



HAL
open science

Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade?

Johannes Ziegler, Ana Petrova, Ludovic Ferrand

► To cite this version:

Johannes Ziegler, Ana Petrova, Ludovic Ferrand. Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2008, 34 (3), pp.643-661. 10.1037/0278-7393.34.3.643. hal-02960241

HAL Id: hal-02960241

<https://hal.science/hal-02960241>

Submitted on 29 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Running Head: FEEDBACK CONSISTENCY

**Feedback Consistency Effects In Visual And Auditory Word Recognition: Where Do
We Stand After More Than A Decade?**

Johannes C. Ziegler^{1,2}, Ana Petrova³ and Ludovic Ferrand^{2,4}

¹Aix-Marseille Université

²Centre National de la Recherche Scientifique

³Université Paris Descartes

⁴Université Blaise Pascal

Journal of Experimental Psychology: Learning, Memory, and Cognition (2008).
<https://doi.org/10.1037/0278-7393.34.3.643>

Please send correspondence to:

Johannes C. Ziegler, PhD
Laboratoire de Psychologie Cognitive
Pôle 3C, Case D
CNRS and Université de Provence,
3 place Victor Hugo
13331 MARSEILLE Cedex 3

Phone +33 4 88 57 68 99

Fax +33 4 88 57 68 95

Email: Johannes.Ziegler@univ-provence.fr

Abstract

The role of phonology-to-spelling consistency (i.e., *feedback consistency*) was investigated in three lexical decision experiments both in the visual and auditory modality in French and English. No evidence for a feedback consistency effect was found in the visual modality, neither in English nor in French, and despite the fact that consistency was manipulated for different kinds of units (onsets and rimes). In contrast, robust feedback consistency effects were obtained in the auditory lexical decision task using exactly the same items that produced a null effect in the visual modality, both in English and in French. Neural network simulations are presented to show that previous demonstrations of feedback consistency effects in the visual modality can be simulated with a model that is not sensitive to feedback consistency suggesting that these effects might have come from various confounds. These simulations together with our results suggest that there are no feedback consistency effects in the visual modality. In contrast, such effects are clearly present in the auditory modality. Given that orthographic information is absent from current models of spoken word recognition, the present findings present a major challenge to these models.

Keywords: Feedback consistency, Word recognition, Lexical decision, Orthography, Phonology

More than a decade ago, Stone, Vanhoy and Van Orden (1997) put forward one of the most intriguing and counter-intuitive hypotheses in the history of visual word recognition. According to this hypothesis, visual word recognition is not only influenced by the consistency of the mapping between spelling and sound (ie., whether orthography is pronounced consistently) but also by the consistency of the mapping between sound and spelling (whether phonology is spelled consistently). These bidirectional consistency effects were predicted in the context of Stone and Van Orden's (1994) recurrent network theory of word perception (see Figure 1). In a recurrent network, the flow of activation is inherently bidirectional. Consistent symmetrical relations result in stable and fast activation, whereas inconsistent and asymmetrical relations slow down the system on its way to equilibrium (Tuller, Case, Ding, & Kelso, 1994; Van Orden, 2002; Van Orden, Jansen op de Haar, & Bosman, 1997). Thus, according to the model, inconsistency in both directions (spelling-to-sound and sound-to-spelling) should slow down word recognition. In reference to this model, spelling -> sound consistency was called *feedforward consistency*, whereas sound -> spelling consistency was called *feedback consistency*.

Insert Figure 1 here

The novelty of Stone et al.'s proposal was the existence of a feedback consistency effect (feedforward consistency effects had been observed previously but only in reading aloud, e.g., Jared, 2002; Jared, McRae, & Seidenberg, 1990; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Taraban & McClelland, 1987). The existence of feedback consistency effects was demonstrated in the visual lexical decision task. Stone et al. (1997) manipulated feedback consistency at the rime level. That is, in the feedback inconsistent condition, words contained rimes that could potentially be spelled in multiple ways (e.g., /i:p/ could be spelled

EEP or EAP). In the feedback consistent condition, words contained rimes that were always spelled the same way (e.g., /ʌk/ is always spelled UCK). Stone et al. found that lexical decisions to feedback inconsistent words were, on average, 33 ms longer than those to consistent words. Moreover, feedback inconsistent words produced more errors than feedback consistent words (9.8% versus 3.3%).

The existence of a feedback consistency effect was surprising. From a traditional information processing view, processes should only flow downstream, as from spelling to phonology. It should not matter in *visual* word recognition that a pronunciation may have more than one spelling. Thus, the existence of feedback consistency effects was perceived as a major theoretical challenge to more traditional bottom-up theories of word recognition (Norris, McQueen, & Cutler, 2000).

Replications And Failures To Replicate

Shortly after Stone et al.'s (1997) seminal article, Ziegler, Montant and Jacobs (1997) were the first to replicate feedback consistency effects in French, a language with a high degree of feedback inconsistency (see Ziegler, Jacobs, & Stone, 1996). As Stone et al. (1997), they manipulated rime consistency and found that visual lexical decisions to feedback inconsistent words took, on average, 33 ms longer than lexical decisions to feedback consistent words. Moreover, Ziegler et al. (1997) also found a feedback consistency effect in the naming task, but the effect was smaller and less reliable than in lexical decision.

Shortly after these two publications, Peereman, Content and Bonin (1998) challenged the feedback consistency hypothesis. In a series of five experiments conducted in French, they showed that feedback consistency affected orthographic processes in writing but not lexical decision. To address the discrepancy between their findings and those of Ziegler et al., they

successfully replicated Ziegler et al.'s feedback consistency effect when using Ziegler et al.'s items. However, they then argued that Ziegler et al.'s finding was due to a confound between feedback consistency and subjective (rated) familiarity. Indeed, when subjective familiarity was partialled out, the feedback consistency effect disappeared.

In English, evidence for feedback consistency effects looked somewhat stronger. Although Stone et al.'s (1997) experiments were criticized by Peereman et al. (1998) for potential familiarity and frequency confounds, two recent studies reported robust feedback consistency effects in English with frequency and subjective familiarity being controlled for (Lacruz & Folk, 2004; Perry, 2003). Indeed, Lacruz and Folk (2004) reported significant feedforward and feedback consistency effects both in lexical decision and naming. Items were either feedforward inconsistent, feedback inconsistent or both (8 items per group). The three groups were controlled not only for frequency but also for subjective familiarity and a number of other variables, such as word length, orthographic neighborhood, bigram frequency, and summed frequency of friends. In lexical decision, feedback consistency effects were significant for both high- and low-frequency words (32ms and 23 ms, respectively). Exactly the same pattern was found in the naming task. Feedback consistency effects were present for both high-and low-frequency words.

Although the data by Lacruz and Folk (2004) seemed very convincing and robust, their results have been criticized in a recent study by Massaro and Jesse (2005). They argued that Lacruz and Folk (2004) as well as all previous studies have failed to match consistent and inconsistent words on the initial onset cluster. While it is difficult to see how the failure to match for the onset could affect Lacruz and Folk's lexical decision data, it is clearly a problem for the naming results. In their study, Massaro and Jesse (2005) used Stone et al.'s (1997) original 2 x 2 factorial manipulation of feedforward and feedback consistency (18 items per cell). Items were matched for initial phonemes, frequency, subjective familiarity,

and bigram frequency. They also used two types of naming tasks: a post-vocalic immediate naming task and a delayed naming task with variable ISI. Their results showed feedback consistency effects that were significant by subjects (not by items). However, these feedback consistency effects were significant both in immediate and in delayed naming. Massaro and Jesse therefore argued that the feedback consistency effects must be due to post-perceptual (e.g., articulatory) processes.

One inevitable problem is that feedback consistency must be manipulated between items. Therefore, uncontrolled and/or confounded variables might explain part of the feedback consistency effect. This issue is particularly damaging in the study of Lacruz and Folk (2004) because of the small number of items per cell (8 in their study). A solution to this problem is to investigate the existence of feedback consistency effects in large-scale databases (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Because such databases of lexical decision and naming performance contain a large sample of words, confounded factors can be partialled out more effectively. Indeed, Balota et al. (2004) investigated in great detail the effects of feedback consistency both for the onset and the rime unit. They found that feedback consistency explained a small but significant portion of the variance but only in naming not in lexical decision. This finding is opposite to that of Ziegler et al. (1997) who found stronger effects in lexical decision than in naming. Clearly, if lexical decision were to encourage feedback processes for the purpose of spelling verification (Ziegler, Jacobs, & Klueppel, 2001), one would predict stronger feedback consistency effects in lexical decision than in naming.

Finally, Kessler, Treiman & Mullennix (2007) analyzed data from four megastudies. They first conducted an ordinary regression using a variety of quantitative measures as independent variables (e.g., frequency, familiarity, neighborhood, bigram frequency, feedforward consistency etc.). In this first step, feedback consistency was not included. They

then calculated the residual variance (difference between predicted and observed response time values) and checked how much of the residual variance was accounted for by feedback consistency. They found that feedback consistency had a significant effect in only one out of 10 comparisons. However, one problem with this approach is that feedback consistency is naturally confounded with variables, such as orthographic neighborhood or bigram frequency (feedback inconsistent words tend to have few neighbors and low-frequency bigrams). By using such highly inter-correlated variables in the first step of the regression, the chance of finding a significant feedback consistency effect is of course reduced.

