Journal of Memory and Language 85 (2015) 116–134

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Letters don't matter: No effect of orthography on the perception of conversational speech ☆

Holger Mitterer [a,*], Eva Reinisch [b]

[a] L-Universita' ta-Malta, Malta
[b] LMU Munich, Germany

ARTICLE INFO

Article history:
Received 3 January 2015
revision received 20 August 2015
Available online 11 September 2015

Keywords:
Spoken-word recognition
Visual-word recognition
Phonological reduction

ABSTRACT

It has been claimed that learning to read changes the way we perceive speech, with detrimental effects for words with sound–spelling inconsistencies. Because conversational speech is peppered with segment deletions and alterations that lead to sound–spelling inconsistencies, such an influence would seriously hinder the perception of conversational speech. We hence tested whether the orthographic coding of a segment influences its deletion costs in perception. German glottal stop, a segment that is canonically present but not orthographically coded, allows such a test. The effects of glottal-stop deletion in German were compared to deletion of /h/ in German (grapheme: *h*) and deletion of glottal stop in Maltese (grapheme: *q*) in an implicit task with conversational speech and explicit task with careful speech. All segment deletions led to similar reduction costs in the implicit task, while an orthographic effect, with larger effects for orthographically coded segments, emerged in the explicit task. These results suggest that learning to read does not influence how we process speech but mainly how we think about it.

Introduction

Our thinking about speech is massively influenced by our ability to read, and we are not aware of this influence. Readers find it natural to think of speech in terms of letter-like segments and often assume that this is universal. It hence came as a big surprise when Morais, Cary, Alegria, and Bertelson (1979) showed that awareness of phonemes does not arise spontaneously. They tested adults that, for social reasons, had not learned to read at a typical school age. One half of these adults was enrolled in a reading class at adult age, the other not. Critically, only participants following the reading class were able to manipulate words at a phoneme level (e.g., perform tasks as "bread minus b is . . . ? " → "red"). Later research revealed a reciprocal relationship between learning to read and phoneme awareness using simpler tasks that also pre-school children can solve to some degree[1]: Bradley and Bryant (1983) devised an "odd-one out" task, in which the question was which word does not fit in a series like "pin, pat, hill, pit". They found that those pre-reading children who perform well in such tasks turned out to be good readers. This has given rise to the idea that spoken and written language processing influence each other.

The link from spoken to written language is obvious. Normal-hearing children invariably learn to speak a

☆ The first author was supported by a University-of-Malta Research Grant. The second author was supported by a grant from the Alexander von Humboldt Foundation. The authors wish to thank Thomas Geyer for the use of the eye-tracker at the Department of Psychology of the University of Munich.

* Corresponding author at: Department of Cognitive Science, University of Malta, Msida MSD 2080, Malta.

E-mail address: holger.mitterer@um.edu.mt (H. Mitterer).

[1] Attempts to do phoneme-deletion tasks with pre-schoolers can lead to frustration. It has happened, for instance, that the question "what is bread minus b" leads the child to answer "then my lunch is gone" and burst into tears.

http://dx.doi.org/10.1016/j.jml.2015.08.005
0749-596X/

language before they learn to write it. The better the oral language is processed, the easier it is to link written language to it (Melby-Lervåg, Lyster, & Hulme, 2012). The link in the other direction is somewhat less straightforward and more controversial. In this paper, we will present data that may force a re-interpretation of the relation between learning to read and speech perception. As a consequence of the paper by Morais et al. (1979), it has become an underlying assumption that learning to read makes us better at perceiving speech—where "better" means "more segmental". This far-reaching interpretation of the phoneme-awareness data is evident when Dehaene et al. (2010, p. 1362) spoke of "the enhanced phonemic processing that accompanies reading acquisition" or when Pattamadilok et al. (2009, p. 169) argue "Thus it is possible that learning to read is crucially involved in developing fine-grained phonological representations." It has even been suggested that orthographic representations are activated online during speech perception (Ziegler & Ferrand, 1998). Contrasting with this theme, we will argue that in everyday speech processing, confronted with conversational speech in a natural task setting, the role of orthography is negligible. Based on the present finding we suggest that learning to read may only influence meta-linguistic thinking about speech and in fact make us "deaf" to the properties of normal conversational speech.

In the literature, there have been proposals of "on-line" and "off-line" influences of learning to read on speech perception. Off-line, or indirect, influences may arise due to exposure to stylistic language variation that comes with reading. More experience with a wider variety of texts, for instance, seems to influence a listener's ability to predict upcoming words (Mishra, Singh, Pandey, & Huettig, 2012). It has also been argued that reading is important for the expansion of the mental lexicon during elementary school (Stanovich & Cunningham, 2001). Reading is considered important here because infrequent words are more likely to occur in texts than in spoken language (Hayes, 1988). As a consequence, infrequent words are more often encountered during reading than during oral language use. As such, reading will influence listening by expanding the lexicon and changing the number of candidate words that may fit a given input.

Extending this line of thought, it has even been claimed that only with sufficient vocabulary growth through reading do children gain access to phoneme-like units in speech perception (Metsala & Walley, 1998). Following this lead, Dehaene et al. (2010) tested the brain activation patterns of literate and illiterate participants, matched on socio-economic status, during various tasks. They found that literate participants showed an increase of brain activation during listening in the superior posterior temporal gyri compared to illiterate participants. This was interpreted as "enhanced phonemic processing which accompanies alphabetization" (p. 5). Interestingly they cited Morais et al. (1979) as additional evidence for a more phonemic processing of speech although Morais et al. (1979) only showed that *meta-linguistic abilities* change drastically with reading acquisition. In this context, it is also important to note that in models of speech processing in the brain the superior posterior temporal cortex is not part of the core speech perception system but a secondary path that seems to be involved in linking speech sound to articulation (Scott & Wise, 2004). Moreover, recent evidence suggests that the implicit processing of speech does not get "more phonemic" with reading. McQueen, Tyler, and Cutler (2012) showed that pre-school children, before they had learned to read, are able to make use of phoneme-like units in speech perception: When they learn that a given speaker produces /f/ in slightly /s/-like way, they are able to generalize this to new words, which necessitates the assumption of a pre-lexical phoneme-sized unit. Moreover, it has also been shown that dyslexics also show no deficit in such phonemic processing (Groen & McQueen, 2014; Mitterer & Blomert, 2003). This suggests that phoneme-like units are not the consequence of learning to read but are used for speech perception independently of reading experience.

Another off-line influence of reading on speech perception can arise when exposure to written words influences the lexical–phonological representations of specific words. Racine, Bürki, and Spinelli (2014) investigated the processing of French words in which the written form suggests the presence of a schwa vowel that, in spoken form, can be deleted either optionally or obligatorily. When these words were presented auditorily with and without schwa to pre-readers in a recognition task, the results showed a simple spoken-word frequency effect: For the words with optional deletion, reactions were faster to the version with schwa; for words with obligatory deletions, reactions were faster to words without the schwa – in both cases matching the respective more frequent spoken form. However, in contrast to these effects for pre-reading children, for beginning readers (aged 9–10 years) the results showed an overall "boost" for the schwa-bearing forms – matching the orthographic representation. This provides evidence that reading a word can influence the phonological representation (Bürki, Ernestus, & Frauenfelder, 2010; Bürki & Gaskell, 2012; Connine, Ranbom, & Patterson, 2008). This may not be surprising, if one assumes that reading actually involves phonological recoding, so that reading a word with a schwa leads to "implicitly hearing" the same word with a schwa. The reading experience may hence influence the phonological representation of this word, since independent evidence indicates that phonological representations are sensitive to input frequencies of variants (Connine et al., 2008; Pitt, 2009). Implicitly hearing a word during reading may thereby also influence listening by changing the phonological representation but without necessarily activating an orthographic representation during listening. This possibility gains credibility given the evidence that words learned during reading also seem to be added to the mental lexicon for spoken words (Bakker, Takashima, van Hell, Janzen, & McQueen, 2014).

A similar influence of the orthographic form on developing lexical representations has been shown for second language learners (Escudero, Hayes-Harb, & Mitterer, 2008). Escudero et al. trained Dutch learners of English to associate a set of novel English (nonsense) words containing either /æ/ or /ɛ/ with novel shapes. Critically, the sound contrast between /æ/ and /ɛ/ is difficult to distinguish for Dutch learners. This was reflected in the results, as participants who learned these words just auditorily seemed to encode them as containing the same vowel. One group of learners, however, also saw orthographic representations of these words, and the results indicated that these words were represented as having phonologically different vowels. This indicates that seeing orthographic word forms in a second language influences the phonological representations of these words.

The evidence for off-line influences of reading on spoken-word recognition hence seems solid. A more far-reaching claim, however, is that orthography has an on-line influence on speech perception, because orthography is activated routinely even in purely oral-language tasks, including natural interactions. Evidence for such an influence of orthography on speech processing comes from studies showing that auditory lexical decisions are easier for words with consistent than inconsistent sound-to-letter (henceforth, S2L) relationships (Ziegler & Ferrand, 1998). Subsequent research (Ziegler, Ferrand, & Montant, 2004) showed that such effects are replicable, but are not significant in an auditory naming (i.e., shadowing) task. This would seem to indicate that the effect of orthography does not arise automatically in spoken-word recognition, otherwise, the effect should be observed in naming as well.

A typical finding in this area is that effects are easier to demonstrate in by-subject analyses than in by-item analyses (see, e.g., Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008). Since a given word can only be consistent and inconsistent with its orthography, S2L consistency is necessarily a between-item variable, which also means that, technically, all studies on orthographic consistency are not experimental but quasi-experimental in nature. In response to this situation Rastle, McCormick, Bayliss, and Davis (2011) used a learning paradigm, in which participants learned novel words. By using novel words, S2L consistency could be varied experimentally within-items. That is, for each word a consistent and an inconsistent orthographic label was used and consistency was varied for each item over participants. With this design, they found that, after training, auditory lexical decision was slower for inconsistent than for consistent words. However, no effect of S2L consistency was found in shadowing, replicating the earlier finding with existing words (Ziegler et al., 2004). This suggests that the effect of S2L consistency might only arise in a meta-linguistic task.

This potential explanation seems to be ruled out by a number of ERP studies (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Perre, Bertrand, & Ziegler, 2011; Perre, Pattamadilok, Montant, & Ziegler, 2009). These studies compared the ERP response to S2L consistent and inconsistent words under various conditions, using a variety of tasks from lexical decision (Perre et al., 2009) to a go–no go task (Perre et al., 2011), in which a button had to pressed in response to a noise stimulus. In the latter case, participants did not have to make a decision on the critical words; nevertheless, inconsistent words still lead to a more negative ERP than consistent words. Additionally, these experiments show that "early" inconsistencies (i.e., inconsistencies in the first syllable of a word) lead to an earlier negativity in the ERPs (again, comparing inconsistent with consistent words) than "late" inconsistencies.

This latter finding is also useful to illustrate how an orthographic effect on speech perception could be envisioned. In a post-lexical account, an orthographic representation is activated once the auditory word is recognized or has passed a threshold of activation. Such an effect would hardly be an effect "on" speech perception, as the orthographic effect would be a *consequence of* speech perception. Orthographic representation would only become available after the hard problem in spoken-word recognition, the invariance problem, has already been solved.[2] Note that a time-course difference between early and late inconsistencies, as found in the ERP data reviewed above, cannot be explained by such a model, since noting the early inconsistencies early requires a pre-lexical influence of orthography, in which pre-lexical phonological representations activate orthographic representations (that is, hearing a /t/ activates the letter *t*).

