelling information that is potentially available whenever lexical representations are contacted, and to allow this information to play a role in phoneme decisions. Then, our listeners responded more rapidly to the consistent target phonemes [b, m, t], that are always realized in word-initial position as B, M, T, than to the inconsistent phonemes [f, k, s], that have varying realizations. In the condition where there were no irregularly spelled filler words, but only words like *nip*, *garnish* and *wiggle*, these same target sets did not differ.

Our results do not support models involving mandatory involvement of spelling in prelexical processing. We used exactly the same targets in exactly the same words in both the regular-filler and irregular-filler versions of the materials. Mandatory effects should have been equally apparent irrespective of the rest of the experiment. But this was not so; our filler manipulation caused spelling to become salient, and orthographic effects to appear, in half the experiment only.

We suggest that, in general, spelling may be accorded very low weight in phonemic decision-making. When salience of orthographic realization causes the decision-making mechanism to take spelling into account, this will have little effect on RTs in the detection of consistently spelled targets (in the present case, in words beginning with [b, m, t]), since all activated candidates will have the same orthographic realization of the target. Consulting spelling 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 (and this will be equally true whether the word being heard actually has majority or minority spelling).

What this account does not yet explain is why the phonemic decision-making process should take account of orthography, when the effects of doing so were only negative. Likewise, where effects of spelling mapping occur in lexical-level tasks, they have mostly been negative: spelling-inconsistency made decisions about spoken words more difficult, for example (Peereman et al., 2009; Ziegler & Ferrand, 1998). Strategic responding, however, arises when a clear processing advantage is to be gained. Although the advantage of attending to spelling is not apparent in these laboratory tasks, it is sometimes apparent in real life. Such an advantage is found particularly in situations requiring decisions about phonemes—situations occurring naturally, such as the need to learn a new name. Here paying attention to orthography can facilitate, for instance, the generation of appropriate analogies. The phonemic experimental tasks draw on this existing decision system, and inherit its tendency to respond to spelling salience by taking spelling into account in phoneme decisions.

We note that if involvement of orthography were a mandatory part of auditory processing, the pattern of results from experiments so far would suggest an advantage for languages with a highly regular spelling system. This should presumably have led in turn to historical regularization of systems such as that of English. That this has not happened counts as further evidence against automatic involvement of spelling in speech processing. In speech, we know that our tongues and our lungs contribute in equal measure; mandatory involvement of spelling in auditory word recognition, however, could only disadvantage the recognition of words like *tongue* over words like *lung*. Our finding that such involvement is not mandatory, but strategically motivated, moves the debate forward; we suggest that explaining the naturally occurring strategic motivation should form the next research phokuce.

References

BAAYEN, R. H., PIEPENBROCK, R., & van RIJN, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
CHÉREAU, C., GASKELL, M. G., & DUMAY, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition*, **102**, 341–360.
CONNINE, C. M., & TITONE, D. (1996). Phoneme monitoring. *Language and Cognitive Processes*, **11**, 635–645.
CONNINE, C. M., TITONE, D., DEELMAN, T., & BLASKO, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language*, **37**, 463–480.
CUTLER, A., DAVIS, C., & KIM, J. (2009). Non-automaticity of use of orthographic knowledge in phoneme evaluation. *Proceedings of INTERSPEECH 2009* (pp.380–383). Brighton, UK.
CUTLER, A., MEHLER, J., NORRIS, D. G., & SEGUI, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, **19**, 141–177.
DIJKSTRA, T., ROELOFS, A., & FIEUWS, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology*, **49**, 264–271.
FRAUENFELDER, U. H., SEGUI, J., & DIJKSTRA, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 77–91.

GASKELL, M., & MARSLEN-WILSON, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 380–396.

GONTIJO, P. F. D., GONTIJO, I., & SHILLCOCK, R. (2003). Grapheme-phoneme probabilities in British English. *Behavior Research Methods*, *Instruments & Computers*, **25**, 136–157.

