Journal of Psycholinguistic Research, Vol. 35, No. 1, January 2006 (© 2006)
DOI: 10.1007/s10936-005-9004-5

Orthographically Influenced Abstract Phonological Representation: Evidence from Non-rhotic Speakers

Marcus Taft[1]

It is typically assumed that when orthography is translated silently into phonology (i.e., when reading silently), the phonological representation is equivalent to the spoken form or, at least, the surface phonemic form. The research presented here demonstrates that the phonological representation is likely to be more abstract than this, and is orthographically influenced. For example, the claim is made that the word "corn" has an underlying /r/ in its phonological representation, even in non-rhotic dialects. The evidence comes from difficulties observed in judgements about the homophony with a target word of a pseudohomophone whose phonology does not match the putative abstract representation of that word. For example, it is hard to say that the pseudohomophone "cawn" is homophonic with "corn". The conclusion that orthography can shape phonological representation is antithetical to both computational models of the conversion of print to sound and linguistic accounts of phonology.

KEY WORDS: abstract phonology; lexical representation; orthography; phonological representation; pseudohomophone judgement.

INTRODUCTION

A visually presented letter-string can be identified as a word after its orthographic representation is found in lexical memory. This orthographic representation can provide a direct pathway to its pronunciation, and it can also provide a direct pathway to information about its meaning. It has been argued, though, that the pathway from orthography to meaning is typically mediated by the phonological representation (e.g., Frost, 1998; Lukatela, *et al.*, 1993; Van Orden, 1987), with the logic being

The research reported in this paper was supported by a grant awarded to the author by the Australian Research Council.
[1]School of Psychology, University of NSW, Sydney, NSW 2052, Australia. email: M.Taft @unsw.edu.au

that the mental lexicon is accessible through phonology before learning to read and, therefore, reading is parasitic on the phonological system that dominates communication. The assumption appears to be made, then, that the phonological representation that is important in reading is the same as the representation that is used to recognize spoken words, though at a level where phonemes have been abstracted out from the phonetic detail. In fact, such an assumption is formalized in current computational models of reading aloud (e.g., Coltheart *et al.*, 2001; Plaut *et al.*, 1996; Zorzi *et al.*, 1998) where the phonological representation of a word is considered to be its phonemic form.

The argument to be made in this paper opposes this view and suggests that the representation involved in generating sound output from an orthographic input is more abstract in nature than the phonemic form of the word. This idea of an abstract phonological representation goes back a long way in the linguistic literature (e.g., Chomsky & Halle, 1968), where the underlying representation is typically determined from its surface manifestation in morphologically related words. For example, the initial neutral vowel (i.e., schwa) of /sə ′lɑɪva/ (*saliva*), is represented as the full vowel /æ/ in its underlying phonological representation because of its manifestation as an /æ/ in the morphemically related word *salivate*. The proposal to be developed here is rather different to this inasmuch as the major influence on the nature of the abstract phonological representation is claimed to be orthography.

I previously floated this idea about 20 years ago (Taft & Hambly, 1985) on the basis of a task where participants were asked to decide whether or not a specified spoken syllable could be heard at the beginning of a spoken word. All of the words had a schwa as their first vowel (e.g., /lə ′gʊ:n/, i.e., *lagoon*, or /nə ′goʊši:eɪt/, i.e., *negotiate*), whereas the specified syllable had a full vowel (e.g., /læg/ or /nɛg/). It was found that many more erroneous "yes" responses were given when the vowel of the presented syllable matched the orthographic version of the schwa in the target word. So, for example, many more participants thought that /lə ′gu:n/ began with /læg/ than thought it began with /lɛg/, and many more thought that /nə ′goʊši:eɪt/ began with /nɛg/ than thought it began with /næg/. This was true even when there was no morphological basis for the full vowel (i.e., unlike *saliva*). The only reason for making the observed type of error would be if orthographic information had an impact when making the judgement about pronunciation. It was suggested that this might happen because a schwa is represented mentally as a full vowel as indicated by orthography so that, for example, the phonological representation of /lə ′gʊ:n/ is actually /læg′ʊ:n/, with the /æ/ reduced to an /ə/ when stress is assigned in generating the surface pronunciation. Alternatively, however, the orthographic representation of the word may be automatically activated while performing the task (see e.g., Slowiaczek *et al.*, 2003; Ziegler & Ferrand, 1998), with this information being taken into account when attempting to match the spoken syllable to the target word. According to this interpretation, the results say nothing about the nature of phonological representation, despite the fact that the task is a phonological one.