The ERP evidence hence suggests that orthographic effects are ubiquitous and probably automatic. However, due to the tradition of ERP research, no statistical test is used that shows that the effects are consistent over items. This problematic for two reasons. As reviewed above, consistency effects tend to be much clearer in subject analyses, so that the ERP evidence is questionable in terms of the statistical conclusion validity in the absence of a test that shows consistency over items. Secondly, S2L inconsistencies do not occur randomly, but are often the consequence of sound change (e.g., due to /r/ dropping, *soar* and *saw* became homophonous in British English). Words with inconsistencies are therefore anything but randomly selected, and, in the absence of random selection, causal inferences are notoriously difficult.

Others indeed argued that orthography influences mainly meta-linguistic thinking about speech (Cutler & Davis, 2012): S2L relations then only influence the decision component in laboratory tasks. For instance the quick activation of a visual-word image during auditory word comprehension may reinforce a "yes" response in the lexical decision task. This would explain why effects of S2L consistency are found in auditory lexical decision but not in shadowing tasks (Rastle et al., 2011). The latter task does not involve a meta-linguistic decision.

Importantly, S2L inconsistencies appear at two levels but mainly one level has yet been scrutinized. Inconsistencies can arise in the orthography, with English being the prime example of "deep" orthography with many inconsistencies, evidenced in pairs like *mint-pint* which should rhyme according to their orthography but in fact do not rhyme. This has been the main focus of research up to now (Rastle et al., 2011; Ziegler & Ferrand, 1998; Ziegler et al., 2004). It is noteworthy that such inconsistencies are not necessary. With a shallow orthography, such effects are avoidable,[3] and the depth of an orthography

[2] It may be argued that orthography may then still influence speech perception by top-down connections from the lexicon on pre-lexical representations. However, convincing evidence for such top-down effects is still lacking (McQueen, Jesse, & Norris, 2009).

[3] Spelling reforms can massively reduce such S2L inconsistencies, even for English. This is illustrated by the use of English loans in Maltese, which are adapted to the Maltese S2L relations (e.g., *orrajt*, *mowbajl*, and *kju* for *alright*, *mobile*, and *queue*).

has a massive influence on how difficult it is to learn to read (Aro & Wimmer, 2003).

However, another type of inconsistency arises in natural interactions, the type of usage that language mostly has evolved for (Dunbar, 1998). In such natural forms of speech, the phonological form of the word may be inconsistent with the spelling as a consequence of deletion and reduction of phonemes in the spoken form. The word *yesterday* may surface here as "jeshay", massively inconsistent with its spelling. Although we produce and perceive such forms all the time, naïve speakers—readers or not—are happily unaware of their existence (Ernestus, 2013). Notably, variant forms are not restricted to high-frequency words such as *yesterday*; even low-frequency words have on average more than five different possible phonemic transcriptions in normal spontaneous interaction (Keating, 1997). These alternative pronunciations are anything but marginal in conversational speech; they are ubiquitous. Every other word in conversational speech has a "letter" changed or missing (Johnson, 2004) and is therefore S2L inconsistent.

This raises the question how such reduced words are recognized. While this question has attracted a lot of attention lately, it is not yet fully resolved. One consistent finding is that words are recognized less efficiently when they are reduced (Brouwer, Mitterer, & Huettig, 2013; Ernestus, Baayen, & Schreuder, 2002; Janse, Nooteboom, & Quené, 2007). Interestingly, the benefit for full over reduced forms remains even when the reduced form is the more frequent one (Ranbom & Connine, 2007). However, the frequency of the reduced form matters, as the reduction costs lessen the more frequent the reduced form is (Mitterer & Russell, 2013; Ranbom & Connine, 2007). This suggests that reduced forms are in some form stored in the mental lexicon.

These data on the perception of reduced forms rule out a very simple matching process in which the incoming form activates the best matching word (cf. Ernestus, 2014). Another potential account would be that words are recognized without using pre-lexical abstraction, so that words are stored as grainy spectrograms (Goldinger, 1998; Pierrehumbert, 2002), and there would be exemplars of both full and reduced forms stored in the mental lexicon to which the input would be matched. This account, however, is challenged by the finding that pre-lexical abstraction contributes to the recognition of reduced forms. Poellmann, Bosker, McQueen, and Mitterer (2014) presented Dutch listeners with multiple /b/-initial words, in which the initial /b/ was reduced to an approximant (e.g., they hear the Dutch word /bəxɔnə/, Engl. *started*, as [ʋexɔ nə]. This experience made it easier for these listeners to later recognize *other* words that had the same reduction compared to a control group that had heard the earlier words in unreduced form. This indicates two things. First of all, listeners make use of prelexical abstraction and are thus able to generalize from one word to another (which argues against a purely episodic-storage model, see Cutler, Eisner, McQueen, & Norris, 2010). Secondly, listeners make use of letter-sized segments in pre-lexical speech processing, even though segments are not as reliably present as often thought. As such, the comprehension of reduced forms is in principle conceivable in classic word-recognition models such as TRACE (McClelland & Elman, 1986) or Shortlist (Norris, 1994), even though some modifications would have to be made, like the use of multiple phonological forms for a given word. Because these models feature letter-sized segments in pre-lexical processing, it is in principle possible that, despite phonological reductions, there is a pre-lexical cross-talk between phonological and orthographic representations in speech perception.

Critically, S2L inconsistencies due to reductions are inherent to speech, and, unlike the deep-orthography effect, they cannot be solved by spelling reforms. Given the massive amount of S2L inconsistencies in conversational speech and their unavoidability, an orthographic influence in speech perception would hence come with a huge burden. If an S2L inconsistency slows down word recognition, the recognition of every other word in a dialogue would be additionally slowed down. This, in turn, questions the assumption that orthography should play a role in speech perception in natural listening situations. Activating orthographic forms during speech processing even though they are quite often inconsistent with the spoken input would be a bad design feature that would make spoken-word recognition inefficient in a dialogue. Note that participants in a dialogue must be able to function with a high level of temporal precision (De Ruiter, Mitterer, & Enfield, 2006), so that an additional burden for word recognition would be quite costly.

Therefore, we tested the role of orthographic coding for words when they occur in reasonably natural speech. Importantly, by experimentally controlling the nature of deletions we can test whether orthography matters during the perception of conversational speech. This is possible because not all speech sounds are coded equally well in orthography. A sound that has no orthographic counterpart is the German glottal stop (/ʔ/) that typically appears at the onset of orthographically vowel-initial words (e.g., *Affe* [ʔɑfə], *ape*) (Kohler, 1996). Under the assumption that S2L inconsistencies influence speech perception, the deletion of glottal stop in German should have smaller reduction costs than the deletion of other segments.

This raises the question which orthographically coded segments should be compared with the unscripted German /ʔ/. We here present two comparisons. We compare German /ʔ/ with German /h/ (grapheme "h"), that is, we compare the deletion costs of two different segments in the same language. The choice for German /h/ was guided by the finding that reduction costs are lower for reductions that occur frequently than for rarely occurring reductions (Mitterer & McQueen, 2009). A survey of the Kiel Corpus of Spontaneous Speech indicated that in word-initial position, the most frequent position of German /ʔ/, only one other segment is deleted with any consistency, namely /h/ (deleted rate 8.7%, compared to 32% for /ʔ/). If, in addition we control for lexical stress – a major constraining factor for segment deletions (see Mitterer, 2011) – /ʔ/-deletion is reduced to 7.3% and /h/ to 7.6% – numbers that can be deemed comparable.[4] This is in line with impressionistic

[4] This huge change is due to the fact that many unstressed function words are/ʔ/ initial (e.g., *ich*, *und*, Engl. *I*, *and*), which inflates the raw deletion rate for /ʔ/.

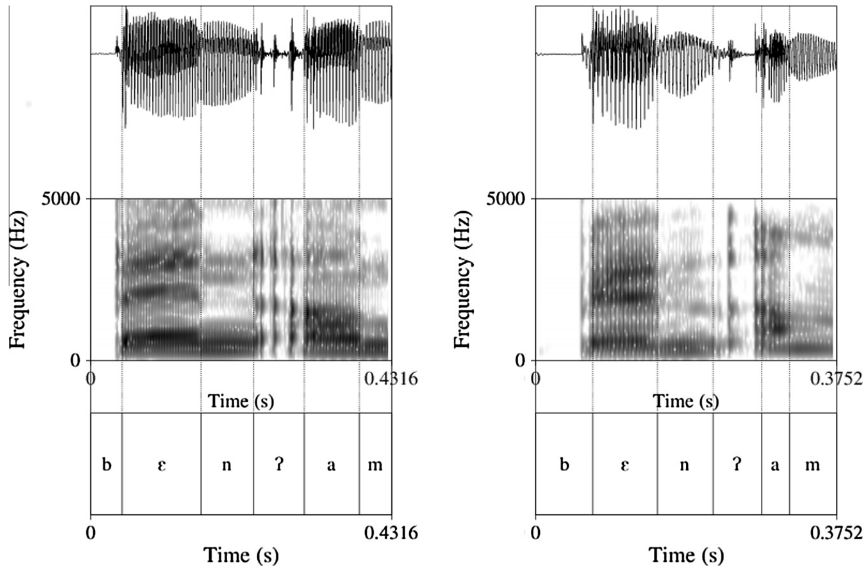

Fig. 1. Examples of the sequence [bɛnʔam] in Maltese (left panel, Engl., *Ben* [proper name] *rose*) and German (right panel, Engl. *Ben at*). Note that this was not an experimental item, since the experimental items did not contain examples of two phonetically similar words in both languages.

claims that /h/ and /ʔ/ behave similarly in German (Wiese, 1996).

The comparison of /h/ and /ʔ/ brings us to another relevant issue: the status of /ʔ/ in German. There are different opinions whether German /ʔ/ should be considered a phoneme (Maas, 1999; Wiese, 1996). The majority seems to argue that /ʔ/ is not a phoneme, because it can be inferred by a rule (prefix any vowel-initial foot with a glottal stop, see Wiese, 1996, p. 58–61). One huge problem with this account is that the predictability of /ʔ/ is contingent on granting phonemic status to /h/. The decision to grant phonemic status to /h/ rather than /ʔ/ is essentially an arbitrary one, since the Phonology of German could be re-written by assuming that /ʔ/ is a phoneme and predict the presence of /h/ by the same rule that now applies to /ʔ/ (insert /h/ to any now vowel-initial foot). How this would work can be explained with an analogy to orthography, where the same reasoning applies. Currently, the word /ʔaus/ is written as *aus* (Engl., out) while /haus/ is written as *Haus* (Engl., *house*). That is, the glottal stop is not coded by a separate letter while /h/ is. However, it would also be possible to write /ʔaus/ as *qaus* (*q* is often used as a grapheme for the glottal stop cross-linguistically) while /haus/ is written as *Aus*[5] and the insertion of aspiration, rather than glottalization, has to be inferred for orthographically vowel-initial words. Since we test the consequences of presenting /ʔ/- and /h/-initial words with and without the initial segment present, our data hence also have bearings on the issue of the relative weight of /h/ and /ʔ/ from a phonological point of view. If /h/ is a phoneme and /ʔ/ is not, this would also predict greater costs for deletion of /h/ than /ʔ/.

However, comparing two different segments leads to unavoidable confounds. It may be the case that /h/ is acoustically less salient than /ʔ/. It has been argued that reductions in conversational speech tend to be constrained by perceptual salience, so that reduction occurs more often for segments that are less salient (Hura, Lindblom, & Diehl, 1992; Mitterer, Csépe, Honbolygo, & Blomert, 2006; Steriade, 2001). It is also conceivable that the reduction of /h/ may be less salient than the reduction of /ʔ/ so that the additional reduction costs caused by the orthographic inconsistency are counteracted by perceptual factors. Indeed, /h/ has been argued to be perceptually weak and therefore often deleted (Mielke, 2002), which may counteract any orthographic effect (which should be larger for /h/ than for /ʔ/). This issue is addressed by comparing deletion of /ʔ/ in German to the deletion of /ʔ/ in a language in which it is coded orthographically.