Feedback Consistency Effects In The Auditory Modality

The difficulty to obtain reliable feedback consistency effects in the visual modality contrasts with the stability of the effect in the auditory modality. Up to now, a dozen of studies have reported such effects in different languages and tasks (Chereau, Gaskell, & Dumay, 2007; K. M. Miller & Swick, 2003; Pattamadilok, Morais, Ventura, & Kolinsky, 2007; Pattamadilok, Perre, Dufau, & Ziegler, 2008; Perre & Ziegler, 2008; Slowiaczek, Soltano, Wieting, & Bishop, 2003; Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008; Ventura, Morais, & Kolinsky, 2007; Ventura, Morais, Pattamadilok, & Kolinsky, 2004; Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2004; Ziegler & Muneaux, 2007; Ziegler, Muneaux, & Grainger, 2003). In a first study, Ziegler and Ferrand (1998) manipulated the feedback consistency of the rime in an auditory lexical decision task in French. Note that “feedback consistency” is defined similarly regardless of the modality. That is, both in the visual and the auditory modality, feedback consistency is defined as the consistency between phonology and spelling (see Figure 1). In the feedback inconsistent condition, spoken words had rimes that could be spelled in multiple ways. In the feedback consistent condition, words had rimes that could only be spelled one way. Ziegler and Ferrand

(1998) found that auditory lexical decisions to inconsistent words took longer and yielded more errors than those to consistent words. These effects were obtained despite the fact that rated familiarity was taken into account.

The existence of feedback consistency effects in the auditory modality was replicated in numerous experiments. First, in a study in Portuguese, Ventura et al. (2004) found that words with rimes that can be spelled in two different ways (inconsistent condition) produced longer auditory lexical decision latencies and more errors than did consistent words. The consistency effect was not obtained in the shadowing task, suggesting that lexical involvement is required for the effect to occur.

Second, the auditory feedback consistency effect was replicated in an auditory lexical decision task with a graded consistency manipulation (Ziegler et al., 2004). In this study, rime phonology was held constant and the probability/frequency with which phonology mapped onto spelling was manipulated. Half of the words had a dominant (frequent) spelling pattern, whereas the other half had a subdominant (rare) spelling pattern. The results showed that lexical decisions to words with subdominant spellings took longer and were more error-prone than lexical decisions to words with dominant spellings. Because rime phonology was identical for dominant/subdominant pairs, this finding not only replicated the auditory feedback consistency effect but also excluded the possibility that phonetic/phonological factors would explain its occurrence.

Finally, Ziegler and Muneaux (2007) showed that orthographic influences on spoken word recognition are tightly linked to the acquisition of literacy. In a developmental study, they showed that prior to literacy, auditory lexical decisions were not influenced by the spellings of spoken words. In contrast, as soon as literacy developed, spoken word recognition became affected by the orthographic similarity/consistency of spoken words (see also

Goswami, Ziegler, & Richardson, 2005). As early as grade 1, the size of the feedback consistency effect was predicted by the reading level of a child.

Two explanations have been proposed to explain these effects. The first assumes that orthography is activated on-line during spoken word recognition (Perre & Ziegler, 2008). The idea is that during learning to read, strong and permanent associations form between orthography and phonology. From then on, the processing of visual and spoken language is tightly linked through a single network that binds the orthographic and phonological aspects of words. As a consequence, orthographic information is co-activated on-line whenever we hear a spoken word. If a word has multiple spellings, the different orthographic patterns compete slowing down spoken word recognition. The second explanation assumes that orthographic consistency plays a major role during the restructuring of phonological representations (Muneaux & Ziegler, 2004; Ziegler & Goswami, 2005). The idea is that orthographically consistent words develop better and more fine-grained phonological representations in the course of reading development. Thus, orthography would not be co-activated in an on-line fashion but rather influence the quality of phonological representations at an earlier stage. Which of these explanations turns out to be correct is still a matter of debate.

In sum, in contrast to the feedback consistency effect in the visual modality, feedback consistency effects in the auditory modality appear quite robust. Of course, the feedback consistency effect in the auditory modality is somewhat more direct than the one in the visual modality because the auditory effect only requires a simple “one-way” feedback from phonology to orthography (i.e., co-activation). The feedback consistency effect in the visual modality is less direct because it necessitates two steps: phonology needs to be activated first before it can feed back to orthography (see Figure 1).

The Present Study

The goal of the present study was to reinvestigate feedback consistency effects in the visual and auditory modality. One possible explanation for the weakness of the feedback consistency effect in the visual modality is related to the *nature* of the units that have been manipulated. Indeed, most previous studies manipulated feedback consistency at the level of the rime (e.g., Stone et al., 1997; Ziegler et al., 1997). Clearly, the rime is a salient unit in phonology (De Cara & Goswami, 2002) but not necessarily in orthography (Peereman & Content, 1997; Rey, Ziegler, & Jacobs, 2000). Thus, manipulating the orthographic properties of the rime might produce more stable results in the auditory modality than in the visual modality.

The nature of the units might also explain why the feedback consistency effect in the visual lexical decision task appears to be more robust in English (three studies found reliable effects) than in French. That is, in English, the rime plays a special role not only in phonology but also in orthography (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). The reason for this is that the orthographic rime disambiguates the inconsistency of the smaller vowel units. In English, the degree of consistency (consistency ratio) of the vowel unit is only .51, on average. However, if one takes into account the consonants that follow the vowel, the consistency increases to around .80. This situation is different in French because the vowel unit is much more consistent in French (i.e., .91). Therefore, the benefits of taking the rime into account are necessarily smaller in French than in English (the consistency ratio in French increases from .91 to .95 by taking the rime into account, see Peereman & Content, 1997). Thus, cross-language differences could explain why feedback consistency effects in the visual modality are less robust in languages other than English (see also Goswami et al., 2005).

The goal of the present study, therefore, was to investigate feedback consistency effects in the visual and the auditory domain by studying both onset and rime consistency. Indeed, many studies suggest that word initial information plays an important role in both visual and auditory processing (Coltheart & Rastle, 1994; Coltheart, Woollams, Kinoshita, & Perry, 1999; Connine, Blasko, & Titone, 1993; Forster & Davis, 1991; Lima & Inhoff, 1985; Marslen-Wilson & Welsh, 1978; Vitevitch, 2002). For example, in the visual modality, the spelling-to-sound regularity of the onset is more important than that of units occurring after the onset (Rastle & Coltheart, 1999). In the auditory modality, words with a competitor in word-initial positions produce stronger interference than words with a competitor in word-final positions (Allopenna, Magnuson, & Tanenhaus, 1998).

In Experiment 1, we used a factorial design to manipulate both the consistency of the onset (consonant-vowel cluster) and the consistency of the rime (vowel-consonant cluster) in a visual lexical decision task. If onset consistency is more important than rime consistency, in French, then we should obtain a reliable feedback consistency effect for the onset. In Experiment 2, we directly contrast feedback consistency effects in the visual and auditory domain using the same items in a visual and an auditory lexical decision task. None of the published experiments has directly compared the feedback consistency in the visual and the auditory modality using the same items. Finally, in Experiment 3, we test the idea that feedback consistency effects in the auditory domain are generally more stable because of the fact that, in the auditory domain, feedback from phonology to spelling is more direct than in the visual domain. If bidirectional feedback effects are generally difficult to obtain, regardless of modality, then feedforward (spelling-to-sound) consistency effects in the auditory domain should not be easier to obtain than feedback (sound-to-spelling) consistency effects in the visual domain. This experiment was conducted in English because the manipulation of bidirectional consistency is more limited in French.

Experiment 1

The goal of the experiment was to investigate whether one could obtain robust feedback consistency effects in the visual lexical decision task in French when the consistency manipulation was extended to the beginning of words. Thus, feedback consistency was manipulated in a 2×2 factorial design with onset consistency and rime consistency as factors.

Methods

Participants

Seventy-six native speakers of French took part in Experiment 1, 36 participated in the lexical decision task and 40 gave familiarity ratings for the items. The participants were psychology students at the Université René Descartes in Paris.