GRAINGER, J., DIEPENDAELE, K., SPINELLI, E., FERRAND, L., & FARIOLI, F. (2003). Masked repetition and phonological priming within and across modalities. *Journal of Experimental Psychology: Learning*, *Memory and Cognition*, **29**, 1256–1269.

HALLÉ, P., CHÉREAU, C., & SEGUI, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language*, **43**, 618–639.

JAKIMIK, J., COLE, R. A., & RUDNICKY, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language*, **24**, 165–178.

JOHANSSON, S., & HOFLAND, K. (1989). *Frequency analysis of English vocabulary and grammar*. Oxford: Clarendon Press.

KEMPS, R., ERNESTUS, M., SCHREUDER, R., & BAAYEN, H. (2004). Processing reduced word forms: The suffix restoration effect. *Brain and Language*, **90**, 117–127.

KOUIDER, S., & DUPOUX, E. (2001). A functional disconnection between spoken and visual word recognition: Evidence from unconscious priming. *Cognition*, **82**, B35–B49.

McCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1–86.

McQUEEN, J. M. (2007). Eight questions about spoken-word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp.37–53). Oxford: Oxford University Press.

McQUEEN, J. M., & CUTLER, A. (1998). Morphology in word recognition. In A. M. Zwicky & A. Spencer (Eds.), *The handbook of morphology* (pp.406–427). Oxford: Blackwell.

MIRMAN, D., McCLELLAND, J. L., & HOLT, L. L. (2005). Computational and behavioral investigations of lexically induced delays in phoneme recognition. *Journal of Memory and Language*, **52**, 416–435.

NORRIS, D., & McQUEEN, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, **115**, 357–395.

NORRIS, D. G., McQUEEN, J. M., & CUTLER, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, **23**, 299–325.

PATTAMADILOK, C., KOLINSKY, R., VENTURA, P., RADEAU, M., & MORAIS, J. (2007). Orthographic representations in spoken word priming: No early automatic activation. *Language and Speech*, **50**, 505–531.

PEEREMAN, R., DUFOUR, S., & BURT, J. S. (2009). Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review*, **16**, 363–368.

SEIDENBERG, M. S., & TANENHAUS, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, **5**, 546–554.

SLOWIACZEK, L. M., SOLTANO, E. G., WIETING, S. J., & BISHOP, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, **56A**, 233–262.

TAFT, M., CASTLES, A., DAVIS, C., LAZENDIC, G., & NGUYEN-HOAN, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language*, **58**, 366–379.

van OOIJEN, B., CUTLER, A., & NORRIS, D. G. (1992). Detection of vowels and consonants with minimal acoustic variation. *Speech Communication*, **11**, 101–108.

VENTURA, P., MORAIS, J., PATTAMADILOK, C., & KOLINSKY, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, **19**, 57–95.

ZIEGLER, J. C., & FERRAND, L. (1998). Orthography shapes the perception of speech. The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, **5**, 638–689.

Appendix: Experimental materials

(a) Target-bearing words, majority spelling.

B: *bucket bandit barnacle basket burnish budget baronet beverage bicycle balance basic bacon;* **M:** *mouldy motive monkey magpie marriage mackerel medical mullet mischievous minstrel molecule manager;* **T:** *turban tangible tiny topical tavern tipsy toddler tunnel towel technical terrible turnip;* **F:** *fumble formula follicle funnel ferry final filament fodder fasten furtive function federal;* **C:** *corner carrot caravan cabin corduroy culprit couple cousin carton canopy consonant cabbage;* **S:** *supper single sickle sonnet seven socket sabotage supple serpent sovereign sapphire silhouette*

(b) Target-bearing words: minority spelling.