As such a language, we used Maltese, which is a Semitic language spoken on the Mediterranean island of Malta, south of Sicily. In Maltese, the glottal stop is scripted as "q" and also has an undisputed phonemic status. In contrast with the German glottal stop—and German /h/—the Maltese glottal stop has few phonotactic constraints regarding the position it can occur in. It occurs in onset (*qattus* /ʔɑttus/ Engl., cat) as well as in coda position (*triq* /triʔ/, Engl., *street*) and clusters with all other types of consonants, even independent of phonological voicing (*dqiq*, /dʔiʔ/, Engl. *flour*, *qtates* /ʔtates/, Engl., *cats*). Fig. 1 provides examples of the glottal stop in Maltese and German. If orthography influences speech perception, the deletion of the orthographically coded Maltese glottal stop should have stronger consequences than the deletion of the German glottal stop, despite their phonetic similarity.

Experiment 1

In order to measure the reduction costs of /h/ and /ʔ/, we used a visual-world eye-tracking paradigm that has repeatedly been shown to reflect reduction costs in a

[5] The capitalization follows German orthographic rules which requires that nouns should be capitalized.

graded manner (Mitterer & McQueen, 2009; Mitterer & Russell, 2013). Additionally, the visual-world paradigm has the advantage of probing language comprehension in a fairly natural way, as natural interactions often require finding a referent in the visual environment (e.g., "Can you hand me the salt please?").

To implement this task, we selected German and Maltese picturable nouns that are either /h/- or /ʔ/-initial and presented them in sentences. These sentences contained typical discourse markers to convey a conversational speech style. This is exemplified in (1) and (2) where a German sentence in a formal style is contrasted with a respective informal version of the sentence, using the target *Ampel* /ʔampəl/ Engl., *traffic light*.

(1) Formal sentence: Er fuhr trotz der roten Ampel über die Kreuzung
Word-by-word translation: he drove despite the red light over the junction
(2) Conversational: Der ist echt trotz der roten Ampel über die Kreuzung gefahren
Word-by-word translation: This one has really despite the red light over the junction driven
Translation: This guy really drove over the junction despite the red light

Psycholinguistics has a tradition of using formal sounding sentences, as they would typically occur in written language, but that would be unlikely in spoken language (e.g., "the secret was whispered", see Sammler et al., 2013). However, here we intended to approximate an ecologically valid situation by using sentences that contained discourse markers and contractions as they occur in spontaneous speech (such as German *an dem* → am, Engl., *on the*, wir *haben* → *wir ham*, Engl., we have).

All sentences were presented with and without /ʔ/ (German and Maltese) and /h/ (German only) respectively, and we measured how well participants' eye-movements converged on the respective target pictures. If the orthographic coding plays a role for target recognition, then for German the absence of the orthographically coded /h/ should weigh heavier than the absence of /ʔ/. In addition, the absence of the orthographically coded /ʔ/ in Maltese should weigh heavier than the absence of uncoded /ʔ/ in German. Importantly, we also allowed for participants to reject that an item was present by, first, having filler trials in which none of the visual objects was mentioned in the sentence. Second, participants were instructed to click on an empty part of the screen if none of the visual objects was mentioned in the sentence.

Method

Participants

22 native speakers of German participated in the experiment. They were students or junior research staff at the University of Munich Additionally, 22 native speakers of Maltese, students at the University of Malta, participated. They were aged 18–32 and reported no hearing problems. They were paid for their participation.

Materials

For the auditory and visual materials, we identified 37 German /ʔ/-initial nouns and 34 German /h/-initial nouns that were pictureable, plus 48 pictureable filler items starting with other consonants. Each noun was embedded in a carrier sentence that allowed some prediction of the upcoming target (see Table 1 for an example). For the Maltese condition, we identified 36 Maltese picturable /ʔ/-initial nouns, plus 48 pictureable filler items starting with other consonants. For each noun, a sentence was generated.

For each sentence a visual display containing three pictures was generated. Pictures were retrieved using Google image search and two native speakers chose the best fitting ones. One exception was *qassata*, a typical Maltese pastry, for which a photo was taken since the online search did not return a good match. For each target, two additional objects were selected, one of which could also fill the slot in the sentence (e.g., for the German target *Eimer*, Engl., *bucket*, the sentence was *There is a hole in the*… and *tyre* was selected as competitor; note that German uses the word *Loch* for the concept *puncture* as well). For half of the filler items, none of the three pictures on the screen matched the sentence. Moreover, the sentences contained discourse markers (such as *like, well, etc.*) and contractions (*you've* rather than *you have*) to convey a conversational style. These sentences were recorded by native female speakers of German and Maltese, who were instructed to speak the sentences as if talking to a friend in a pub.

Filler items were recorded once. For the critical target items, the sentences were recorded at least three times, and more often if the speaker had problems to produce

Table 1
Explanation of the cross-splicing for the critical items (example for German). The presented stimuli contained the base sentence (italics) with the critical juncture cross-spliced from an "initial-segment deleted" utterance (bold) or an "initial-segment realized" segment (underlined).

Recording	Content					
Base recording	Orthography	Da	ist	ein	Loch	im Eimer
	IPA	*da*	*ɪs*	*ɑin*	*lɔx*	*i*m ʔɑi*m*ɐ
Deleted recording	IPA	*da*	ɪs	ɑin	lɔx	i**m ɑi**mɐ
Realized recording	IPA	*da*	ɪs	ɑin	lɔx	im ʔɑimɐ
Stimulus	**Content**					
Deleted	IPA	*da*	*ɪs*	*ɑin*	*lɔx*	*i***m ɑi***mɐ*
Realized	IPA	*da*	*ɪs*	*ɑin*	*lɔx*	*i*m ʔɑi*mɐ*

the deletion. Note that conscious deletion—especially of the unscripted glottal stop—can be difficult for speakers, since they tend not to be aware of it (for more discussion on this point, see the General discussion).

The recordings were then prepared for the experiment as follows. The filler sentences were used in their unaltered form. For the experimental items, we used cross-splicing. For each sentence, three recordings were selected to generate the experimental stimuli. One recording, which was produced fluently and in a conversational style was used as a base, and the critical juncture (see Table 1) was cross-spliced from two other recordings, once with and once without the initial segment. Since all targets were embedded after voiced segments, a continuous f0 contour was taken as a clear sign that no glottal gesture was present, and such items were used for the segment-deleted version. All target sentences were cross-spliced and the only difference was the presence or absence of the initial segment.

Participants were presented with each target item once. The number of critical items with and without the initial segment was counterbalanced across participants. A different random order was generated for each participant that balanced the number of times each quadrant contained the targets over participants and conditions.

Procedure

Participants were seated in front of a computer screen and an Eyelink SR 1000 eye-tracker in desktop set-up was calibrated (German participants were tested at the Department of Psychology at the University of Munich; Maltese participants at the Department of Cognitive Science at the University of Malta). They were instructed that they would see three pictures scattered over the four quadrants of the computer screen and hear a sentence over headphones. They were asked to click on one of the pictures if it was mentioned in the sentence. If none of the objects on the screen was mentioned in the sentence, they should click on the empty quadrant. Such target-absent trials were added to allow participants to reject critical items in which the initial segment was missing.

Design and analysis

Since the comparison of German /ʔ/ versus German /h/ is a within-participant analysis, and the comparison of German /ʔ/ versus Maltese /ʔ/ is necessarily between participants, we present two separate analyses. In both analyses the independent variables were Segment (German /ʔ/ vs. German /h/ or German /ʔ/ vs. Maltese /ʔ/) and Deletion (initial segment deleted versus present). The dependent variables were fixation proportions on the target item and the acceptance rate and latency derived from the click responses. With acceptance rate, we mean that participants actually click on the intended object. We use the term acceptance rather than error rate, since participants may have reasons to say that a word was not produced if its initial segment is missing.

All results were analysed with linear-mixed effects models with participant and item as random factors and a maximal random effect structure (Barr, Levy, Scheepers, & Tily, 2013). In the analysis of the acceptance rate, we deviated from the maximum random effect structure due to convergence problems, as noted in the relevant analysis. For acceptance rate, a logistic linking function was used (using the function glmer from the package lme4, v1.1.7); RTs were log-transformed, and fixation proportions were transformed into logOdds for the analysis. Degrees of freedom were estimated using the R package lmerTest (v2.0.25). Note that for the analyses of fixation proportion and acceptance rate, the regression weights can be interpreted as effect sizes.

In eye-tracking research, as in ERP research, the decision of which time windows to analyse is critical. As in the present study the focus is on the (possible) influence of orthography on spoken words, we used a time-window of 200–700 ms after word onset, which reflects the initial stages of word recognition (Allopenna, Magnuson, & Tanenhaus, 1998). An orthographic influence after the word has been recognized would hardly constitute an orthographic influence on spoken-word recognition but rather constitute a post-lexical effect.[6] This choice is also in line with the claim that early inconsistencies mainly lead to early effects in word recognition (Perre et al., 2011). Since our inconsistencies are at word onset, they should be reflected in an early-time window.

We also took into account the estimated word frequency using the SUBTLEX-D corpus (similar to SUBTLEX-NL, see Keuleers, Brysbaert, & New, 2010) for "spoken" (based on frequency of a word in media subtitles) and written forms (Google Books corpus, also provided by the SUBTLEX database). Only written frequencies were used for Maltese, since there is no corpus that allowed us to assess spoken frequency (note that there is no subtitling on Malta, so that it is not possible to generate a corpus analogue to the Methods of Keuleers et al., 2010). For written frequencies, we used the Maltese Language Resource Server (MLRS). Word frequencies were used as log frequency per million and centred to sum to zero for the analysis (see Baayen, 2008, for a rationale why centering is important in such cases).

Results

German /h/ versus German /ʔ/

The first two rows of Table 2 show the parameters of the acceptance rate and RT measures. Fig. 2 shows the eye-tracking data for trials in which the participants clicked on the intended item. The results show clear reduction costs with fewer clicks on the intended target, longer reaction times and fewer fixations on the critical targets if the initial segment was deleted. These observations were supported by statistical analyses: The analysis of acceptance rate showed a marginal effect of Deletion ($B = 0.86$, $SE(B) = 0.47$, $p = .07$), but no effect of Segment ($B = 0.40$, $SE(B) = 0.92$, $p > .2$) and no interaction ($B = -0.23$, $SE(B) = 0.94$, $p > .2$). The analysis of latencies showed the same pattern (Deletion: $B = -0.045$, $SE(B) = 0.018$, $t(53) = -2.45$, $p < .05$, Segment: $B = -0.027$, $SE(B) = 0.060$, $t(70) = 0.46$;

[6] As a reviewer requested the analysis for a later time window, we present the analogue analysis for a time window 700–1200 ms in Appendix B.

Table 2
Acceptance rates and latencies of click responses in the visual-world task depending on target word and actual production (with or without the first segment), with standard deviations by participants in brackets. (Note that standard deviations for percentages should be interpreted cautiously, especially when this leads to expectancies that some participants perform better than 100% correct, as in the current one). The results show deletion costs in all cases, even though German /ʔ/ has no orthographic counterpart. The exception is the latency measure for Maltese /ʔ/, in which the effect is likely washed out due to the overall much longer latencies.