Items

The factorial manipulation of onset and rime consistency resulted in four groups: 1/ Onset and Rime consistent (ON con/RI con)¹, 2/ Onset inconsistent and Rime consistent (ON inc/RI con), 3/ Onset consistent and Rime inconsistent (ON con/RI inc) and 4/ Onset and Rime inconsistent (ON inc/RI inc). Each group contained 25 words (see Appendix A for a complete list). Items were in the low-to medium frequency range. We avoided extremely low-frequency items (< 2 per million) because previous research has shown that such items produce error rates up to 80% (Peereboom et al., 1997). Such high-error items would then need to be excluded anyway. Items were selected on the basis of statistical analyses of feedback consistency in French (Ziegler et al., 1996). The four groups were matched on number of

¹ Note that in this and all following experiments some consistent words might occasionally have a very infrequent *enemy*, which is the reason why the consistency ratio is not quite 1.0. However, all consistent items have at least 9 times more spelling friends than enemies, that is, a consistency ratio greater than .90.

letters, word frequency according to two word frequency counts (LEXIQUE, New, Pallier, Ferrand, & Matos, 2001; LEXOP, Peereman & Content, 1999), orthographic neighborhood, and feedforward consistency. Finally, we asked participants to rate the familiarity of the words on a 7-point Likert scale. Item characteristics are given in Table 1.

Insert Table 1 here

One hundred orthographically legal and pronounceable nonwords were created by changing a single letter (either at the beginning, the middle or the end) from existing words, which were matched in frequency and length to the critical words. Thus, the nonwords were tightly matched to real words in terms of length and neighborhood characteristics. On average, the nonwords had 3.8 orthographic neighbors, which is comparable to the neighborhood characteristics of the real words. Also, none of the nonwords contained any illegal bigrams. No consistency manipulation was performed on the nonwords.

Procedure

Participants were tested individually in a sound-proof room. A classic lexical decision task without feedback was used. This and the following experiments were controlled using DMDX software (Forster & Forster, 2003). Both speed and accuracy were emphasized. Each trial was preceded by a fixation cross (500ms). The stimulus remained visible on the screen until the participant responded using the computer keyboard. Participants received 10 practice trials, for which feedback was provided.

Results

Five items were excluded due to excessively high error rates (> 50%, see Appendix A). The four word groups were still perfectly matched on frequency and all other potentially confounding variables. In fact, the item statistics reported in Table 1 were re-computed after eliminating these 5 items. The remaining data were trimmed according to a three standard deviation (SD) cut-off (1.6% outliers). Data were analyzed in a 2×2 analysis of variance (ANOVA) with *Onset Consistency* and *Rime Consistency* as factors. Analyses were performed for participants (F1) and items (F2). Results are presented in Table 2.

Insert Table 2 here

In the latency data, there was no significant effect of either Onset Consistency, $F(1, 35) = .24, p > .60$, $F(1, 91) = .12, p > .70$, or Rime Consistency, $F(1, 35) = .23, p > .60$; $F(1, 91) = .07, p > .70$. The interaction between the effects of Onset and Rime consistency failed to reach significance, $F(1, 35) = 2.8, p > .10$; $F(1, 91) = .26, p > .60$.

In the error data, the main effect of Onset Consistency was not significant, $F(1, 35) = 1.89, p > .10$; $F(1, 91) = .52, p > .40$. The main effect of Rime Consistency was significant by participants but not by items, $F(1, 35) = 10.51, p < .01$; $F(1, 91) = 1.8, p > .18$. The interaction between these two variables was not significant, $p > .70$.

In summary, the present experiment showed no convincing evidence for the existence of feedback consistency effects in the visual lexical decision task. Only the error data exhibited a small effect of Rime Consistency (4% versus 6%, for feedback consistent versus inconsistent words, respectively). However, this effect was only significant by participants not by items.

One potential problem with the present null finding is the fact that the present experiment used items from the low-to-medium frequency range, whereas previous studies

tended to use items of extremely low frequency (e.g., Stone et al., 1997; Ziegler et al., 1997). Given that consistency effects tend to be more reliable for low-frequency words (Seidenberg et al., 1984, but see Jared, 2002), it could be possible that we missed a feedback consistency effect in the present experiment. Because of the relatively large number of items in each cell, we were able to address this problem in a post-hoc analysis using only items with a frequency smaller than 20 per million resulting in an average frequency of 6.0 per million according to LEXIQUE and 7.2 per million according to LEXOP. This selection left us with 16 items per cell, which is still twice as many as in Lacruz and Folk's (2004) study. Also, the four groups were still perfectly matched for frequency on both databases and subjective familiarity (all $ps > .20$). The results of this post-hoc analysis are presented in Table 3.

Insert Table 3 here

As before, the data were submitted to a 2×2 ANOVA with Onset Consistency and Rime Consistency as factors. The latency data showed no significant effect of either Onset or Rime consistency, and no significant interaction between the two factors, all $F_s < 1$, $ps > .50$. Similarly, the error data showed no significant main effects or interactions, all $F_s < 1.8$, $ps > .18$.

Discussion

The goal of the present experiment was to give feedback consistency a second chance. That is, in the present experiment, we manipulated feedback consistency of the onset as well as the rime. If previous studies in French yielded unreliable results because onset consistency had not been controlled for or because onsets are more important units than rimes, then we

should have found clear feedback consistency effects for either unit. However, this was not the case. The present experiment showed no compelling evidence for the existence of a feedback consistency effect in the visual lexical decision task in French. Note that items in the present experiment were more carefully controlled than items in previous experiments. For example, the groups were matched on recent frequency counts, and, unlike the items of Ziegler et al. (1997), they were matched on rated familiarity. Thus, the results of the present experiment join those of Peereman et al. (1998) to suggest that feedback consistency effects in the visual modality are unreliable.

Experiment 2

Fragile feedback consistency effects in the visual lexical decision task stand in sharp contrast with robust feedback consistency effects in the auditory lexical decision task (see references in the introduction). The goal of this experiment was to investigate feedback consistency using the same items in the visual and auditory lexical decision task. If the feedback consistency effect were modality-specific, then we should obtain robust feedback consistency effects in the auditory lexical decision task with the same items that produce no feedback consistency effects in the visual lexical decision task. Experiment 2 replicated the design of Experiment 1. A partially new item set was selected for Experiment 2. This was done because the four groups had to be matched for phonological neighborhood size in order to use the same items in the auditory lexical decision task. Indeed, phonological neighborhood has a strong inhibitory effect on auditory lexical decision (Goldinger, Luce, & Pisoni, 1989; Luce, Pisoni, & Goldinger, 1990; Vitevitch & Luce, 1998; Ziegler et al., 2003). Apart from these changes, the design of Experiment 2 was identical to that of Experiment 1.

Methods

Participants

Ninety-five university students participated in the experiment (38 in the visual LDT and 57 in the auditory LDT). None of them had participated in Experiment 1.

Items

We used the same design as in Experiment 1 but we matched the four groups much more tightly on phonological variables than was the case in Experiment 1. Item characteristics are presented in Table 4, all items are listed in Appendix B. It was very difficult to simultaneously match for orthographic and phonological neighborhood. Thus, the choice was made to give preference to the control of phonological neighborhood. As a consequence, feedback consistent words had slightly more orthographic neighbors than inconsistent words. This is not a problem, however, because words with many orthographic neighbors produce faster latencies in the visual lexical decision task than words with few orthographic neighbors. Therefore, if anything, the fact that consistent words had more orthographic neighbors should be in favor of finding feedback consistency effects in the visual lexical decision task. The nonwords were the same as in Experiment 1.

Insert Table 4 here

Procedure

The procedure was identical to Exp 1. In the auditory LDT, the clock was started at the onset of the auditory stimulus and was stopped when the participant responded.

Results

In visual lexical decision task, three items were excluded because of error rates above 50% (see Appendix B). Data were trimmed according to a 3 SD cut-off, which affected 2.01% of the auditory data and 1.6% of the visual data.

The results for both the auditory and visual modality are presented in Table 5. The results clearly show feedback consistency effects for both units in the auditory modality but not the visual modality. The data were analyzed in two ways. First, we conducted a global three factorial ANOVA with Modality (visual versus auditory), Onset Consistency (consistent versus inconsistent) and Rime Consistency (consistent versus inconsistent) as factors. Then, for each modality, we conducted a 2×2 ANOVA with Onset Consistency and Rime Consistency as factors.

Insert Table 5 here

Global Analyses

The latency data exhibited a main effect of modality, $F(1, 93) = 341.6, p < .0001$; $F(1, 89) = 1506.1, p < .0001$, which reflects the fact that RTs were faster in the visual than in the auditory modality. The main effects of Onset Consistency and Rime Consistency were significant by subjects but not by items (Onset: $F(1, 93) = 7.78, p < .05$; $F(1, 89) = .37, p > .50$; Rime: $F(1, 93) = 16.81, p < .0001$; $F(1, 89) = 1.18, p > .30$). More importantly, the main effects of Onset and Rime Consistency were qualified by significant interactions with Modality (Onset \times Modality, $F(1, 93) = 26.21, p < .0001$; $F(1, 89) = 3.84, p < .05$; Rime \times Modality, $F(1, 93) = 28.63, p < .0001$; $F(1, 89) = 4.91, p < .05$). These interactions confirm that feedback consistency effects were present in the auditory modality but absent in the

visual modality. Onset Consistency interacted with Rime Consistency in the analysis by participants, $F(1, 93) = 19.23, p < .0001$; $F(1, 89) = .88, p > .30$. The triple interaction between the effects of Modality, Onset Consistency and Rime Consistency was significant because doubly consistent words were slowest in the visual lexical decision task but fastest in the auditory task, $F(1, 93) = 63.90, p < .0001$; $F(1, 89) = 6.21, p < .05$.