PH: *pharmacy phantom physical phony pheasant photo;* **K:** *kidney kerosene kitchen kingdom kangaroo kennel;* **C:** *cedar centre celery cipher circuit citadel*

(c) Filler words: irregular spelling.

aghast almond alphabet amoeba anchor architect asphalt aunt autumn awry beauty benign bomb busy calf calm campaign cello cement certain chagrin chalet champagne chandelier chaos chaperone character charade charisma charlatan chasm chassis chateau chauffeur chauvinist chef chemical chemistry chic chiffon chivalry chlorine cholesterol chorus chrome chronic chronicle cigarette circle circus climb cognac comb comfort cough crate creche crescent debt deny diesel diocese doubt dove dynasty echo folk foreign genre ghastly gherkin ghetto ghost ghoul gnarled gnash gnat gnome guarantee guard guardian gymnast half hasten honest honour hypnotism impugn island issue khaki knack knapsack knee kneel knife knock knocker knoll knotty knowledge knuckle lamb listen llama love lyric machine martyr mechanic mocha morphine mortgage mystery myth naive numb nymph ocean ochre often orchid orphan palm parachute phrase physics pint plumber pneumatic pressure prove psalm psychology pyjamas pyramid quiche rhapsody rhetoric rheumatism rhinoceros rhubarb rhumba rhyme rhythm salmon scenery sceptre schedule scheme scholar school science scissors scythe sign sphere sphinx sugar sure sword talk tomb tongue tough vignette walk wheel whim whistle white wholesale wholesome wraith wrangler wrapper wreath wreck wreckage wren wrench wrestle wretch wretched wriggle wrinkle wrist writhe writing wrong yacht yoghurt

(d) Filler words: regular spelling.

airman alert algebra ankle armadillo aster atheism aver baboon banana beer begin behold bell belong bench bet biology bite blink blunder board boat bolt brand bring brood bunk butter camel campus candid cart chaff chain challenge chamber chancellor channel chapel chapter charter chatter cheese chest chicken chief child chilly chip chipmunk chocolate choice choke chopstick chore chum church chutney cold contract cordon cranium crate damp dash delegate demand diaper dim dive divine doll dome dribble duck dump dune duty dwelling dynamo endless fairy farm feeble fellow finger finish fleet fling flounder flowery foal font fool forbid forest fortunate fresh fret friendly frost fungus funny gain gallop gambit garden garner garnish giddy gift gill girdle gizzard gloom glum goodness grain grape grave groan groove grumpy grunt gulf gun gush gutter habit hack halo hammer happy harm hate hefty heroin hickory hinder history hold home honk hope hover hunting *impute jam jamboree jargon jaw joiner jump jungle kite lake lamp lawn leader lemon lessen letter lid lift ligament liniment list lobby lock lone lonely lump lung lute malaise map marine matter mayor mecca medallion meddle mend mending method million mime mix model mood mosaic muffin music nail name napkin neck nest nimble nine nip noble nod noisy normal north notch note number nut octave opaque open oracle organ pagoda painter paradise parasite park passable pasture peach penicillin pepper pious pitcher place placemat plank plant platform pleasant plucky poem poetry pond popular pork porter potassium praise press prevalent private privilege probable procedure programme prone proper public punish punt pylon radius ramble repeat respectable riddle rivalry rock root rope rosary ruby saga salad salve sample sandal sandy sane satin scoff scrape scribble seal season session shabby shade shadow shaggy shambles shanty shelter shine ship shipping shiver shopping shower shrapnel shunt shuttle sister sling snuff sonata sound soup south spend spill spite spook spout stand stanza store sudden survey swoop talcum tarmac teller thankful thematic theory thermal thicken thong thorn thorny thread thrifty thrill thrush thump time toad tool top topic total transfer tremble trench trial tribe troops trousers trunk truthful tub turbine tyrant valve vanity vest volume vote wager wagon wake wall wallet wallow wander wax wide widow wife wiggle wild willing willow wind wine winter wisdom wisp witch wizard wobble wombat wooden yard yellow yoke zigzag zodiac*

(e) Practice lists.

target /g/:	*super oatmeal fight stem golf*
target /l/:	*condone awake bear bicker grey*
target /v/:	*vast blossom*
target /w/:	*shuffle rain parrot vanish hand eel*
target /v/:	*port nudge singer beggar strand*
target /l/:	*jeans level corrode*