	Acceptance rate Initial segment		Latency Initial segment	
	Present	Absent	Present	Absent
German /h/	96.8% (4.7%)	94.3% (5.3%)	1317 (270)	1361 (221)
German /ʔ/	97.2% (3.7%)	95.8% (4.0%)	1325 (275)	1349 (1292)
Maltese /ʔ/	93.5% (5.1%)	86.6% (10.0%)	2044 (370)	2047 (369)

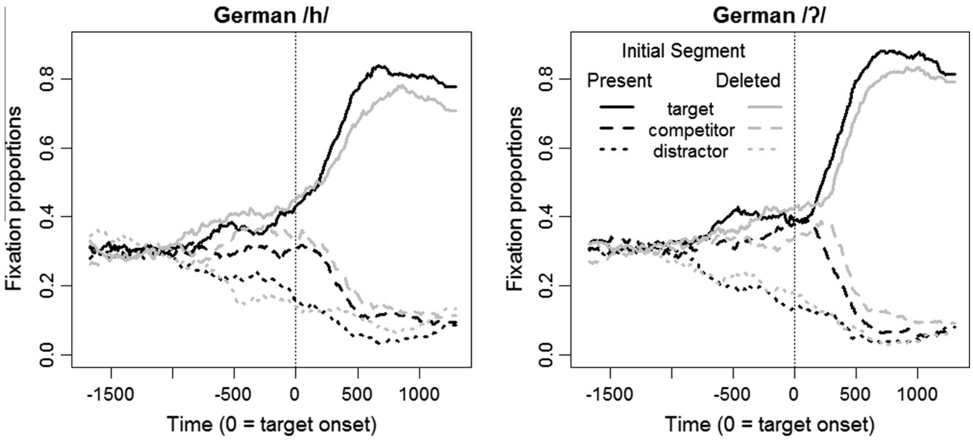

Fig. 2. Fixation proportion for the German participants to pictures with /h/-initial (left panel) and glottal stop initial (right panel) targets, depending on whether the initial segment of the target is produced or not.

Segment ∗ Deletion: $B = -0.004$, $SE(B) = 0.037$, $t(26) = -0.12$). Both of these patterns were unchanged when Written Frequency (deletion: Deletion: $B = -0.045$, $SE(B) = 0.018$, $t(54) = -2.46$, $p < .05$; Segment ∗ Deletion: $B = -0.001$, $SE(B) = 0.01$, $t(27) = -0.15$) or Spoken Frequency (deletion: Deletion: $B = -0.045$, $SE(B) = 0.019$, $t(54) = -2.45$, $p < .05$; Segment ∗ Deletion: $B = -0.004$, $SE(B) = 0.03$, $t(27) = -0.15$) was added as predictor, including an interaction of frequency and reduction (all of which also were not significant, $p_{min} > .2$).

For the analysis of the eye-movement patterns, we first tested whether either of the targets (i.e., /ʔ/ or /h/-initial) was better predictable from the sentence context by looking at the time window from 600 ms before target onset until target onset. Fig. 1 shows that indeed already before word onset participants looked more at the target and/or the second item fitting the sentence context (i.e., the semantic competitor) than at the third, unrelated, item. But there was no differential prediction of the target depending on Segment (i.e., /ʔ/ vs. /h/; $B = -0.04$, $SE(B) = 0.28$, $t(43) = -0.177$). This shows that our /ʔ/ and /h/ targets were comparable. To test the effect of reduction on word recognition, a linear mixed-effects model was run on the fixation proportions (time window 200–700 ms). Due to convergence issues, the random slopes were entered as uncorrelated random effects. The analysis revealed a main effect of Deletion ($B = -0.78$, $SE(B) = 0.16$, $t(68) = 7.62$, $p < .001$), with fewer fixations on the target if the initial segment was missing. But there was again neither an effect of Segment ($B = -0.03$, $SE(B) = 0.30$, $t(58) = -0.10$, $p > .2$) nor an interaction between the two factors ($B = -0.05$, $SE(B) = 0.40$, $t(20) = 0.135$, $p > .2$). Again these patterns were unchanged when spoken or written frequencies were taken into account.

Two additional models were fit to evaluate whether target prediction influenced the reduction costs, again using the 200–700 ms time window. The two analyses differed in how *prediction* was operationalized. One analysis used the mean fixation proportion before target onset *on a given trial* to predict the fixation proportion after target onset. The other analysis used the item predictability *aggregated over all trials*. This aggregated predictability was implemented as the item-specific random effect from the earlier analysis of the pre-target time window, that is, to what extent a given word was, on average, looked at more in the pre-target window. In both analyses, there was an effect of prediction on target fixation in the 200–700 ms time window (trial-specific prediction: $B = 0.14$, $SE(B) = 0.02$, $t = 6.95$, $p < .001$); aggregated predictability: $B = 0.76$, $SE(B) = 0.17$, $t = 4.48$, $p < .001$), but only aggregated predictability influenced the reduction costs (prediction × reduction interaction; trial-specific prediction: $B = 0.04$, $SE(B) = 0.04$, $t = 1.02$, $p > .2$; aggregated predictability: $B = -0.57$, $SE(B) = 0.21$, $t = 2.71$, $p < .01$). This indicates that deletion costs were smaller if the target word was more predictable, independent of the deleted segment.

Maltese /ʔ/ versus German /ʔ/

For the Maltese data, two items were removed from all analyses, because the majority of participants did not click on them, independent of deletion of the initial segment. One was *qafas*, Engl., *frame*, which is more often used in a metaphorical than concrete meaning, the second was *qarrej*, Engl., *reader*, which was often confused with its competitor *prezentatur*, Engl., *presenter*, *newsreader*. As Fig. 3 and Table 2 show, the Maltese performed less efficiently than the German participants as indicated by the longer response latencies. This is probably due to the fact that the Maltese participants were less experienced with this kind of experiment than the German participants. Note that our instruction did not specifically ask for either speed or accuracy, because we wanted to avoid any pressure typical for laboratory settings. However, our German participants were "used" to do experiments[7] and generalized the typical "respond as fast and accurate as possible" instruction to the current experiment.

The analysis of acceptance rate showed an effect of Deletion ($B = 0.79$, $SE(B) = 0.27$, $p < .01$), reflecting higher acceptance rates if the initial segment was present, an effect of Segment ($B = 1.9$, $SE(B) = 0.64$, $p < .01$), reflecting a higher level of acceptance rates by the German participants, but no interaction of these two factors ($B = -0.39$, $SE(B) = 0.56$, $p > .2$). This pattern did not change when written frequency was added to the model (note that no spoken frequency data are available for Maltese).

The absence of an interaction between Deletion and Segment may be surprising given that, in percentages, the acceptance rates drop much more in the Maltese data (by 7%) than in the German data (by 1.5%). Note, however, that a percentage scale should not be treated as an interval scale (Dixon, 2008; cf. Jaeger, 2008), since differences in raw percentages close to floor or ceiling are more meaningful than in the middle range. A 1.5% drop near 100% has to be considered as larger than a similar drop around 90% correct responses. We corrected for this using a logit analysis that takes this into account. If interpreted that way, the absence of an interaction in logistic space is justified. Note also, that many German participants accepted all items in some conditions, making the estimation of the mean quite difficult. Therefore the main focus of our argument is on the eye-tracking data, since eye-tracking data has been shown to be more sensitive than perceptual decisions with regard to finding deletion costs (Mitterer & Ernestus, 2006; Mitterer & McQueen, 2009).

For the latency of the correct responses, we find a marginal effect of Deletion ($B = -0.04$, $SE(B) = -0.02$, $t(45) = -1.83$, $p = .07$), reflecting slightly longer latencies if the initial segment is deleted, a clear effect of Segment ($B = -0.42$, $SE(B) = 0.08$, $t(96) = -5.22$, $p < .001$), reflecting faster responses in the German data set and no interaction ($B = -0.02$, $SE(B) = -0.04$, $t = -0.07$, $p > .2$). Again, this pattern did not change when written frequency was added to the model (note that no spoken frequency data are available for Maltese: $B_{Deletion} = -0.034$, $SE(B) = 0.019$, $t(35) = 1.83$, $p = .7$; $B_{Segment} = -0.42$, $SE(B) = 0.08$, $t(95) = -5.16$, $p < .01$; $B_{DeletionxSegment} = -0.029$, $SE(B) = 0.038$, $t(45) = -0.77$).

For the analysis of the eye-movement patterns, we again tested whether either the Maltese or the German targets were better predictable by looking at the time window from 600 ms before the target onset until target onset. There was no differential prediction of the target depending on Segment ($B = 0.07$, $SE(B) = 0.28$, $t(60) = .27$). To test the effect of reduction on word recognition, the time window from 200 to 700 ms after target onset was used, reflecting the initial lexical access (see Appendix B for an analysis of a later time window). This revealed a main effect of Deletion ($B = -0.80$, $SE(B) = 0.22$, $t(35) = 3.65$, $p < .001$), with fewer fixations on the target if the initial segment was deleted, a main effect of Segment ($B = -1.1$, $SE(B) = 0.34$, $t(74) = -3.12$, $p < .01$), with fewer looks to the target in the Maltese data set, but no interaction of the two factors ($B = 0.26$, $SE(B) = 0.31$, $t(37) = .81$). This patterns did not change when written frequency was added to the model ($B_{Deletion} = -0.81$, $SE(B) = 0.22$, $t(35) = 3.65$, $p < .001$; $B_{Segment} = -1.1$, $SE(B) = 0.34$, $t(74) = -3.10$, $p < .01$; $B_{DeletionxSegment} = 0.25$, $SE(B) = 0.32$, $t(68) = 0.79$). That is, deletion costs appeared equal for orthographically uncoded German and coded Maltese /ʔ/.

We again conducted additional analyses to investigate the influence of predictability on reduction costs, using trial-specific or aggregated measures of prediction. As in the comparison of German /h/ with German /ʔ/, both measures influence the amount of target fixation in the 200–700 ms time window for the German vs. Maltese comparison (trial specific prediction: $B = 0.18$, $SE(B) = 0.03$, $t = 5.524$, $p < .001$; aggregated predictability: $B = 0.78$, $SE(B) = 0.23$, $t = 3.47$, $p < .01$). In the current analyses, however, there was no significant reduction of deletion costs by predictability (trial specific prediction: $B = 0.02$, $SE(B) = 0.05$, $t = 0.42$, $p > .2$; aggregated predictability: $B = -0.41$, $SE(B) = 0.26$, $t = -1.56$, $p > .1$). To summarize, we find evidence of reduction costs in all three dependent measures, but no interaction with Segment. That is, deletion of German /ʔ/ is just as harmful for target recognition as deletion of Maltese /ʔ/.

Discussion

In a comparison of reduction costs for a word-initial segment that is coded in German orthography (i.e., /h/) vs. not (/ʔ/) we found evidence of reduction costs in all three dependent measures (acceptance rates, click latencies, and eye movements), but no interaction with Segment. That is, deletion of German /ʔ/ is just as harmful to spontaneous speech processing as deletion of German /h/. The same result was obtained in the comparison of German /ʔ/ with the Maltese /ʔ/, the latter of which is also orthographically coded. This has several theoretical consequences.

First, it indicates that ascribing phonemic status in German to /h/ but not to German /ʔ/ may be questionable. Both segments behave similarly in terms of their distribution and realization, and the current data indicate

[7] It is important to note here that the Psychology Department at the University of Malta does not perform laboratory based research, which means that only very few participants had previously participated in similar studies.