The idea that phonological representation does not correspond to surface pronunciation was revealed in an entirely different setting (Taft, 1984), using a manipulation of morphemic structure. Participants were shown a list of words and were asked to decide for each word whether there existed another English word pronounced identically to it. They were asked not to read the word aloud, but could repeat it to themselves sub-vocally. This task was a relatively easy one when the presented word had the same morphemic structure as the word to be generated (e.g., given *weighted*, it was easy to generate *waited* and hence to say "yes"). However, it proved difficult to think of the homophonic partner of words like *ode* and *mustered* because of incongruity in morphemic structure. That is, while *ode* is monomorphemic, its homophonic partner *owed* is bimorphemic; and while *mustered* is bimorphemic, its homophonic partner *mustard* is monomorphemic. The most straightforward way to perform the homophone judgement task is to generate the phonological version of the presented word and determine whether this is associated with a different orthographic representation to the one that was presented. The only reason that it should matter whether the presented word and the generated word differ morphemically would be if the phonological representation is morphemically structured, and if it is morphemically structured, this means that it is more abstract than its surface pronunciation. For example, *ode* might be represented as /oʊd/ while *owed* is represented as /oʊ/ + /D/ (where /D/ refers to the underlying form of the past tense). Birch *et al.* (1998) came to a similar conclusion on the basis of the same finding with pseudohomophones (i.e., nonwords homophonic with real words, like *glode*, cf. *glowed*).

While the Taft (1984) and Birch *et al.* (1998) results imply the existence of an abstract phonological representation, it could be argued that morphological structure is a special case of abstractness because a polymorphemic word is decomposed within lexical memory (see e.g., Taft, 2004). As such, it does not bear upon the issue of whether monomorphemic forms are represented abstractly. Moreover, although morphemic structure is reflected in the spelling of a word, conclusions cannot be drawn about any orthographic influence upon phonological representation from these studies.

The same thing holds for the finding of Hallé *et al.* (2000) that, although a *b* is pronounced /p/ before an *s* in French (as in *subséquent*), it is harder to detect the phoneme /p/ when it corresponds to such an underlying *b* than when it corresponds to an actual *p* (as in *rhapsodie*). This finding indicates that the phonemic judgement is based on an underlying representation, and it is possible that this underlying representation is shaped by orthography. However, it is alternatively possible that it is determined by morphemic considerations, because a *b* that precedes an *s* is always the final consonant of a prefix (e.g., *sub*).

In one of their experiments, Birch *et al.* (1998) examined monomorphemic items and obtained a result that could potentially be explained in terms of an orthographically influenced underlying representation. The relevant nonword items had a pronunciation that was homophonic with a real word only at the surface phonetic level, mostly as a result of "alveolar flapping." For example, in casual American speech, *battle* is pronounced as if it were spelled *baddle*. Homophony was found to be harder to detect when it only existed at the phonetic level (like *baddle*) than when it existed at the phonemic level as well (e.g., *afrade* homophonic with *afraid*), a result that indicated a level of phonological representation that is not phonetic. While all processing models assume the phonological representation to be phonemic rather than phonetic, examples like *battle* suggest that orthography might have an impact on the determination of the phonemic representation (i.e., in determining that [d] should be represented by /t/ rather than /d/). However, it may alternatively be the case that the phonemic representation /t/ may exist simply because it has not always been experienced as a flapped [d], notably when the word has been carefully enunciated.

The experiment to be reported here, albeit a small-scale one, addresses the question of whether the phonological representation of a word is more abstract than its simple phonemic form, and whether orthography, not just morphology, determines that representation. The study makes use of the fact that Australian English is a "non-rhotic" dialect, meaning that an /r/ is not overtly pronounced when following a vowel. For example, *car* is pronounced /kɑː/ rather than /kɑːr/, and *court* and *caught* are homophonic (both being pronounced /kɔːt/).

The non-rhotic *r* provides an excellent opportunity to test whether phonological representations are orthographically influenced. The suggestion being made is that when a post-vocalic *r* appears in the spelling of a word, it may be represented by an /r/ in the phonological representation of that word even though it does not occur phonemically for non-rhotic speakers. It is only when the surface pronunciation of the word is generated from this underlying representation that the post-vocalic /r/ disappears for the non-rhotic speaker. Thus, it is being suggested that the underlying phonological representation is universal across different dialects of English, with any differences in pronunciation (i.e., accent) arising at the stage of generating the surface output.