The error data showed a similar pattern with a main effect of Modality, $F(1, 93) = 12.36, p < .001$; $F(1, 89) = 4.99, p < .05$, reflecting the fact that participants made more errors in the auditory modality. The effects of Onset Consistency and Rime Consistency failed to reach significance by items (Onset: $F(1, 93) = 10.37, p < .01$; $F(1, 89) = .71, p > .50$; Rime: $F(1, 93) = 1.46, p > .20$; $F(1, 89) = .36, p > .50$). As before, the main effects of Onset and Rime Consistency interacted with Modality (Onset \times Modality: $F(1, 93) = 24.43, p < .0001$; $F(1, 89) = 4.02, p < .05$; Rime \times Modality: $F(1, 93) = 3.78, p < .05$; $F(1, 89) = 1.45, p > .10$). These two-way interactions reflect the fact that feedback consistency effects were present in the auditory modality but not in the visual modality. The Onset Consistency by Rime Consistency interaction was significant by participants, $F(1, 93) = 5.28, p < .05$; $F(1, 89) = .62, p > .40$, and the triple interaction was not significant, all $F_s < 1$.

In summary, the present analyses clearly confirm that onset consistency and rime consistency interact with modality (significant by subjects and items). These interactions reflect the fact that feedback consistency effects were present in the auditory modality but not in the visual modality. Below, the feedback consistency effects are analyzed for each modality, separately.

Visual Modality Only

The data were submitted to a 2×2 ANOVA with Onset Consistency and Rime Consistency as factors. In the latency data, Onset Consistency approached significance by participants but the effect went in the wrong direction with slower responses to consistent than to inconsistent words, $F(1, 37) = 3.12, p = .085$; $F(1, 89) = .61, p > .40$. The Rime Consistency effect was not significant, all $F_s < 1$. The Onset \times Rime interaction was significant by participants but not by items, $F(1, 37) = 9.42, p < .01$; $F(1, 89) = .66, p > .40$. In the error data, none of the main effects were significant, all $p_s > .20$. The Onset \times Rime interaction was significant by participants but not by items, $F(1, 37) = 5.24, p < .05$; $F(1, 89) = 1.10, p > .30$.

Auditory Modality Only

The latency data exhibited main effects of Onset Consistency, $F(1, 56) = 33.05, p < .0001$; $F(1, 92) = 2.85, p < .10$, Rime Consistency, $F(1, 56) = 58.60, p < .0001$; $F(1, 92) = 2.26, p > .10$, and a significant interaction between the effects of Onset and Rime Consistency, $F(1, 56) = 74.64, p < .0001$; $F(1, 92) = 4.19, p < .05$. This interaction reflects the fact that the effects of onset and rime consistency were not additive. The error data showed a significant effect of Onset Consistency, $F(1, 56) = 33.62, p < .0001$; $F(1, 92) = 5.35, p < .05$, Rime Consistency by participants, $F(1, 56) = 4.76, p < .05$; $F(1, 92) = .89, p > .30$, and no significant interaction, $F(1, 21) = 1.17, p > .20$; $F(1, 92) = .15, p > .60$.

In summary, the within-modality analyses confirm the results from the global ANOVAs. That is, no reliable feedback consistency effects were obtained in the visual lexical decision task (the two-way interaction is marginally significant but the effects go in the wrong direction), whereas significant main effects and significant two-way interactions were

obtained in the auditory lexical decision task. The significant two-way interaction reflects the fact that doubly inconsistent words did not produce more costs than singly inconsistent words.

Fast Auditory Versus Slow Visual Responders

One possible explanation for the finding that feedback consistency effects are more reliable in the auditory than in the visual modality is the fact that auditory lexical decisions take, on average, 300 ms longer than visual lexical decisions. This is of course due to the fact that participants do not have all the information to make an auditory lexical decision "right away" as they do when the words are presented visually. If feedback effects need time to build up, then it would be expected that feedback consistency effects in auditory lexical decision are more reliable than in visual lexical decision. This hypothesis can directly be tested in a post-hoc analysis by comparing the fastest participants from the auditory lexical decision task with the slowest participants from the visual lexical decision task, thus matching for overall response latency. In order to match for overall response latencies across modalities, we had to collapse the visual lexical decision data from Experiments 1 and 2. The results of this post-hoc analysis are presented in Table 6.

Insert Table 6 here

As can be seen in Table 6, feedback consistency effects were again obtained in the auditory modality but not the visual modality. In the auditory modality, significant effects were obtained for Onset Consistency ($F(1, 9) = 40.70, p < .0001$) and Rime Consistency ($F(1, 9) = 35.36, p < .0001$). In contrast, in the visual lexical decision task, neither Onset

Consistency ($F(1, 9) = 0.0, p > .90$) nor Rime Consistency ($F(1, 9) = .66, p > .40$) produced significant effects². None of the two-way interactions were significant.

Linear Regression Analyses

Although consistent and inconsistent items were carefully matched on a number of word recognition variables, it is always possible that there are uncontrolled item-specific effects that either mask or amplify the desired effect. One way to address this issue is by running linear regression analyses to determine the unique amount of variance of a critical variable after entering all control variables in the first step of the linear regression analysis (Balota et al., 2004; Kessler et al., 2007). Such an analysis has the additional advantage of treating feedback consistency as a continuous (i.e., consistency ratios) rather than a categorical variable.

We therefore performed two linear regression analyses on the item means of Experiment 2, one for the visual modality and one for the auditory modality. Word frequency, word length, and neighborhood were entered in step 1 of the regression equation. In the visual modality, word length was the number of letters and neighborhood was the number of orthographic neighbors. In the auditory modality, word length was the number of phonemes and neighborhood was the number of phonological neighbors. We then determined how much unique variance was accounted for by feedback consistency in step 2. The results showed that feedback consistency accounted for 5.2% unique variance in the auditory modality ($\Delta F(1, 91) = 5.22, p = .025$), whereas it accounted for a small and nonsignificant amount of variance in the visual modality (0.4%; $\Delta F(1, 87) = .35, p = .55$). Given that one of the fundamental differences between visual and spoken word recognition is that spoken words are presented

² Note that exactly the same results were obtained when we restricted the post-hoc analysis to the 10 slowest readers of Experiment 2 only (onset consistency, $F(1, 9) = 1.6, p > .20$; rime consistency, $F(1, 9) = .12, p > .70$).

sequentially, we conducted an additional regression analyses taking into account duration, uniqueness point and frequency in step 1 of the regression equation. Feedback consistency still accounted for 4.1% of the unique variance in the auditory modality ($\Delta F(1, 91) = 5.7, p = .018$).

Increasing The Power

One concern is that we might not have had enough power to detect small feedback consistency effects in the visual modality. The *power* of Experiment 2 to detect a potential feedback consistency effect of 30 ms was .70 (Faul, Erdfelder, Lang, & Buchner, 2007). However, the power can be increased to .95 by combining the visual lexical decision data from Experiments 1 and 2 (total of 74 participants). We conducted an ANOVA on the collapsed data with Onset Consistency and Rime Consistency as factors. As before, there was no hint of a feedback consistency effect neither for the onset nor the rime (both $F_s < 1$).

Discussion

The novelty of the present experiment is that feedback consistency was investigated in the visual and auditory modality using the same items. Therefore, the present results provide convincing evidence that feedback consistency effects are strong and reliable in the auditory modality but weak and unreliable in the visual modality. This cross-modal dissociation is not due to the fact that auditory lexical decisions take longer than visual lexical decisions. When the fastest participants from the auditory experiment were compared with the slowest participants from the visual experiment (thus equating response speed across modalities), we still found a feedback consistency effect in the auditory but not the visual modality. Finally, in

a linear regression analysis, feedback consistency accounted for a significant amount of unique variance in the auditory but not in the visual modality.

Experiment 3

The previous experiment failed to show feedback consistency effects in the visual modality with exactly the same items that produced strong feedback consistency effects in the auditory modality. One explanation for the difference between modalities is that feedback consistency effects in the auditory modality are more “direct” than feedback consistency effects in the visual modality. Indeed, as can be seen in Figure 1, feedback consistency in the auditory modality could be considered *first-order* feedback: phonology is activated by the auditory input and feeds back to orthography (P->O feedback). However, feedback consistency in the visual modality could be considered *second-order* feedback: orthography is activated by the visual input, then orthography has to feed forward to phonology (O->P) before the *computed* phonology can feed back to orthography (P->O). One could argue that what is “feedback” in the visual modality is only “feed-forward” in the auditory modality.