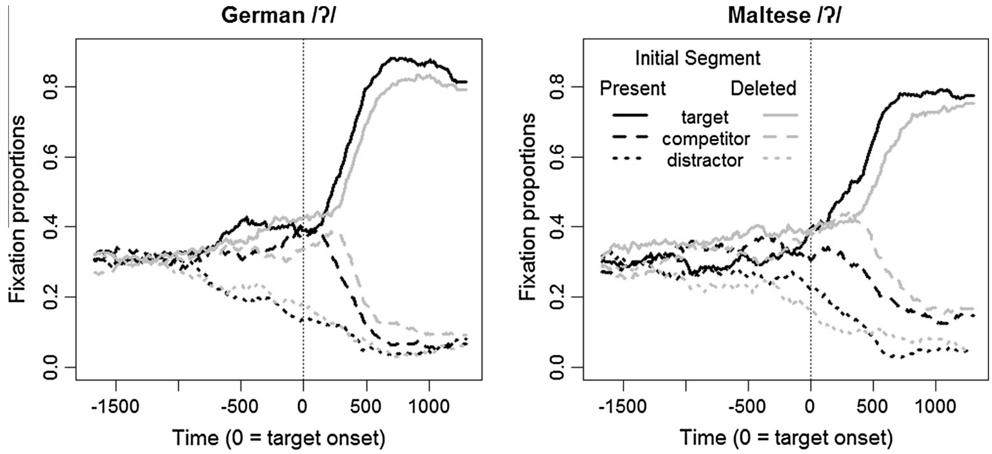

Fig. 3. Fixation proportion to pictures with German glottal stop-initial (left panel) and Maltese glottal stop-initial (right panel) targets, depending on whether the initial segment of the target is produced or not.

that deletion of these segments also leads to similar effects. This shows that there is little empirical basis to make a categorical distinction between a "phoneme /h/" and a "boundary marker /ʔ/". This is reinforced by the fact that there is no obvious difference in the consequence of deletion of German /ʔ/ and Maltese /ʔ/, the latter of which is undeniably a phoneme. This resonates with recent trends in linguistics to view phonemic status as gradient (Hualde, 2004; Scobbie & Stuart-Smith, 2008) and to allow marginal phonemes instead of making categorical distinctions between "phonemes" and non-contrastive variation.

The second theoretical consequence concerns our main question. The results indicate that the orthographic coding of German /h/ and Maltese /ʔ/ does not lead to additional processing costs for its reduction, as would be expected if an inconsistency between sound and spelling would negatively influence spoken-word recognition. However, it may be argued that, even though deletion rates of /h/ and /ʔ/ are similar, the perceptual consequences of their absence may differ. Our data hence provide a counterpoint to theoretical thinking that arose out of the paper by Morais et al. (1979). As reviewed above, they had shown that awareness of phonemes does only arise with reading instruction. Since then, it has become an undercurrent in cognitive science that leaning to read makes us perceive speech more segmentally. However, the case of German glottal stop shows that learning to read does not make us perceive speech in terms of segments. If so, German speakers should become aware of the glottal stop that occurs frequently in German, but our intuition and experience is that German speakers without phonetic training are completely unaware of this segment of German.[8]

This observation about meta-linguistic knowledge about glottal stop indicates that there might be a huge divide in how orthography influences speech processing in a normal conversation versus a more explicit setting, with orthography having little effect on normal speech processing. However, this is based on informal observations regarding the knowledge of German readers on the sound structure of German. Therefore, Experiment 2 tried to show that this translates into experimental effects once we use formal speech and an explicit task. Note that this also deals with another possible objection to our preliminary conclusions. In Experiment 1, we failed to find an effect of orthography, which basically is a null-effect. We used a power calculator for experiments with crossed random effects (Westfall, Kenny, & Judd, 2014) and found that for our counterbalanced design (i.e., each participant heard a given target only as either reduced or unreduced) with 22 participants and 71 items for the within-German comparison, we have a power of 0.86 to detect a medium-sized effect. Assuming a medium-sized effect is justified based on earlier evidence: Ziegler et al. (2004) report a 62 ms effect of orthographic consistency in auditory lexical decision. With a standard error of the mean of 11 ms for 22 participants, this leads to a standard deviation of 52 ms, which gives rise to an effect size of $d = 1.2$ (62 ms/52 ms). Our design would have a power of 1 to reveal such an effect (i.e., according to the power calculator provided by Westfall et al.). While this indicates that our study was not underpowered, the best way to counter this argument is by showing that an effect of orthography arises with similar items (see below for details) and a similar number of participants.

Experiment 2

For this experiment, we re-recorded the same items used in Experiment 1 but with two changes. First of all, the items were embedded in a minimal context rather than in a full sentence. Secondly, speakers were asked to speak clearly rather than casually. While it would be ideal to record these words as one-word utterances as in most

[8] Part of this work was presented at AMLaP 2014 in Edinburgh, where a reviewer of the conference abstract (bravely) admitted to being a native German speaker (and a psycholinguist) without any explicit knowledge about the role of glottal stop in German.

other experiments on orthographic effects in speech perception, a minimal context was necessary to make the glottal stop audible. Starting from the sentences used in Experiment 1, the target words were therefore embedded in the smallest phrase from that sentence that could conceivably be an (elliptic) answer to a question. For instance, for the German target *Eimer* (Engl., *bucket*) the chosen phrase was *in dem Eimer* (Engl., *in the bucket*) which is a possible answer to *Wo ist ein Loch drin?* (Engl., *Where do we have a hole*?). To give the participants an explicit task, they were asked to judge how well the word was pronounced and, as in Experiment 1, the target words were presented with and without the initial segment.

Method

Participants

22 German native speakers and 21 Maltese native speakers from the same populations as in Experiment 1 (students at the Universities of Munich and Malta) participated in the experiment. They were paid for their participation.

Materials and procedure

The same items as in Experiment 1 were recorded with the same speakers, but now with the instruction to speak clearly. Additionally, the minimal phrases were formal in terms of style in that no contractions were applied (hence the German item *im Eimer* from Experiment 1, was produced as *in dem Eimer* in Experiment 2; analogue to the English versions of you're vs. you are). The items were again recorded with and without the initial segment and crossed-spliced as in Experiment 1. That is, again only the first syllable of the target word differed between segment-present and segment-absent versions and both segment-present and segment-absent versions were spliced.

These stimuli were then presented in an explicit judgment task. Participants first saw a written version of the target word on the screen, and then heard the short phrase containing the target word. They were instructed to rate how well the target word was pronounced in this short phrase by pressing a number key from 1 to 7. On the scale, "7" was labelled as "very good" (*sehr gut* and *tajba ħafna*), 5 was labelled as "good" (*gut* and *tajba*), 3 was labelled as "bad" (*schlecht* and *ħazina*), and 1 as "very bad" (*sehr schlecht* and *ħazina ħafna*).

In the German version of this experiment, each participant heard the 34 /h/-initial and the 37 /ʔ/-initial items, half of which were presented with and half without the initial segment, again counterbalanced across participants. For the Maltese version, we presented the 36 /ʔ/-initial items plus 36 fillers starting with other consonants (these recordings were not manipulated). Fillers were necessary for Maltese since otherwise all target words would have started with the grapheme 'q' which could have given away the critical manipulation. In this way, in both versions, there was about one quarter of trials in which a scripted segment was missing. The experiment lasted about 6–10 min (participants were not instructed to react as fast as possible, leading to variation in how fast they made their decision).

Results and discussion

Fig. 4 presents the mean goodness ratings. The error bars indicate the standard error of the mean based on a by-participant estimate. The figure shows that leaving out the first segments lead to a much worse pronunciation rating compared to the items with all segments present. Importantly, this effect was modulated by the orthographic status of the missing segments. While the ratings for Maltese /ʔ/ (written as 'q') and German /h/ (written as 'h') went from 6 (between *very good* and *good*) to 2 (between *very bad* and *bad*), deletion of the unscripted German /ʔ/ only lead to a change from 6 to 4.

As in Experiment 1, we carried-out a within-participant analysis for the German data and a between-participant analysis for the data regarding the glottal stop in German vs. Maltese. In both analyses a linear mixed-effects model with the fixed effects Segment, Deletion (contrast coded to -0.5 and 0.5) and their interaction was run (with the package *lmerTest* v2.0.25). A maximal random-effect structure was included.

In the comparison of German /h/ and /ʔ/, we found a main effect of Deletion ($B = 3.11$, $SE(B) = 0.15$, $t(42) = 21.04$, $p < .001$), an effect of Segment ($B = 1.02$, $SE(B) = 0.14$, $t(60) = 7.12$, $p < .001$), and, critically, an interaction of the two factors ($B = -1.89$, $SE(B) = 0.26$, $t(51) = -7.16$, $p < .001$). That is, deletion lowers the pronunciation ratings but more so for the orthographically coded segment /h/ than uncoded /ʔ/.

The same pattern emerges for the comparison of German versus Maltese /ʔ/, with an effect of Deletion ($B = 3.10$, $SE(B) = 0.18$, $t(60) = 17.95$, $p < .001$), an effect of Segment ($B = 0.77$, $SE(B) = 0.18$, $t(60) = 4.17$, $p < .001$), and an interaction of the two factors ($B = -1.86$, $SE(B) = 0.35$, $t(60) = -5.38$, $p < .001$).

The present results show that in an explicit pronunciation judgment task with careful speech listeners are influenced by the orthographic coding of the target words. In conjunction with Experiment 1, they thus demonstrate a dissociation between explicit processing with careful

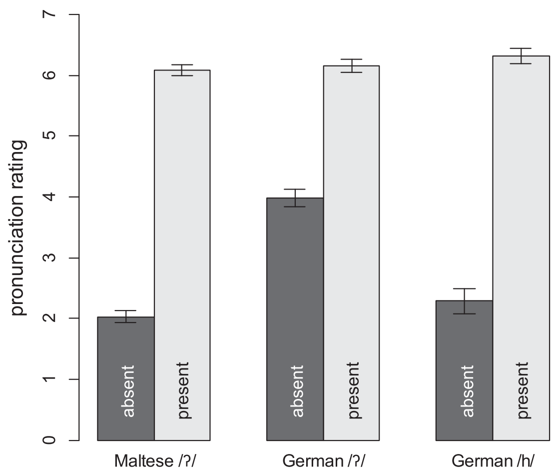

Fig. 4. Goodness ratings on the pronunciation of the target words depending on whether the first segment is present or absent and the nature of the segment.

speech and implicit processing when confronted with normal, casual speech. This is unlikely to be due to a power problem, since the number of participants and items were similar in both experiments.

General discussion

Our results demonstrate that segmental deletions in casual speech make word recognition harder (for a review, see Ernestus, 2014). As such they replicate the finding that reduced forms are recognized less efficiently than full forms (Ernestus, 2014). However, these deletion costs are not moderated by the orthographic coding of the deleted segment. Based on the assumption that orthography influences speech perception online, we had predicted that the deletion costs would be larger for the orthographically coded German /h/ and Maltese /ʔ/ as compared to the uncoded German /ʔ/. This prediction was not borne out, instead, the deletion costs were comparable. This suggests that, in perceiving conversational speech, orthography has little role to play. This would seem functional given the massive difference between rather variable phonological forms in conversational speech and invariant orthographic forms: While "jeshay" may be an acceptable phonological form of *yesterday*, it is certainly not an acceptable orthographic form. In conversational speech, inconsistencies between spelling and pronunciation are the norm.