A pseudohomophone judgement task was used to examine the possibility that non-rhotic speakers have an underlying post-vocalic /r/ in their phonological representation. The experiment examined the accuracy of making a homophony decision when the presented pseudohomophone and the to-be-generated word (i.e., the “base” word) conflicted in terms of the presence or absence of a post-vocalic *r* in their orthography. For example, if there is an underlying /r/ in the phonological representation of *corn*, then it should be hard to generate *corn* as a homophone of the nonword *cawn*. When presented with *cawn*, participants would need to undertake a phonological transformation (e.g., via grapheme–phoneme conversion rules, Coltheart *et al.*, 2001), and then determine if the output of this transformation matched any phonological representation in lexical memory. The homophony judgement would then be based on the success or otherwise of this matching process. If the phonological representation of *corn* includes an /r/, it is predicted that a “yes” response to *cawn* will be given less often than to a pseudohomophone, like *soke*, whose phonology is more similar to the phonological representation being sought (i.e., /soʊk/, representing *soak*). This is because the phonological representation generated from *cawn* (i.e., /kɔ:n/) would not match the phonological representation stored for *corn* (i.e., /korn/), whereas the phonological representation generated for *soke* would match that stored for *soak* (i.e., /soʊk/).

The above prediction holds primarily when the presented letter-string is not overtly pronounced. If readers do pronounce the nonword aloud and listen to what they are saying, homophony should now be judged on the basis of the surface form of the word. This means that the pronunciation /kɔ:n/, generated from *cawn*, should be recognized as a word through the normal speech recognition system. The abstractness of the internally generated phonological representation should have its maximum impact when it is not dominated by an acoustic signal that can enter lexical memory through the speech recognition system. This idea is tested in the present experiment by having participants initially perform the pseudohomophone judgement task without being allowed to read the letter-strings aloud, and then to perform it again while reading the nonwords aloud and listening to themselves. It is only under the silent conditions that an effect of the mismatched *r* should be observed.

METHOD

Participants

Participants were 18 native Australian English speakers who were postgraduates and undergraduates at the University of New South Wales.

Materials

The items of interest were 18 nonwords whose pronunciation was homophonic with a real word (see Appendix). These pseudohomophones were of two types: Those that conflicted with their base words in terms of the presence or absence of a post-vocalic *r* (the "R" condition, e.g., *cawn* homophonic with *corn*, *forl* homophonic with *fall*), and those that did not conflict with the putative underlying phonological representation of their base word (the "Non-R" condition, e.g., *soke* homophonic with *soak*, *troo* homophonic with *true*). The R and Non-R items were matched on length (with a mean of 4.2 letters for each condition), frequency of their base word (134 vs. 150 per million, according to Carroll *et al.*, 1971), and number of letters differing from their base word, taking position into account (1.78 vs. 2.00).

The 18 items of interest were randomly interspersed with 22 other (Non-R) pseudohomophones, as well as 30 nonwords that were not homophonic with an English word (e.g., *baim*, *snert*). There were also eight practice items, which were a mixture of pseudohomophones and non-homophonic nonwords.

The items were printed in lowercase on individual cards and, apart from the practice items, were presented in a different random order to each participant.

Procedure

Participants were tested individually and told that a series of made-up words would be presented one at a time, some of which were pronounced identically to a real English word and some of which were not. Participants were told to say the letter-string to themselves and then decide whether or not it was pronounced identically to a real word. They were asked not to read the item aloud, but simply say "yes" or "no" according to their judgement. The example of *katt* was given as an item that would require a "yes" response.

Following completion of this task, the items were all presented again and the participants were asked to go through each of them for a second time, but this time saying each letter-string aloud and listening carefully to their pronunciation before responding.

RESULTS

The accuracy data were analysed in a 2 × 2 design where the two factors were Type of Task (i.e., Silent vs. Aloud) and Type of Item (i.e., R vs. Non-R). Both were within-group factors in the analysis of participant means (F_1), but in the analysis of item means (F_2) Type of Task was a within-groups factor while Type of Item was a between-groups factor. The results are presented in Table I.

The main effect of Type of Task was highly significant, with fewer errors being made when the letter-strings were read aloud than when processed silently, $F_1(1, 17) = 71.98, p < 0.001$; $F_2(1, 16) = 27.96, p < 0.001$. Saying the items aloud provided a more comprehensive basis for making a homophone judgement, but in addition, gave a second opportunity to respond correctly to the pseudohomophones. The main effect of Type of Item was also significant, though only in the participant analysis, $F_1(1, 17) = 54.29, p < 0.001$; $F_2(1, 16) = 2.19, p > 0.1$. Of greater importance, however, was the fact that the interaction between the two factors was significant, $F_1(1, 17) = 43.99, p < 0.001$; $F_2(1, 16) = 10.10, p < 0.01$. This came about because the difference in accuracy in saying "yes" to the R and Non-R items was significant in the silent task $F_1(1, 17) = 98.09, p < 0.001$; $F_2(1, 16) = 5.32, p < 0.05$, but not in the aloud task, both F's < 1.