To test the hypothesis that a first-order feedback effect is easier to obtain than a second-order feedback effect regardless of modality, we used Stone et al.’s (1997) original factorial manipulation of feedforward and feedback consistency in the visual and auditory modality. Thus, the critical comparison with regard to second-order feedback is the comparison between feed-forward (O->P) consistency in the auditory modality and feedback (P->O) consistency in the visual modality. In both cases, orthographic and phonological information has to “bounce back” once before affecting its target. If second-order feedback effects are difficult to obtain regardless of the modality, then we should find no sign of a feed-forward (O->P) consistency effect in the auditory modality. The present experiment had to be

done in English because it is difficult to manipulate bidirectional consistency in French using a reasonable number of items and controlling for the major word recognition variables.

Another advantage of running the present experiment in English is that it allows us to check whether our failure to find feedback consistency effects in the visual lexical decision task in Experiments 1 and 2 is due to cross-language differences between English and French.

Methods

Participants

Sixty-three native English speakers from the University of York participated in the experiment, 40 in the auditory lexical decision task and 23 in the visual lexical decision task.

Stimuli

Because previous consistency analyses were done for American English (Ziegler, Stone, & Jacobs, 1997), we recalculated the consistency values for British English using CELEX (Baayen, Piepenbrock, & van Rijn, 1993). Because the feedback consistency effect in English has been successfully obtained manipulating rime consistency (Lacruz & Folk, 2004; Stone et al., 1997), we manipulated rime consistency instead of onset consistency.

Eighty monosyllabic items were selected from the CELEX database, 20 items per group. All items are listed in Appendix C. The four groups resulted from a factorial manipulation of feed-forward (FF) and feedback (FB) consistency: 1/ FF-consistent/FB-consistent, 2/ FF-inconsistent/FB-consistent, 3/ FF-consistent/FB-inconsistent, and 4/ FF-inconsistent/FB-inconsistent. A word was defined as feed-forward consistent when its rime spelling could be pronounced in different ways (regardless of modality). A word was feedback inconsistent when a rime phonology could be spelled in different ways (regardless

of modality). Consistency ratios (friends/friends+enemies) are provided in Table 7. The four groups were matched for word frequency according to CELEX, number of letters, number of phonemes, duration, orthographic neighborhood, number of higher frequency neighbors (Grainger, 1990), body neighbors (Ziegler & Perry, 1998) and phonological neighborhood (Goldinger et al., 1989). There were no significant differences in any of the comparisons (see Table 7 for details). In addition, we verified that our groups did not differ on the semantic variables studied by Balota et al. (2004), that is, imageability, meaningfulness and connectivity. As an imageability measure, we used recent norms that were provided for 3,000 monosyllabic words (Cortese & Fugett, 2004). To assess potential differences between our items in terms of meaningfulness, we used semantic connectivity measures proposed by Steyvers and Tenenbaum (2005). This metric indicates how many connections a given word has to other words in the network and how many words are connected to that given word. This measure was obtained for two large-scale databases: WordNet (G. A. Miller, 1990) and the word association norms of Nelson and colleagues (Nelson, McEvoy, & Schreiber, 1998).

Insert Table 7 here

Nonwords (N=80) were selected using the *ARC nonword database* (Rastle, Harrington, & Coltheart, 2002). This database comes with a search function which allows one to select nonwords with certain properties. This made it possible to match our nonwords very tightly to the critical words in terms of number of letters (4.5), number of phonemes (3.7), orthographic neighborhood (5.2) and phonological neighborhood (8.1). All nonwords were monosyllabic, orthographically legal and pronounceable according to the ARC nonword database. No consistency manipulation was performed on the nonwords.

Results

Two participants were excluded from the auditory lexical decision task because of overall high error rates. One item was excluded from the auditory lexical decision data because of an error rate greater than 50% (squaw, 71% errors). Four items were excluded from the visual lexical decision data because of error rates greater than 50% (squaw 91%, caste 65%, casque 65% and salve 60%). In addition, .89% of the auditory lexical decision data and 1.8% of the visual lexical decision data were excluded due to data trimming. The results from both modalities are presented in Table 8.

Insert Table 8 here

Global Analyses

The data were submitted to a $2 \times 2 \times 2$ ANOVA with Modality (visual versus auditory), FF Consistency (consistent versus inconsistent) and FB Consistency (consistent versus inconsistent) as factors. We predicted that FB consistency effects should be present in the auditory modality but not the visual modality, whereas FF consistency effects should be present in the visual modality but not the auditory modality. Thus, we expected interactions of FF and FB consistency effects with Modality.

The latency data showed a significant effect of Modality, $F(1, 59) = 79.51, p < .0001$; $F(1, 72) = 616.28, p < .0001$, reflecting the fact that RTs were faster in the visual than the auditory modality. More importantly, the RT data exhibited significant effects of FB consistency, $F(1, 59) = 33.71, p < .0001$; $F(1, 72) = 3.50, p < .065$, that were further qualified by the predicted FB Consistency \times Modality interaction, $F(1, 59) = 12.16, p < .001$;

$F2(1, 72) = 3.35, p < .07$. In contrast, no significant effect was obtained for FF consistency, $F1(1, 59) = .51, p > .40$; $F2(1, 72) = .12, p > .70$, and the interaction with modality was not significant, $F1(1, 59) = 1.13, p > .30$; $F2(1, 72) = .00, p > .99$. In the analysis by participants, FF Consistency interacted with FB consistency, $F1(1, 59) = 5.25, p < .05$; $F2(1, 72) = 1.56, p > .20$. The triple interaction failed to reach significance, all $F_s < 1$.

In the error data, the main effect of modality was again significant, $F1(1, 59) = 10.57, p < .01$; $F2(1, 72) = 6.28, p < .05$, with more errors in the auditory than the visual modality. FB consistency had no significant effect and did not interact with Modality, all $F_s < 1$. The effect of FF Consistency was significant by participants but not by items, $F1(1, 59) = 5.64, p < .05$; $F2(1, 72) = .32, p > .50$. Similarly, the FF Consistency \times Modality interaction was significant by participants only, $F1(1, 59) = 4.48, p < .05$; $F2(1, 72) = .29, p > .50$. This interaction was due to the fact that FF inconsistent words produced more errors than FF consistent words in the auditory but not in the visual modality. If anything, we predicted the opposite, a stronger FF consistency effect in the visual modality. As in the latency analyses, FF consistency interacted with FB consistency, $F1(1, 59) = 15.92, p < .0001$; $F2(1, 72) = 2.61, p > .10$. Finally, the triple interaction was not significant, $F_s < 1$.

Visual Modality Only

The data were submitted to a 2×2 ANOVA with FF Consistency and FB Consistency as factors. In the latency data, no significant effects were obtained neither for FB consistency, $F1(1, 22) = 2.68, p > .10$; $F2(1, 72) = .42, p > .50$, nor for FF Consistency, $F1(1, 22) = .84, p > .30$; $F2(1, 72) = .11, p > .70$. The interaction between FF Consistency and FB Consistency failed to reach significance, $F1(1, 22) = 2.84, p > .10$; $F2(1, 72) = 2.80, p > .10$.

The error data showed no significant effects of either FF or FB consistency, all $F_s < 1$, $p_s > .60$. In contrast, the $FF \times FB$ interaction was significant, $F(1, 22) = 15.94$, $p < .001$; $F(1, 72) = 4.75$, $p < .05$. This interaction reflects the fact that the predicted FF consistency effect was only obtained for FB consistent items and that a potential FB consistency effect was limited to FF consistent items.

Auditory Modality Only

Turning to the auditory data, a significant effect of FB consistency was obtained, $F(1, 37) = 5.27$, $p < .0001$; $F(1, 75) = 3.99$, $p < .05$. As predicted, there was no hint of a FF consistency effect, $F(1, 37) = .15$, $p > .70$; $F(1, 75) = .04$, $p > .80$. The interaction failed to reach significance, $F(1, 37) = 2.50$, $p > .10$; $F(1, 75) = .75$, $p > .30$.

In the error data, no significant effect of FB consistency was obtained, $F_s < 1$. The effect of FF consistency was significant by participants but not by items, $F(1, 37) = 10.73$, $p < .01$; $F(1, 75) = 1.49$, $p > .20$, so was the interaction between FF and FB consistency, $F(1, 37) = 5.08$, $p < .05$; $F(1, 75) = 1.13$, $p > .20$.

Fast Auditory Versus Slow Visual Responders

As in Experiment 2, one could argue that feedback consistency effects are more stable in the auditory modality because participants respond more slowly in the auditory LDT. For this purpose, we again analyzed the data from the 10 fastest participants in the auditory task against the 10 slowest participants in the visual LDT. These data are presented in Table 9. Indeed, for this subset, the Modality effect disappeared, $F(1, 18) = 1.05$, $p > .30$, which confirms the efficiency of the matching.

Insert Table 9 here

The results were extremely similar to those of the global analysis. That is, on RTs, a significant effect of FB consistency effect was obtained, $F(1, 18) = 8.31, p < .05$, and the effect was qualified by an interaction with modality, $F(1, 18) = 5.03, p < .05$. In contrast, neither the effect of FF consistency nor the interaction between FF consistency and Modality was significant, all $F_s < 1, p_s > .50$. No other interactions were significant, all $F_s < 1$. In the error data, only Modality had a significant effect, $F(1, 18) = 15.25, p < .001$. None of the other effects or interactions were significant, all $p_s > .20$.