There is another difference between spoken and visual-word recognition that may make a one-to-one mapping problematic. While we argued that both make use of pre-lexical abstract units, the form of these units may be quite different. In spoken-word recognition these units may not be phoneme sized and hence difficult to link to graphemes. The paradigm that has been crucial for showing the reality of pre-lexical units in spoken-word recognition is the "perceptual-recalibration paradigm", in which participants recalibrate their pre-lexical categories to fit the properties of an usual speaker, using either lexical or visual information (Bertelson, Vroomen, & de Gelder, 2003; Norris, McQueen, & Cutler, 2003). Results on generalization of these recalibrated units suggest that Dutch listeners may have at least three different units for /r/ (Mitterer, Scharenborg, & McQueen, 2013) and American English listeners seem to use different /b/s for different vowel contexts (Reinisch, Wozny, Mitterer, & Holt, 2014).[9] As such, it would not be as straightforward to achieve a link between orthographic and phonological pre-lexical representations, because the latter may be much more specific than orthographic units in visual-word recognition.

We have argued in the introduction that a penalty for S2L inconsistent words in perceiving casual speech would be a bad design feature. While the literature usually conceptualizes the orthographic effect in speech perception as a burden for inconsistent words, it is possible to re-conceptualize this as a *benefit* for consistent words, since there is no neutral control condition. This would mean that the activation of orthography would not have detrimental effects on spoken-word recognition. However, there would also not be a huge benefit either, because this benefit only arises for words that are produced carefully and canonically. In this case spoken-word recognition is easy anyway, so that this benefit would not lead to a huge improvement in the efficiency of spoken-word recognition.

Another question that follows from our findings is why do earlier studies and our Experiment 2 find evidence that orthography influences speech perception (Ziegler & Ferrand, 1998)? Such studies usually rely on single-word presentations of clear utterances. Such school-like stimuli may indeed put participants into a "school-like" task setting, in which orthography would have a role to play. The stimuli in these experiments typically have a syllable rate of about 2 Hz (Taft et al., 2008), which is twice as slow as normal speech, which tends to have a rate of 4–5 Hz. A speech rate of 2 Hz is hence reminiscent of slow, early reading, which might facilitate the activation of orthography in response to such stimuli.

Our data alone, however, allow an alternative interpretation in terms of explicitness of the task, since the effect of orthography was observed in an explicit task with careful speech but not in an implicit task with casual speech. That is, the contrast between the current two experiments confounds speech type with task. However, our central theoretical proposition is that, in everyday speech perception, listeners do not automatically activate orthographic representations. Everyday speech perception, too, tends to confound casual speech and a natural listening situation. The Visual-World eye-tracking task is relatively natural in asking participants to find a referent to what they hear in the visual environment and hence models this natural listening situation. Although we cannot rule out that, by imposing an unnatural task, orthographic effects may be observed even when listening to conversational speech, this is would be of little theoretical consequence. The current data would then still suggest that such orthographic effects are, with conversational speech, restricted to this particular task setting. In fact, it is not unlikely that both the use of careful lab speech and explicit tasks may foster the use of orthographic representations in speech perception. While further research is necessary to delineate how effective both the task variable and the type of speech are in activating orthographic representations, the current data strongly suggest that, in normal conversations, interlocutors do not make use of orthographic representations for speech perception.

Another potential factor that may mediate effects of orthography on spoken-word recognition may be how given words are used. It has often been noted that effects of orthography tend to be rather variable over items, and effects are often only significant by participants but not by items (Rastle et al., 2011). One factor may be that word choices differ substantially in written and spoken language (Hayes, 1988). As a consequence, there are words that are "spoken-dominant" and "written-dominant", that is, more frequent in one of the modalities. With the visual-world paradigm, we are limited to use pictureable nouns, which

[9] The qualifications for a given language are supplied, because we deem it likely that these findings are highly dependent on the exact articulatory settings for different segments in the given languages.

are more likely to be "spoken-dominant".[10] It may very well be the case that orthography is automatically activated for words that are "written-dominant", that is, words that are often found in texts but not used often in spoken language.

As already alluded to in the discussion of Experiment 1, the current results indicate that reading does not make us perceive speech in a more segmental manner. Instead, it becomes more and more apparent that learning to read only *leads us to believe* that we know something about speech. For instance, as Ernestus (2013) remarked, it is rather curious that we say words like *prowly* all the time, we hear them all the time, yet we are completely unaware that this is a frequent phonological form of *probably*. If learning to read makes us perceive speech in terms of segments, we should realize that such forms as *prowly* exist, because we should hear them as /p/+/r/+/oʷ/+/l/+/ɪ/. However, we clearly do not. In fact, listeners will even say that a form such as *prowly* contains a /b/ (Kemps, Ernestus, Schreuder, & Baayen, 2004), probably due to the fact that knowing how the word *probably* is written influences our intuitions about which segments are present in the speech stream. Reading may hence rather make us deaf for the real properties of normal, conversational speech. In a way, the view of speech that we acquire with learning to read is as accurate as how Plato thought of sensory evidence in his "allegory of the cave". Letters represent the phonetic reality just as badly as shadows of figures on a cave wall represent reality.

Another aspect of our data also indicates that learning to read does not make us listen to speech in terms of segments (as assumed by Dehaene et al., 2010; Pattamadilok et al., 2009, see the relevant quotations in the introduction). If it would, German native speakers should have discovered that the segment glottal stop occurs in German. This can be deduced from the current data as follows. Experiment 1 shows that, in perception, German glottal stop is as important a segment for word recognition as /h/. If learning to read makes us discover that speech consists of such segments, German listeners should not only become consciously aware of /h/ but also of the glottal stop. However, German speakers usually have no idea that glottal stop exists in German, unless explicitly taught. Consider Fig. 1, with the phrase [bɛnʔɑm] in Maltese and German. Despite being phonetically very similar, German speakers will typically say that it contains 5 segments (/b/, /ɛ/, /n/, /ɑ/, /m/), while Maltese speakers will typically say that it contains 6 segments (/b/, /ɛ/, /n/, /ʔ/, /ɑ/, /m/). The example of the German glottal stop then indicates that learning to read may cause phonetic deafness to anything that is not coded in the orthography. An example of that is given (unwillingly) by Ohk (2006) in a Maltese–German dictionary. In an overview of how the Maltese graphemes are to be pronounced, she writes about *q* (the grapheme used in Maltese for the glottal stop) that the corresponding sound does not exist in German (p. xviii: "Dieser Laut [glottal stop] existiert im Deutschen nicht"). This suggests that learning to read influences *thinking about* speech, but not the *processing of* speech. If learning to read would make us perceive speech in terms of segments, German speakers (including Ohk, 2006) should know of /ʔ/. Phoneticians (with a non-German L1) have often remarked to us how widespread and clear the use of glottal stop is in German. Yet this does not become evident to German native speakers by learning to read. Instead, learning to read may in fact cloud our view of speech, so that we are unaware of reductions and segments that are not coded in the orthography.

Finally, the current data also indicate that, in spoken-word recognition, German /ʔ/ is as important as German /h/ and Maltese /ʔ/. This provides evidence that, in Phonology, German /ʔ/ should not be treated categorically different from German /h/ and Maltese/ʔ/. While it is obvious that there are differences in phonological status between, for example, the German alveolar stop /t/, which can occur in onset and coda position and in various clusters, and the phonotactically more restricted glottal stop in German, such differences are probably not best described by making a categorical difference between the two. Such a categorical difference is even less called for between German /h/ and German glottal stop, which both are quite restricted phonotactically. As others have remarked, phonemic status is a concept for which grey areas need to be recognized, so that we should consider marginal phonemes as a normal, rather than exceptional (Hualde, 2004; Scobbie & Stuart-Smith, 2008). German /h/ and /ʔ/ are probably both best conceptualized as weak or marginal phonemes.

In summary, our data indicate that in conversational speech, orthography has little role to play. This is probably not too surprising, because the computational problem during auditory and visual word recognition is massively different—unless unusual auditory material with slow speech and all segments intact is used. Input variability hardly plays any role in visual word recognition (for instance, it seems that choices of font in studies of visual-word recognition are quite arbitrary and of little consequence). Spoken-word recognition, however, is all about variability. While perception of allophones and "auditory pre-processing" (or the perception of speech in terms of gestures) may reduce the variability in the speech signal to a considerable degree, research on spontaneous speech corpora indicates that variation is often so strong that the phonemic transcription changes (Keating, 1997). The visual analogue to this is an orthographic error, but in speech such variation is normal and not treated as an error. Indeed, the very idea that an inconsistency between pronunciation and orthography should slow down spoken-word recognition would mean a huge burden for functioning in a dialogue. Given these considerations, it is unlikely that, in online processing during normal conversation, the two should interact closely.

Appendix A

This appendix provides the experimental items for the three conditions (Maltese /ʔ/, German /ʔ/, and German /h/). The English translations were done so that the word

[10] We compared the frequency of our words in the spoken and written modality using the SUBTLEX-DE corpus (Brysbaert et al., 2011) comparing the occurrence of our words in subtitles (as an estimate for spoken frequency) and the Google Books 2000–2009 corpus. The results showed that our target words were on average roughly 3 *times* more frequent in spoken than in written communication.

order of the translation was as similar as possible to the word order of the original sentence. The critical target words are in *italics* in both the original sentences and the translations. Brackets are used to indicate reductions and their canonical forms. Also note that some of the concepts are difficult to capture in translation (e.g., the Maltese *qassata*, which refers to a unique shape of pastry).

Maltese /ʔ/ items

Dam sa ma dalam quddiem *il-qabar* t' ommu u ta'missieru
He stayed till dusk in front of *the grave* of his mum and dad

Qas emmintu meta qalli li weġġa saqajh meta rifes fuq *l-qabru*
Can you believe it, he hurt his foot when he stepped on a *crab*

Kien pjuttost magħruf ħafna għad-devozzjoni kbira tiegħu *Il-qaddis* tar-raħal
Well known for his great devotion was the *saint* of the village

Rega' poġġa l-*qafas* li kien xtara fuq il-mejda
They again put the *frame* that they just bought on the table

L-istudenti tal-mediċina kienu ghal darb'oħra qed jistudjaw fuq il-*qalb*
The medical student were again studying the *heart*

Minħabba ir-rih qawwi, bla dubju ta' xejn il-*qala*' beda jixxejjer
Because of the strong wind without a doubt, the *sail* started to flap

Ghal dar'ohra regghet ħaslet il-*qalziet* u naxritu fuq il-bejt
For another time, she again washed the *trousers* and hung them up on the roof

Fil-klassi ta' l-astronomija osservajt l-*qamar* għall-ewwel darba f'ħajti
In the astronomy class, I observed the *moon* for the first time in my life

F'lulju, bhas-soltu il-bdiewa qatgħu il-*qamħ* kollu li kien hemm fl-għalqa
In July as usual, the farmers cut all the *wheat* that was in the fields

Oħti tgħidx kemm xtaqet iżżomm il-*qanfud* li rajna meta morna nimxu fil-kampanja
My sister said how much she wanted to keep the *hedgehog* that she saw this morning when we went for a walk in the countryside

Bħala souvenir tal-vaganza tagħhom, għal sena'oħra il-ġenituri taw Il-*qanpiena* ta' San Pietru
As souvenir from their holidays, for another year the parents gave them (a replica of) the *bells* of Saint Peter

Ommi tant għogobha il-*qaqoċċ* li reggħet sajritu għall-ikel
My mother really liked the *artichoke* that she used for dinner

Għat-tieni darba dan ix-xhar il-kaċċaturi reggħu sabu l-*qarn* ta' annimali selvaġġi fil-foresta
For the second time this month, the hunters again found *antlers* from one of the wild animals in the forest

Is-sajjieda li kienu għajjiena mejta poġġew *il-qarnit* li qabdu għal frisk
The cooks who were really tired served the *octoptus* that they had caught freshly