DISCUSSION

It is apparent from this experiment that non-rhotic pronunciation is latently rhotic. That is, the phonological representation of a vowel followed by an *r* is different to another homophonic vowel that does not include an *r*. So, while the phonological representation generated from *soke* matches the phonological representation stored for *soak*, the phonological representation generated from *cawn* does not match that stored for *corn*. The suggestion is that *corn* is represented phonologically as something like /korn/, reflecting the *r* that is evident in the spelling of the word, and it is therefore hard to activate it as a homophone of *cawn*. It was observed that participants failed to detect the homophony of R items even when they repeated the letter-string a number of times sub-vocally. It was only when the acoustic signal

Table I. Mean Percentage Correct for Non-rhotic Speakers Making Pseudohomophone Judgements to R and Non-R Items When Reading Those Items Silently or Aloud

	R Condition e.g., *cawn*, *forl*	Non-R Condition e.g., *soke*, *troo*
Silent	53.7%	79.1%
Aloud	85.7%	87.0%

was overtly generated and fed through the speech recognition system that the surface pronunciation sufficiently dominated the abstract phonological representation to allow accurate performance.

This improvement in performance with overt articulation demonstrates that the R items were truly homophonic with their base words, and that the failure of participants to realize this when performing the task silently indicates that such homophony was hard to establish from the abstract phonological representation alone.

It is not just the stored phonological representation of the word that appears to be orthographically influenced, but also the phonological representation that is generated from the visually presented letter-string. The reason for saying this is that difficulty in detecting homophony occurs when it is the nonword rather than base word that includes the post-vocalic *r*. For example, only four of the eighteen participants thought that *forl* was homophonic with *fall* when the task was performed silently (though all but three changed their minds when performing the task aloud). If *forl* were converted into a phonological representation via standard grapheme–phoneme conversion rules (e.g., Coltheart *et al.*, 2001), the grapheme *or* would be translated into the phoneme /ɔ:/ and would thus allow a ready match to be made with the existing phonological representation of *fall* (i.e., /fɔ:l/). Instead, it seems that an underlying pronunciation is generated from the letter-string whereby the post-vocalic *r* is translated into an /r/ despite its absence from the overt non-rhotic pronunciation.

Instead of saying that the orthography of a word shapes its phonological representation, it might be argued that the reason for incorporating an /r/ into the abstract phonological representation is because it has been heard in other English dialects. For example, the phonological representation of *corn* might include an /r/, not because *corn* is spelled with an *r*, but because it has been frequently heard with an /r/ in the pervasive American media. This suggestion is countered, however, by the high error-rate on those R items where it is the nonword and not the base word that includes the post-vocalic *r*, like *forl*. Here, it is not the presence of an /r/ in the stored representation that creates the problem, but rather the presence of an /r/ in the phonological form that is sub-lexically generated from the letter-string. So if the grapheme *r*, when following a vowel, is translated phonologically into /r/ by a non-rhotic speaker, whether or not this arises from exposure to rhotic dialects, the resultant phonological representation is based upon orthographic information, by definition.

Although this is a small-scale experiment, the implications are quite far-reaching. If phonological information is represented in lexical memory in a form that is not veridical with pronunciation, then attempts that have been made to model print-to-sound computationally are founded on incorrect assumptions (e.g., Coltheart *et al.*, 2001; Plaut *et al.*, 1996; Zorzi *et al.*, 1998). In these models, the only phonological representation that can be generated from orthographic input is identical to the surface pronunciation. There may or may not be problematical issues arising if an intermediate abstract phonological representation is introduced into such models, but at the very least, any changes of this sort need to be entered into a simulation in order to determine their impact on all aspects of the model's functioning. Certainly, the notion of an abstract representation is antithetical to connectionist models that simply learn associations between input orthography and output phonology (e.g., Plaut *et al.*, 1996; Seidenberg & McClelland, 1989) because a specified abstract representation, along with its conversion into a surface form, is not something that would emerge through mere exposure to the statistical characteristics of the language.