Linear Regression Analyses

As in Experiment 2, we performed stepwise regression analyses to determine the amount of unique variance explained by feedback consistency after partialing out word frequency, word length, and neighborhood in step 1 of the regression analysis. As before, in the visual modality, word length was the number of letters and neighborhood was the number of orthographic neighbors. In the auditory modality, word length was the number of phonemes and neighborhood was the number of phonological neighbors. Feedback consistency was entered in step 2 as a continuous variable (i.e. consistency ratios). The results showed that feedback consistency accounted for 6.4% unique variance in the auditory modality ($\Delta F(1, 74) = 5.54, p = .020$), whereas it accounted for a small and nonsignificant amount of variance in the visual modality (0.3%; $\Delta F(1, 71) = .29, p = .59$). Similarly, when frequency, duration and length were entered in step 1, feedback consistency still accounted for 4.4% unique variance in the auditory modality ($\Delta F(1, 74) = 4.94, p = .029$).

Discussion

The main results of the present experiment can be summarized as follows. First, we failed to replicate Stone et al.'s (1997) results because the orthogonal manipulation of feedforward and feedback consistency in the visual lexical decision task showed no significant effect of either variable. The absence of a feedforward consistency effect might seem puzzling. However, before Stone et al. (1997), no previous study reported consistency effects in lexical decision unless the items included strange words (Berent, 1997; Seidenberg et al., 1984). The only significant effect in the visual modality was an interaction between feedforward and feedback consistency on error rate. Note that the same interaction was also present for one of the semantic variables, namely the connectivity measure (see Table 7). It is thus possible that differences in connectivity produced this interaction. Second, with regard to auditory lexical decision, we found a robust feedback consistency effect in the latency data. This finding perfectly replicates the results of Experiment 2 showing that feedback consistency effects are present in the auditory modality but absent in the visual modality. As in the French experiment, the cross-modal dissociation is striking given that the same items were used across the two modalities. Third, as predicted, we did not find strong evidence for second-order feedback consistency effects in the auditory modality, which seems to suggest that second-order feedback effects are as difficult to find in the auditory modality as they are in the visual modality. Finally, feedback consistency accounted for a significant amount of unique variance in the auditory modality, whereas it did not account for a significant amount of unique variance in the visual modality. In sum then, the present experiment, done in English, confirms the existence of a reliable feedback consistency effect in the auditory modality and the absence of feedback consistency effects in the visual modality.

General Discussion

The present study investigated the role of feedback consistency in the visual and auditory lexical decision task in French and English. The main results were straightforward and consistent across the three experiments. 1) There was no evidence for a feedback consistency effect in the visual modality neither in English nor in French, and regardless of whether consistency was manipulated for the onset or the rime (French data only). 2) Robust feedback consistency effects were obtained in auditory lexical decision using the same items that produced a null effect in the visual modality, both in English and in French. 3) Fast as well as slow participants showed feedback consistency effects in the auditory modality but not the visual modality, which rules out an explanation of the cross-modal dissociation in terms of overall response speed.

While the finding of a feedback consistency effect in the auditory modality replicates an increasing number of studies (see references in the introduction), the absence of a feedback consistency effect in the visual modality contradicts the results from Stone et al. (1997) and those of two recent studies, which found feedback consistency effects in the visual modality while controlling for all major word recognition variables including subjective familiarity (i.e., Lacruz & Folk, 2004; Perry, 2003). Before turning to a discussion of our results, these discrepancies need to be addressed.

Turning to Stone et al.'s (1997) study first, there are two problems, which seem to cast doubt on the reliability of their results. The first has to do with objective word frequency measures. Stone et al. (1997) controlled for word frequency using the Kucera & Francis corpus, which is based on about 1 million words. When Peereman et al. (1998) calculated the word frequencies of Stone et al.'s stimuli using the much larger CELEX corpus (16 million

words), they found that Stone et al.'s consistent words were significantly more frequent than their inconsistent words.

The second problem is that Stone et al. (1997), as Ziegler et al. (1997), failed to control for subjective familiarity. This is of concern because word frequency tables tend to be less reliable for low-frequency than for high-frequency words (Gernsbacher, 1984). One could, of course, argue that familiarity ratings are themselves affected by feedback consistency. The argument being that a familiarity rating is a dependent variable, much like latencies in a lexical decision experiment. Thus, controlling for familiarity would make the feedback consistency effect in lexical decision go away. However, this argument has been directly addressed by Peereman et al. (1998) in various analyses. Most importantly, they conducted an experiment (Exp. 5c), for which they selected 160 words that were matched for frequency but that varied in feedback consistency in a graded fashion (10 consistency classes were used). They then asked participants to judge the familiarity of these items. If feedback consistency influenced familiarity ratings, words from lower consistency classes should have been judged less frequent than words from higher consistency classes. The results showed a strong correlation between objective frequency and familiarity ($r = .90$) but no significant correlation between feedback consistency and familiarity. Thus, the failure to control for subjective familiarity needs to be taken seriously.

Thus, we are left with two studies, which seemed to show reliable feedback consistency effects while controlling for subjective familiarity (Lacruz & Folk, 2004; Perry, 2003). One way to test whether their results unequivocally support the existence of feedback consistency effects is to run their stimuli through an implemented word recognition model that does not "know" about feedback consistency. This is the case of the recent CDP+ model (Perry, Ziegler, & Zorzi, 2007). The CDP+ model is a connectionist dual process model; it is not sensitive to feedback consistency at a sublexical level because the mapping between

orthography and phonology is performed by a feedforward associative learning network, the two-layer assembly network of Zorzi, Houghton and Butterworth (1998). It has been shown that the CDP+ model is currently the most powerful reading aloud model because it explains more than twice as much of the item-specific variance in large scale databases than its competitors (i.e., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). Thus, we submitted the feedback inconsistent words of Lacruz and Folk (2004, Exp 2) and Perry (2003) along with their consistent controls to the CDP+ model. Given that the CDP+ is not sensitive to feedback consistency, it should not simulate an effect of feedback consistency. However, if the CDP+ model were to show a feedback consistency effect, we could be fairly confident that a feedback mechanism is not strictly necessary and that variables other than feedback consistency might have produced the feedback consistency effect. The results of the simulations along with the original data are presented in Figure 2.

Insert Figure 2 about here

As can be seen in Figure 2, the CDP+ showed a clear feedback consistency effect for both data sets. As concerns the Lacruz and Folk (2004) data, the CDP+ displayed exactly the same pattern as the human data, that is, a significant feedback consistency effect ($F(1, 28) = 5.12, p < .05$), a significant frequency effect ($F(1, 28) = 47.11, p < .001$), and no significant interaction between feedback consistency and frequency. As concerns the data of Perry (2003), the CDP+ model exhibited a feedback consistency effect although this effect was only marginally significant ($t(48) = 1.5, p < .07$, one-tailed). Thus, in both cases, a model without a sublexical feedback mechanism predicted the correct pattern. Therefore, the feedback consistency effect reported by Lacruz and Folk (2004) and Perry (2003) could result from

variables other than feedback consistency. As a matter of fact, Perry (2003) noticed that his feedback consistency manipulation was confounded with the frequency of spelling-to-sound correspondences. Given that the CDP+ model is extremely sensitive to spelling-to-sound consistency (for details, see Perry et al., 2007), it is likely that the robust feedback consistency effects by Lacruz and Folk (2004) and Perry (2003) were at least partially due to this factor.

It is also puzzling that Balota et al. (2004) reported small but reliable feedback consistency effects in naming in their large-scale database analysis of reading aloud and lexical decision. However, recent simulation work suggests that CDP+ perfectly predicts the feedback consistency effects found in those large-scale databases (M. Yap, *personal communication*, January 22, 2008). So, again, the correlations between feedback consistency and latencies could well be due to other factors than feedback consistency.

The fact that CDP+ can simulate the published feedback consistency effects without using a sublexical feedback mechanism made us wonder whether the model was simply overly sensitive to all sorts of consistency effects. If this were the case, the model would not provide a useful tool for testing whether previously reported feedback consistency effects are really due to feedback consistency. The critical test, therefore, was to run the items of our Experiment 3 through the model. If the model again were to predict a feedback consistency effect despite the fact that our items were controlled much more tightly on all kinds of lexical and sublexical variables, this would strongly weaken the above arguments. Thus, the 80 words of Experiment 3 were run through CDP+. The model made no naming errors. Importantly, there was no hint of a feedback consistency effect in the latency data (106.7 versus 105.1 cycles for feedback consistent versus feedback inconsistent words, respectively, $F < 1$). Thus, the CDP+ model correctly predicts the absence of a feedback consistency effect in Experiment 3. In sum then, the simulation results seem to suggest that the correct pattern to

expect is a null effect of feedback consistency in the visual modality in English as well as in French.