F'ħin minnhom waqt il-laqa' il-*qarrej* xeraq u kellhu bżonn jixrob
At some time during the meeting, the *reader* choked and had the need to drink something

Fid-dokumentarju li rajt il-bieraħ qal li f'Kuba hemm ħafna sigar tal-*qasab*
In the documentary that I watched yesterday, it was said that in Cuba there are many trees of (sugar) *canes*

Miskin hija ma jistax jgħum, minħabba il-*qasma* li għandu f'saqajħ
This poor guy cannot go swimming because of the *cut* he has in his foot

Ommi tgħidx kemm għoġbitha l-*qasrija* li xtrajtilha għal 'Jum l-Omm'
My mum said how much she liked the *flowerpot* that I bought her for mother's day

Fl-ahhar poġġejt naqra bilqiegħda u kilt il-*qassata* bil-kwiet
At last, I sat down and ate the *pastry* quietly

Għal program ta din il-gimgħa, il-presentatur stieden lil *qassis*
For the program this week, the presenter had invited a *priest*

Missieri rega' kiel il-*qastan* kollu
My father again ate the *chestnut* in its entirety

Meta sema' il-ħsejjes, il-*qattus* tgerrex u telgħa b'ġirja waħda għal fuq is-siġra
When it heard the noises, the *cat* was scared away and with one swift move climbed the tree

Kelli nħares barra mit-tieqa u rajt il-*qawsalla*
I had to look outside the window and saw a *rainbow*

Fl-istorja tat-tfal il-*qawwies* kellu bil-fors isalva lill-prinċipessa
In the kids' story, the *archer* obviously had to save the princes

Il-bidwi kellu jerga' jieħu l-*qażquż* il-marid għand il-veterinarju
The farmer again had to take the sick *piglet* to the veterinary

(*continued on next page*)

Il-mara kienet pronta u poġġiet il-*qoffa* vojta ma' l-art
The women was ready and put the empty *basket* on the floor

Sfortunatament il-bennej mar l-isptar għax weġġa l-*qorriegħa*
Unfortunately, the builder went to hospital because he had hurt his *forehead*

Għar-raba gurnata nfila rega' kien hemm folla kbira tan-nies quddiem il-*qorti*
For the fourth day, there was again a large crowd of people in front of the *court*

It-tfal imqarqca reggħu ħallew il-*qoxra* tal-larinġa fuq il-mejda ta' l-ikel
The kids again left the *peel* of the orange of the dinner table

F'dan l-istagun il-bdiewa qatgħu *l-qoton* li kien hemm fl-għalqa
In this season, the farmers cut the *cotton* that was in the fields

Bhas-soltu il-hawker mesaħ il-*qorq* meta wasal lura d-dar
Normally, the hawker scrapped his *sandals* clean when he arrived back at home

It-tfal m'għobdewx minn ommom u xtraw il-*qubbajt* qabel ma telqu mill-festa
The kids did not please their mother and bought some *nougat* before they left the feast

Fir-rumanz li qrajt l-ahhar il-*qaddej* tas-sinjuri kien ragel xwejjah
In the novels I read lately, the *servant* of the rich men was the culprit

Bil-guħ li kellna kilnih kollu il-*qarabagħli* li xtrat ommi
With the hunger that we had, we ate all of the *courgette* that my mother had bought

Missieri prova kemm segħha biex ikabbar il *qargħa ħamra*
My father tried how large he can grow the *pumpkin*

Hu rega kellu jbiddel il-qmis minħabba il-*qatra* tal-inbid aħmar
He again had to change his shirt because of a *drop* of red wine

---

German /ʔ/

Der Typ repariert heute endlich den *Abfluss* im Bad
This guy tonight finally repairs the *drain* in the bathroom

Da hinter'm (=hinter dem) Hügel konnte man am Himmel einen Adler sehen
There behind that hill you can see an *eagle* in the sky

Um sechs zündete sie wie immer die Kerzen auf'm *Adventskranz* an
At six, she normally lit he candles on the *Advent wreath*

Ganz gebannt schaut das Kind im Zoo dem *Affen* zu
All in awe the kid observed the *apes* in the zoo

Am liebsten spielte er eigentlich auf'm *Akkordeon*
In the end, he preferred to play the *accordion*

Jochen hat mal wieder vergessen, die Dokumente in'en [=in den] *Aktenschrank* zu legen
Jochen had again forgotten to put the documents into the *filing cabinet*

Im dem Kanal lag echt ein *Alligator*
In this channel, there really was an *alligator*

Stell Dir vor, Petra wurde gestern von'ner giftigen *Ameise* gebissen
Believe it or not, yesterday Petra was bitten by a poisonous *ant*

Der ist echt trotz der roten *Ampel* über die Kreuzung gefahren
This guy really drove across the junctin despite the red *traffic light*

Für'n Cocktail brauchte er neben *Ananas* auch noch Kokos
For the cocktail, he needed not only *pineapple* but also coconut

Er hat sich dann doch fürn blau(e)n *Anzug* entschieden
He finally decided to go with the blue *suit*

Da sind total viele Vitamine in dem *Apfel*
There are lots of vitamins in this *apple*

Er brauchte noch einen Reiniger für sein *Aquarium*
He still needed some detergent for his *aquarium*

Sie probierte den Schmutz aus ihrem *Auge* zu kriegen
She tried to get the dirt out of her *eye*

Ausgerechnet heute hatte er seinen *Ausweis* verloren
Today of all days he had lost his *ID card*

Er dreht bei dem schö(ne)m Wetter eine Runde mit dem *Auto*
With this good weather he takes the *car* out for a spin

Er machte heute den gesunden, aber nicht so leckeren Salat mit den *Avokados*
Today he makes the healthy, but not really tasty, salad with *avocados*

Auf'm Baum hatte sich ein *Eichhörnchen* eingenistet
On this tree there lived a *squirrel*

Da war ein Loch im *Eimer*
There is a whole in the *bucket*

In dem neu(e)n Fantasy-Film kam en (=ein) *Einhorn* vor
In this new fantasy movie there also is a *unicorn*

Nebenbei legte er die Nudeln in den *Einkaufswagen*

Casually he put the noodles into the *shopping trolley*

Im kalten Wasser war tatsächlich ein *Eisbär* zu sehen
In this cold water there really was an *ice bear* to be seen

Durch d'n Sturm gestern nacht fiel auch en (=ein) Eiszapfen auf den Boden
Because of the storm last night, an *icicle* fell on the floor

Mitten im Wald steht ein *Elch*
In the middle of the forest there is a *moose*

An der Wasserstelle war nirgends ein *Elefant* zu sehen
At the watering hole, there was no *elephant* to be seen

Das Bild im Wohnzimmer zeigte unter anderem einen *Engel*
The painting in the living room displayed amongst other things an *angel*

Für sein neues Rezept hatte er nebm (=neben) Ente auch Spargel benutzt
For his new recipe, he used not only *duck* but also asperges.

Morgens früh fütterte er zuerst immer den Esel
In the early morning, he first fed the *donkey*

Im Garten hatte er schon lange keinen *Igel* mehr gesehen
In the garden, he had not seen an *hedgehog* for a long time

An den wenigen schönen Wintertagen bauten sie 'n Iglu
On the few beautiful days of winter, they built an *igloo*

Sie liebte den Film mit dem *Indianer*[a]
She loved the movie with the *native Americans*

Dann holt er die Torte aus'm *Ofen*
Then he got the cake out of the *oven*

Sie suchte ewig nach dem zweiten Ohrring
She took forever looking for the second earring

Er fügt noch die Schale 'ner unreifen *Orange* hinzu
He added the peel of an unripe *Orange*

Paula suchte in dem *Ordner* vergeblich nach der Rechnung
Paula searched unsuccessfully in that *folder* for the bill

Er denkt echt er könnte sie mit 'ner teuren *Uhr* beeindrucken
He really thinks he would be able to impress her with an expensive *watch*

Die zweite Bombe war auf'm *Unterseeboot* versteckt
The second bomb was hidden on the *submarine*

[a] In the German cultural context, the term *Indianer* is not a loaded term and still the most common reference term for *Native Americans*. At the time of this writing, there is also an entry in the German Wikipedia on "Indianer" that describes the history of the indigenous people of the Americas.

German /h/-initial words

Er findet das Sprichwort mit dem *Hahn* total doof
He thinks the saying with the *rooster* is totally stupid

Er wollte unbedingt den Film mit dem *Hai* sehen
He definitely wanted to see the movie with the *shark*

Dann hatte sie die Marke ganz vorsichtig am *Halsband* befestigt
Then she had fastened the badge really carefully on the *collar*

Er fluchte laut als ihm ein *Hammer* auf die Füsse fiel
He cursed loudly when the *hammer* fell on his feet

Als Haustier wollten sie unbedingt einen *Hamster*
As a pet, they definitely wanted a *hamster*

Er hatte gestern früh im Bus 'nen *Handschuh* gefunden
Yesterday morning on the bus, he had found a *glove*

Fürs Sportstudio brauchte er 'n *Handtuch* zum trainieren
For the gym, he needed a *towel* in order to train

Er liest noch eben schnell seine Mails auf'm *Handy*
He quickly checks his emails on his *phone*

Heute trainierte er zur Abwechslung mal mit der kleinen *Hantel*
Today for a change, he trained with the small *weight*

Einlich (=eigentlich) spielte sie am liebsten auf der alten *Harfe*
Actually, she preferred to play on the old *harp*

Er hatte noch nie gesehen wie schnell ein *Hase* laufen kann
He never had seen before how fast a *hare* could run

Rechts neben dem *Haus* hatten sie eine Reihe Bäume gepflanzt
On the right side of the *house*, they had planted a row of trees

Um die Beleuchtung einzuschalten, musste er den *Hebel* da umlegen
To turn on the lighting, he had to throw that *lever* there

Das Schloss war von ner hohn *Hecke* umgeben
The castle was surrounded by one of these high *hedges*

Der Schüler kritzelte die ganze Zeit in seim *Heft* rum
The pupil was constantly doodling in his *notebook*

Er musste den Handwerker schon wieder wegen der kaputten *Heizung* anrufen
He had to call the craftsman yet again because of the faulty *heating*

Beim Skifahren hatte er immer einen *Helm* auf
While skiing, he always was wearing a *helmet*

Jetzt hatte er schon wieder einen Fleck auf seinem *Hemd*
Now he had yet again a stain on his *shirt*

(*continued on next page*)

In den Zeitungscomics mit dem Wikinger kommt manchmal so'n *Henker* vor
In these Sunday funnies with the wiking, there sometimes is this *executioner*

Der Koch fand nicht wirklich Gefallen an seinem neuen *Herd*
The chef did not really get to like his new *stove*.