It is possible that an early reader behaves like a machine that responds to the input-output characteristics of the language, but that the advanced reader develops a more sophisticated phonological system with abstract representations that mediate between orthography and pronunciation. Such representations may optimize the efficiency of the lexical processing system, perhaps not only for reading, but for orthographic output (i.e., spelling) and spoken language processing as well. Indeed, if orthography influences the nature of the phonological representation, it means that the phonological processing of an advanced reader will be qualitatively different to that of a person who is not literate (see e.g., Castro-Caldas & Reis, 2003; Morais *et al.*, 1979). Moreover, because the orthographic influence would only make sense for alphabetically scripted languages, it is possible that bilinguals whose native orthographic system is not alphabetic may phonologically process the words of an alphabetically scripted language in a non-optimal manner.

CONCLUSIONS

The finding that homophony judgements that involve post-vocalic /r/ are relatively hard to make for a non-rhotic speaker, implies that the phonological representation includes an /r/ even when it is not phonemically revealed. Such a conclusion entails that phonological representations, or at least the ones involved in the processing of visually presented text, are moulded by orthographic considerations. The involvement of orthography in a sound-based task is placed within the phonological system, which is different to saying that orthographic representations are automatically activated when phonology is processed (e.g., Slowiaczek *et al.*, 2003; Ziegler & Ferrand, 1998), though this may also be true. It is hard to see how the non-rhotic /r/ effect in the pseudohomophone task could be located within the orthographic system alone. If, for example, the letter-string *troo* is able to activate the orthographically stored representation for *true*, then *forl* should equally be able to activate the orthographically stored representation for *fall*. It is only in the phonological modality that a mismatch between *forl* and *fall* can have its impact.

However, it must also be acknowledged that with the small number of items in the study, it is possible that certain item-specific factors may provide an alternative explanation for the results. For example, if the task were influenced by feedback from phonology to orthography (e.g., Stone *et al.*, 1997), then it can be argued that *oo* → /uː/ → *ue* is a natural translation while *or* → /ɔː/ → *a* is not, because it is only true when *a* is followed by *ll*. Note, though, that research suggests that such feedback actually occurs at a higher-level than the phoneme-to-grapheme level, namely where the vowel is combined with its following consonants (e.g., Stone, *et al.*, 1997; Ziegler & Ferrand, 1998). At this level, the translation /ɔːl/ → *all* is not at all unnatural. So, any activation of links between phonology and orthography at this level, even if in parallel with phoneme-grapheme links, should allow the ready generation of *fall* from *forl*.

The idea that orthography might mould phonological representation has consequences not just for psychological models of the conversion of print to sound, but also for linguistic accounts of phonological representation. The suggestion that something as unnatural as orthography can play a role in shaping phonological representation is antithetical to any linguistically-based accounts of phonology, regardless of whether they be framed in terms of traditional generative phonology (e.g., Chomsky & Halle, 1968) or more recent optimality theory (e.g., Archangeli, 1997). Yet the present findings imply that orthographic considerations are indeed important, at least if one is wanting to say something about human language performance.

APPENDIX

The Following are the R and Non-R Items with their Pseudohomophone Judgement Accuracy Scores Under Silent and Aloud Conditions.

	Silent %	Aloud %
R items		
cawn (corn)	39	83
fawm (form)	39	94
chork (chalk)	83	100
parm (palm)	61	89
harf (half)	34	56
bort (bought)	61	89
parth (path)	83	100
pawk (pork)	61	83
forl (fall)	22	83
Non-R items		
soke (soak)	100	100
laik (lake)	72	88
sneaz (sneeze)	89	100
gane (gain)	100	100
lyne (line)	100	100
foan (phone)	40	40
thinc (think)	100	100
beme (beam)	39	83
troo (true)	72	72

Base Word in Parentheses

REFERENCES

Archangeli, D. (1997). *Optimality Theory*. Malden: Mass. Blackwell.

Birch, S., Pollatsek, A., & Kingston, J. (1998). The nature of the sound codes accessed by visual language. *Journal of Memory and Language*, *38*, 70–93.

Castro-Caldas, A., & Reis, A. (2003). The knowledge of orthography is a revolution in the brain. *Reading and Writing*, *16*, 81–97.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage Word Frequency Book*. Boston: Houghton-Mifflin.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*, 71–99.

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

Lukatela, G., Lukatela, K., & Turvey, M. T. (1993). Further evidence for phonological constraints on visual lexical access: TOWED primes FROG. *Perception & Psychophysics*, *53*, 461–466.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise simultaneously? *Cognition, 7*, 323–331.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.

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*, *56A*, 233–262.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36*, 337–359.

Taft, M. (1984). Evidence for an abstract lexical representation of word structure. *Memory and Cognition*, *12*, 264–269.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology. 57A*, 745–765.

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language*, *24*, 320–335.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181–198.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*, 683–689.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1131–1161.