In Holland wollte er unbedingt mal geräucherten *Hering* essen
When in Holland, he absolutely wanted to eat one of this smoked *herrings*

Zur Sicherheit wollte er auch sein *Herz* checken lassen
Just to be sure, he also wanted to check his *heart*

Das nervige Kinderlied handelt von so'ner kleinen *Hexe*
The annoying children's song was about some small *witch*

Opa mag am liebsten den Kuchen mit den *Himbeeren*
Grandpa definitely prefers the cake with the *raspberries*

Im ganzem Wald konnte man den *Hirsch* beim Balzen zuhören
In all of the forest, one could hear the *stag* doing his mating display

Er hatte im Supermarkt echt ewig nach dem *Honig* gesucht
He took like forever searching for the *honey* in the supermarket

Auf'm Bild war er mit seiner einzigen blaunen *Hose* zu sehen
On this photo, he could be seen wearing his only pair of blue *pants*

Das Unfallopfer wurde mit 'nem *Hubschrauber* ins Krankenhaus gebraucht
The victim of the accident was transported by *helicopter* to the hospital

Als Glücksbringer hatten sie schon mal ein *Hufeisen* an die Wand gehängt
As a lucky charm, they had already hung a *horseshoe* on the wall

Er hatte noch nie richtig frischen *Hummer* gegessen
He never had eaten reallly fresh *lobster* before

Das Kind fürchtete sich ein bisschen vor dem grossen *Hund*
The child was a little bit afraid of the large *dog*

Er hat sein Fahrrad mit so'ner lauten *Hupe* ausgestattet
He had equipped his bicycle with one of these loud *horns*

Der alte Mann echt ging nie ohne seinen *Hut* zur Kirche
The old man really never made his way to church without his *hat*

Appendix B

Analyses of late time window from Experiment 1 (700–1200 ms)

German /h/ versus German /ʔ/

Term	Estimate	Standard error	*df*	*t*
(Intercept)	2.77	0.16	32.1	16.93***
deletion	0.55	0.15	135.3	3.65***
phoneme	0.53	0.22	44.66	2.41*
deletion: phoneme	−0.13	0.37	23.8	00.72

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Adding spoken frequency

Term	Estimate	Standard error	*df*	*t*
(Intercept)	2.77	0.16	32.44	16.82***
deletion	0.54	0.15	132.24	3.64***
phoneme	0.52	0.23	57.46	2.31*
FreqSpoken	−0.02	0.08	69.54	−0.23
deletion: phoneme	−0.10	0.38	24.99	−0.27
deletion: FreqSpoken	0.08	0.11	132.96	0.7

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Adding written frequency

Term	Estimate	St error	*df*	*t*
(Intercept)	2.77	0.16	32.31	16.86***
deletion	0.55	0.15	137.15	3.66***
phoneme	0.50	0.23	57.52	2.22*
FreqWritten	−0.07	0.09	68.59	−0.74
deletion:phoneme	−0.07	0.37	25.02	−0.18
deletion: FreqWritten	0.17	0.13	138.54	1.25

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Maltese /ʔ/ versus German /ʔ/

Term	Estimate	Standard error	*df*	*t*
(Intercept)	3.04	0.21	75.07	14.29***
deletionC	0.49	0.21	82.44	2.29*
languageMaltese	−0.83	0.31	76.17	−2.74**
deletionC: languageMaltese	0.11	0.31	86.63	0.35

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

Adding written frequency

Term	Estimate	St error	*df*	*t*
(Intercept)	3.06	0.21	74.99	14.26***
deletionC	0.49	0.22	83.12	2.27*
languageMaltese	−0.85	0.31	75.88	−2.77**
FreqWritten	0.07	0.10	68.55	0.70
deletionC: languageMaltese	0.11	0.31	86.69	0.35
deletionC: FreqWritten	0.00	0.12	204.59	0.01

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439. http://dx.doi.org/10.1006/jmla.1997.2558.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*, 621–635.

Baayen, H. R. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language, 73*, 116–130. http://dx.doi.org/10.1016/j.jml.2014.03.002.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001.

Bertelson, P., Vroomen, J., & de Gelder, B. (2003). Visual recalibration of auditory speech identification: A McGurk aftereffect. *Psychological Science, 14*, 592–597.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*, 419–421.

Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics, 34*, 519–539. http://dx.doi.org/10.1017/s0142716411000853.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*, 412.

Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language, 62*, 421–437. http://dx.doi.org/10.1016/j.jml.2010.01.002.

Bürki, A., & Gaskell, M. G. (2012). Lexical representation of schwa words: Two mackerels, but only one salami. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 617–631. http://dx.doi.org/10.1037/a0026167.

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics, 70*, 403–411. http://dx.doi.org/10.3758/PP.70.3.403.

Cutler, A., & Davis, C. (2012). An orthographic effect in phoneme processing, and its limitations. *Language Sciences, 3*, 18. http://dx.doi.org/10.3389/fpsyg.2012.00018.

Cutler, A., Eisner, F., McQueen, J. M., & Norris, D. (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology 10* (pp. 91–111). Berlin: de Gruyter.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language, 82*, 515–535. http://dx.doi.org/10.1353/lan.2006.0130.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ... Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*, 1359–1364.

Dixon, P. (2008). Models of accuracy in repeated-measures design. *Journal of Memory and Language, 59*, 447–456.

Dunbar, R. (1998). *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.

Ernestus, M. (2013). *Halve Woorden*. Nijmegen, The Netherlands: Radboud Universiteit Nijmegen.

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*, 27–41.

Ernestus, M., Baayen, H. R., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language, 81*, 162–173. http://dx.doi.org/10.1006/brln.2001.2514.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics, 36*, 345–360. http://dx.doi.org/10.1016/j.wocn.2007.11.002.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279. http://dx.doi.org/10.1037//0033-295X.105.2.251.

Groen, M., & McQueen, J. M. (2014). Phonological representations in adults with dyslexia: Evidence from a perceptual learning paradigm. In *Abstracts from AMLaP XX*. Edinburgh.

Hayes, D. P. (1988). Speaking and writing: Distinct patterns of word choice. *Journal of Memory and Language, 27*, 572–585.

Hualde, J. I. (2004). Quasi-phonemic contrasts in Spanish. In *WCCFL 23: Proceedings of the 23rd west coast conference on formal linguistics* (pp. 374–398). <http://prosody.beckman.illinois.edu/jihualde/objects/pubs/Hualde-WCCFL23_3.pdf>.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*, 1–10.

Hura, S. L., Lindblom, B., & Diehl, R. (1992). On the role of perception in shaping phonological assimilation rules. *Language and Speech, 35*, 59–72.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. http://dx.doi.org/10.1016/j.jml.2007.11.007.

Janse, E., Nooteboom, S., & Quené, H. (2007). Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition. *Language and Cognitive Processes, 22*, 161–200.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama, & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis. Proceedings of the 1st session of the 10th international symposium* (pp. 29–54). Tokyo, Japan: The National International Institute for Japanese Language.

Keating, P. (1997). Word-level phonetic variation in large speech corpora. In B. Pompino-Marschal (Ed.), *An issue of ZAS working papers in linguistics*. Available as http://www.Humnet.Ucla.edu/humnet/linguistics/people/keating/berlin1.Pdf. <http://www.linguistics.ucla.edu/people/keating/berlin1.pdf>.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, H. R. (2004). Processing reduced word forms: The suffix restoration effect. *Brain and Language, 90*, 17–127. http://dx.doi.org/10.1016/S0093-934X(03)00425-5.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*, 643–650. http://dx.doi.org/10.3758/brm.42.3.643.

Kohler, K. J. (1996). Glottal stop and glottalization – A prosody in European languages. *AIPUK, 30*, 207–216.

Maas, U. (1999). *Phonologie. Einführung in die funktionale Phonetik des Deutschen*. Opladen/Wiesbaden: Westdeutscher Verlag.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–86. http://dx.doi.org/10.1016/0010-0285(86)90015-0.

McQueen, J. M., Jesse, A., & Norris, D. (2009). No lexical–prelexical feedback during speech perception or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language, 61*, 1–18.

McQueen, J. M., Tyler, M. D., & Cutler, A. (2012). Lexical retuning of children's speech perception: Evidence for knowledge about words' component sounds. *Language Learning and Development, 8*, 317–339. http://dx.doi.org/10.1080/15475441.2011.641887.

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*, 322.

Metsala, J. L., & Walley, A. C. (1998). *Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability*. <http://psycnet.apa.org/psycinfo/1998-07737-004>.

Mielke, J. (2002). Turkish /h/ deletion: Evidence for the interplay of speech perception and phonology. In *Proceedings-NELS* (Vol. 32, pp.

383–402). <http://www.zas.gwz-berlin.de/fileadmin/material/ZASPiL_Volltexte/zp28/zaspil28-mielke.pdf>.

Mitterer, H. (2011). Recognizing reduced forms: Different processing mechanisms for similar reductions. *Journal of Phonetics, 39*, 298–303. http://dx.doi.org/10.1016/j.wocn.2010.11.009.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics, 65*, 956–969. http://dx.doi.org/10.3758/BF03194826.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science, 30*, 451–479. http://dx.doi.org/10.1207/s15516709cog0000_57.

Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/s that speakers lenite: Evidence from /t/-lenition in Dutch. *Journal of Phonetics, 34*, 73–103. http://dx.doi.org/10.1016/j.wocn.2005.03.003.

Mitterer, H., & McQueen, J. M. (2009). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 244–263. http://dx.doi.org/10.1037/a0012730.

Mitterer, H., & Russell, K. (2013). How phonological reductions sometimes help the listener. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 977–984. http://dx.doi.org/10.1037/a0029196.

Mitterer, H., Scharenborg, O., & McQueen, J. M. (2013). Phonological abstraction without phonemes in speech perception. *Cognition, 129*, 356–361. http://dx.doi.org/10.1016/j.cognition.2013.07.011.

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

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189–234. http://dx.doi.org/10.1016/0010-0277(94)90043-4.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238. http://dx.doi.org/10.1016/S0010-0285(03)00006-9.

Ohk, B. (2006). *Dizzjunarju Ġermaniż -Malti: Wörterbuch Deutsch-Maltesisch*. Marsa, Malta: Grima Print.

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience, 21*, 169–179.

Perre, L., Bertrand, D., & Ziegler, J. C. (2011). Literacy affects spoken language in a non-linguistic task: An ERP study. *Frontiers in Psychology, 2*. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3198050/>.

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research, 1275*, 73–80.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp. 101–139). Berlin: Mouton de Gruyter.

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language, 61*, 19–36.

Poellmann, K., Bosker, H. R., McQueen, J. M., & Mitterer, H. (2014). Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch. *Journal of Phonetics, 46*, 101–127.

Racine, I., Bürki, A., & Spinelli, E. (2014). The implication of spelling and frequency in the recognition of phonological variants: Evidence from pre-readers and readers. *Language, Cognition and Neuroscience, 29*, 893–898.

Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language, 57*, 273–298.

Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1588–1594. http://dx.doi.org/10.1037/a0024833.

Reinisch, E., Wozny, D. R., Mitterer, H., & Holt, L. L. (2014). Phonetic category recalibration: What are the categories? *Journal of Phonetics, 45*, 91–105. http://dx.doi.org/10.1016/j.wocn.2014.04.002.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H.-J., ... Schulze-Bonhage, A. (2013). Co-localizing linguistic and musical syntax with intracranial EEG. *NeuroImage, 64*, 134–146.

Scobbie, J. M., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In *Contrast: Perception and acquisition: Selected papers from the second international conference on contrast in phonology* (pp. 87–113). Berlin: Mouton de Gruyter. <http://books.google.com/books?hl=en&lr=&id=LDMxpO1PMk0C&oi=fnd&pg=PA87&dq=fuzzy+inventory+scobbie&ots=X2vMD_flpA&sig=i57s5VqvHXloKEBaE1Y9U8m2y5o>.

Scott, S. K., & Wise, R. J. S. (2004). The functional neuroanatomy of prelexical processing in speech perception. *Cognition, 92*, 13–45.

Stanovich, K. E., & Cunningham, A. E. (2001). What reading does for the mind. *Journal of Direct Instructions, 1*, 137–149.

Steriade, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). New York, NJ: Academic Press.

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.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*, 2020–2045. http://dx.doi.org/10.1037/xge0000014.

Wiese, R. (1996). *The phonology of German*. Oxford: Clarendon Press.

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.

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition, 32*, 732–741.