Turning to our results, the pattern that needs to be explained is a null effect of feedback consistency in the visual lexical decision task and a robust effect of feedback consistency in the auditory lexical decision task when using exactly the same items. Three types of explanations can be put forward. The first explanation assumes that the bidirectional coupling between orthography and phonology is still a core mechanism both in visual and auditory word recognition (Frost & Ziegler, 2007; Grainger & Ferrand, 1996; Grainger & Ziegler, 2007; Van Orden & Goldinger, 1994) but that feedback consistency effects are extremely difficult to detect in the visual modality. There are two potential reasons for why feedback consistency effects could be difficult to detect: (i) rapid clean-up and ii) underspecified representations.

According to the first possibility (rapid clean-up), inconsistency in the coupling between phonology and spelling can be “cleaned up” quickly and efficiently because the visual stimulus remains present on the screen. For example, when people are presented with the word BEEF, the resonance framework predicts that rime phonology will activate multiple spelling candidates, such as “EEF” and “EAF”. However, because the correct visual stimulus is still present, one can imagine that wrong spelling candidates are quickly removed from the competition. This is different for the auditory modality. A spoken word is presented sequentially, and the acoustic information cannot be re-accessed once the word has been played. Thus, multiple spellings, if they are activated automatically after an initial pass, cannot be cleaned up and will therefore continue to affect spoken word recognition. One way to test the idea that feedback consistency effects are present but hard to detect would be to use more sensitive on-line techniques, such as event-related brain potentials. Such techniques

might make it possible to detect early and transient effects of feedback inconsistency (Holcomb & Grainger, 2006; Perre & Ziegler, 2008; Sereno & Rayner, 2003).

According to the second possibility (underspecified representations), the computation of phonology in the visual lexical decision task might be underspecified and coarse (e.g., Berent, 1997; Frost, 1998). If the initial computation of phonology is coarse or underspecified (e.g., vowels might not be assembled at first), then feedback consistency effects are difficult to show because these effects require phonology to be fully activated. The situation is different in the auditory modality because phonology does not need to be computed but is fully activated to begin with. Thus, if phonological codes in visual word recognition are underspecified, this would provide an elegant explanation for why feedback consistency effects are obtained in the auditory but not the visual modality. However, if we maintain the idea of a bidirectional coupling, we still need to explain why we did not find second-order (bidirectional) feedback consistency effects in the auditory modality (see Experiment 3).

The second explanation dismisses the idea of a bidirectional coupling. Instead, it is suggested that there is co-activation between orthography and phonology but no reverberation or resonance (feedback loops). That is, phonology would be co-activated whenever we process a visual word (e.g., Ferrand & Grainger, 1994; Rayner, Sereno, Lesch, & Pollatsek, 1995; Van Orden, 1987; Ziegler, Van Orden, & Jacobs, 1997), and orthography would be co-activated whenever we process a spoken word (e.g., Ziegler & Ferrand, 1998; Ziegler et al., 2004). However, the co-activated information would not feed back to influence the initial patterns of activation. That is, there would simply be competition between co-activated units but no additional feedback. As a consequence, one would predict first-order feedback consistency effects but no second-order feedback consistency effects. This prediction is broadly compatible with the present data except for the fact that we did not find feedforward consistency effects in the visual lexical decision task. However, keep in mind that

feedforward consistency effects in the visual lexical decision task are difficult to obtain possibly because the computation of phonology in silent reading is underspecified (e.g., Berent, 1997; Seidenberg et al., 1984).

Finally, one could argue that feedback consistency effects are easier to detect when the incoming information is somewhat ambiguous or noisy (e.g., Frost, Repp, & Katz, 1988). Clearly, incoming “real-world” acoustic information suffers from noise interference (telephones ringing, computer fan noise, jack hammering, background conversation etc.) more than incoming orthographic information. In addition, as mentioned above, orthographic information can be cleaned up more efficiently than acoustic information. Thus, feedback loops or resonance might be particularly useful in the auditory modality to exclude noise. In other words, it could be the case that the influence of both lexical and orthographic feedback is amplified in the auditory modality compared to the visual modality because feedback helps to stabilize the transient acoustic information. Indeed, lexical feedback has been shown to be one of the core features of auditory word recognition and speech perception (Grossberg, Boardman, & Cohen, 1997; McClelland, Mirman, & Holt, 2006; Pitt & Samuel, 1995).

In conclusion, there is very little evidence, neither empirically nor computationally, for feedback consistency effects in the visual modality. In contrast, strong evidence for feedback consistency effects was obtained in the auditory modality. This striking dissociation highlights important differences in the role of feedback in visual and spoken word recognition. Given that orthographic information is absent from current models of spoken word recognition (e.g., Grossberg et al., 1997; McClelland et al., 2006), the existence of orthographic effects on spoken word recognition presents a challenge to these models.

References

- Alloppenna, 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 & Language*, 38, 419-439.
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX Lexical Database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133, 283-316.
- Berent, I. (1997). Phonological priming in the lexical decision task: regularity effects are not necessary evidence for assembly. *Journal of Experimental Psychology: Human Perception & Performance*, 23, 1727-1742.
- Chereau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition*, 102, 341-360.
- Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 1197-1211.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.
- Coltheart, M., Woollams, A., Kinoshita, S., & Perry, C. (1999). A position-sensitive Stroop effect: Further evidence for a left-to-right component in print-to-speech. *Psychonomic Bulletin & Review*, 6, 456-463.

- Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory & Language*, *32*, 193-210.
- Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavioral Research Methods Instruments & Computers*, *36*, 384-387.
- De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: The special status of rimes in English. *Behavior Research Methods, Instruments & Computers*, *34*, 416-423.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.
- Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *47A*, 365-382.
- Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory & Language*, *30*, 1-25.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior, Research Methods, Instruments & Computers*, *35*, 116-124.
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*, 71-99.
- Frost, R., Repp, B. H., & Katz, L. (1988). Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory & Language*, *27*, 741-755.
- Frost, R., & Ziegler, J. C. (2007). Speech and Spelling Interaction: The Interdependence of Visual and Auditory Word Recognition. In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 107-118). Oxford: Oxford University Press.

- Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256-281.
- Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory & Language*, *28*, 501-518.
- Goswami, U., Ziegler, J. C., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*, *92*, 345-365.
- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory & Language*, *29*, 228-244.
- Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory & Language*, *35*, 623-647.
- Grainger, J., & Ziegler, J. C. (2007). Cross-code consistency effects in visual word recognition. In E. L. Grigorenko & A. Naples (Eds.), *Single-word reading: Biological and behavioral perspectives* (pp. 129-157). Mahwah, NJ: Lawrence Erlbaum Associates.
- Grossberg, S., Boardman, I., & Cohen, M. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception & Performance*, *23*, 481-503.
- Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*, 1631-1643.

- Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory & Language*, 46, 723-750.
- Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory & Language*, 29, 687-715.
- Kessler, B., Treiman, R., & Mullennix, J. (2007). Feedback consistency effects in single-word reading. In E. L. Grigorenko & A. Naples (Eds.), *Single-word reading: Behavioral and biological perspectives* (pp. 159–174). Mahwah, NJ: Erlbaum.
- Lacruz, I., & Folk, J. (2004). Feedforward and feedback consistency effects for high- and low-frequency words in lexical decision and naming. *Quarterly Journal of Experimental Psychology A*, 57, 1261-1284.
- Lima, S. D., & Inhoff, A. W. (1985). Lexical access during eye fixations in reading: effects of word-initial letter sequence. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 272-285.
- Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives. ACL-MIT Press series in natural language processing* (pp. 122-147). Cambridge, MA: The MIT Press.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29-63.
- Massaro, D. W., & Jesse, A. (2005). The Magic of Reading: Too Many Influences for Quick and Easy Explanations. In T. Trabasso, J. Sabatini, D. W. Massaro & R. C. Calfee (Eds.), *From orthography to pedagogy: Essays in honor of Richard L. Venezky* (pp. 37-61). Mahwah, NJ: Lawrence Erlbaum Associates.
- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10, 363-369.

- Miller, G. A. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography*, 3, 235-312.
- Miller, K. M., & Swick, D. (2003). Orthography influences the perception of speech in alexic patients. *Journal Cognitive Neuroscience*, 15, 981-990.
- Muneaux, M., & Ziegler, J. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbor generation task. *Language & Cognitive Processes*, 19, 641-660.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. from <http://www.usf.edu/FreeAssociation/>
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L'Année Psychologique*, 101, 447-462.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299-325.
- Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, 22, 1-27.
- Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2008). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience*.
- Peereman, R., & Content, A. (1997). Orthographic and phonological neighborhoods in naming: Not all neighbors are equally influential in orthographic space. *Journal of Memory & Language*, 37, 382-410.
- Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments & Computers*, 31, 376-379.

- Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory & Language*, *39*, 151-174.
- Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*, *1188*, 132-138.
- Perry, C. (2003). A phoneme-grapheme feedback consistency effect. *Psychonomic Bulletin & Review*, *10*, 392-397.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review*, *114*, 273-315.
- Pitt, M. A., & Samuel, A. G. (1995). Lexical and sublexical feedback in auditory word recognition. *Cognitive Psychology*, *29*, 149-188.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56-115.
- Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception & Performance*, *25*, 482-503.
- Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *55A*, 1339-1362.
- Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science*, *6*, 26-32.
- Rey, A., Ziegler, J. C., & Jacobs, A. M. (2000). Graphemes are perceptual reading units. *Cognition*, *75*, B1-B12.

- Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, 23, 383-404.
- Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences*, 7, 489-493.
- Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 56A, 233-262.
- Steyvers, M., & Tenenbaum, J. (2005). The Large Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, 29, 41-78.
- Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 1248-1268.
- Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory & Language*, 36, 337-359.
- 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 & Language*, 58, 366-379.
- Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory & Language*, 26, 608-631.
- Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107-136.

- Tuller, B., Case, P., Ding, M., & Kelso, J. A. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 3-16.
- Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181-198.
- Van Orden, G. C. (2002). Nonlinear dynamics and psycholinguistics. *Ecological Psychology*, *14*, 1-4.
- Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 1269-1291.
- Van Orden, G. C., Jansen op de Haar, M. A., & Bosman, A. M. T. (1997). Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. *Cognitive Neuropsychology*, *14*, 131-165.
- Ventura, P., Morais, J., & Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: from sublexical to lexical involvement. *Cognition*, *105*, 547-576.
- Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, *19*, 57-95.
- Vitevitch, M. S. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, *28*, 270-278.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*, 325-329.

- 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.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, 131, 3-29.
- Ziegler, J. C., Jacobs, A. M., & Klueppel, D. (2001). Pseudohomophone effects in lexical decision: Still a challenge for current word recognition models. *Journal of Experimental Psychology: Human Perception & Performance*, 27, 547-559.
- Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers*, 28, 504-515.
- Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory & Language*, 37, 533-554.
- Ziegler, J. C., & Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: a developmental study. *Psychonomic Bulletin & Review*, 14, 75-80.
- Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, 48, 779-793.
- Ziegler, J. C., & Perry, C. (1998). No more problems in Coltheart's neighborhood: Resolving neighborhood conflicts in the lexical decision task. *Cognition*, 68, B53-B62.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments & Computers*, 29, 600-618.

Ziegler, J. C., Van Orden, G. C., & Jacobs, A. M. (1997). Phonology can help or hurt the perception of print. *Journal of Experimental Psychology: Human Perception & Performance*, 23, 845-860.

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

Appendix A: Items, latencies (RTs) and errors (%) in Experiment 1

	Onset	Rime	Item	Errors	RTs		Onset	Rime	Item	Errors	RTs
1	Con	con	pointe	0.0	588	51	Inc	con	tranche	0.0	621
2	Con	con	figue	8.3	662	52	Inc	con	capre	30.6	789
3	Con	con	lustre	5.6	656	53	Inc	con	hanche	2.8	631
4	Con	con	lune	0.0	576	54	Inc	con	cèdre	5.6	658
5	con	con	prune	2.8	552	55	Inc	con	harpe	5.6	657
6	con	con	porche	0.0	612	56	Inc	con	cycle	0.0	566
7	con	con	cuve	5.6	662	57	Inc	con	tombe	2.8	595
8	con	con	juge	5.6	652	58	Inc	con	herbe	2.8	597
9	con	con	porte	0.0	562	59	Inc	con	hymne	2.8	764
10	con	con	grille	0.0	573	60	Inc	con	rhume	0.0	582
11	con	con	gorge	0.0	625	61	Inc	con	halte	2.8	709
12	con	con	fraude	11.1	708	62	Inc	con	daube	-	-
13	con	con	prince	0.0	617	63	Inc	con	singe	2.8	573
14	con	con	meute	-	-	64	Inc	con	toast	16.7	618
15	con	con	fourche	0.0	676	65	Inc	con	lymphe	8.3	702
16	con	con	coffre	0.0	615	66	Inc	con	bombe	0.0	623
17	con	con	feutre	0.0	666	67	Inc	con	bègue	-	-
18	con	con	norme	5.6	618	68	Inc	con	pompe	0.0	601
19	con	con	frange	11.1	719	69	Inc	con	jambe	0.0	550
20	con	con	louve	11.1	743	70	Inc	con	lèvre	2.8	584
21	con	con	trouble	0.0	581	71	Inc	con	chèvre	0.0	542
22	con	con	spectre	0.0	650	72	Inc	con	cinq	8.3	605
23	con	con	loge	8.3	631	73	Inc	con	timbre	0.0	598
24	con	con	pulpe	5.6	646	74	Inc	con	zeste	13.9	813
25	con	con	fiche	0.0	642	75	Inc	con	chef	0.0	564
26	con	inc	cuir	2.8	564	76	Inc	inc	graine	0.0	643
27	con	inc	gland	11.1	672	77	Inc	inc	soeur	0.0	548
28	con	inc	spot	8.3	803	78	Inc	inc	cèpe	25.0	748
29	con	inc	noce	2.8	662	79	Inc	inc	brie	13.9	722
30	con	inc	gang	19.4	690	80	Inc	inc	thème	2.8	562
31	con	inc	front	2.8	635	81	Inc	inc	seigle	-	-
32	con	inc	bord	2.8	629	82	inc	inc	dôme	19.4	767
33	con	inc	gag	27.8	733	83	inc	inc	moelle	2.8	651
34	con	inc	grippe	0.0	616	84	inc	inc	bûche	2.8	602
35	con	inc	feuille	0.0	579	85	inc	inc	guêpe	0.0	592
36	con	inc	flamme	0.0	562	86	inc	inc	poêle	0.0	589
37	con	inc	stock	2.8	552	87	inc	inc	thèse	0.0	581
38	con	inc	sphère	0.0	630	88	inc	inc	chlore	0.0	648
39	con	inc	chute	2.8	596	89	inc	inc	grêle	2.8	692
40	con	inc	griffe	2.8	573	90	inc	inc	câble	2.8	633
41	con	inc	date	5.6	569	91	inc	inc	kyste	-	-
42	con	inc	jarre	11.1	773	92	inc	inc	greffe	2.8	684
43	con	inc	nord	0.0	585	93	inc	inc	mœurs	2.8	688
44	con	inc	zone	5.6	586	94	inc	inc	phoque	0.0	640
45	con	inc	juin	0.0	618	95	inc	inc	style	2.8	552
46	con	inc	dièse	22.2	757	96	inc	inc	plaie	5.6	586
47	con	inc	sport	0.0	540	97	inc	inc	flair	30.6	728
48	con	inc	science	2.8	611	98	inc	inc	graisse	2.8	623
49	con	Inc	square	5.6	663	99	inc	inc	tête	0.0	558
50	con	Inc	score	2.8	586	100	inc	inc	clerc	30.6	755

Appendix B: Items, latencies (RTs) and errors (%) in Experiment 2

Onset	Rime	Item	Auditory		Visual		Onset	Rime	Item	Auditory		Visual	
			RT	Errors	RT	Errors				RT	Errors	RT	Errors
con	con	foudre	823	1.8	539	9.1	inc	con	bègue	1074	28.6	-	-
con	con	couche	722	0.0	566	0.0	inc	con	bombe	853	0.0	529	4.5
con	con	cuve	858	1.8	520	9.1	inc	con	capre	989	30.4	579	4.5
con	con	feutre	835	1.8	570	0.0	inc	con	cèdre	971	14.0	513	0.0
con	con	fiche	927	17.9	619	18.2	inc	con	chef	904	8.8	597	31.8
con	con	figue	837	1.8	586	18.2	inc	con	chèvre	946	0.0	490	0.0
con	con	fourche	810	0.0	576	4.5	inc	con	cinq	998	1.8	527	4.5
con	con	frange	1034	7.1	677	0.0	inc	con	cycle	989	8.9	565	9.1
con	con	fraude	921	3.5	632	22.7	inc	con	daube	1024	24.5	-	-
con	con	gorge	909	1.8	540	0.0	inc	con	halte	938	19.3	658	9.1
con	con	grille	967	7.3	540	0.0	inc	con	hanche	980	8.8	591	4.5
con	con	jugé	893	0.0	543	4.5	inc	con	harpe	803	7.0	563	4.5
con	con	loge	950	12.3	616	4.5	inc	con	herbe	801	1.8	562	0.0
con	con	louve	983	16.1	696	13.6	inc	con	jambe	880	0.0	491	0.0
con	con	lune	805	0.0	530	0.0	inc	con	lèvre	935	1.8	570	9.1
con	con	lustre	965	8.8	584	9.1	inc	con	lymphe	1020	25.5	640	0.0
con	con	meute	960	19.3	785	13.6	inc	con	pompe	811	5.3	533	4.5
con	con	norme	915	0.0	568	4.5	inc	con	rhume	844	3.6	576	0.0
con	con	coude	730	1.8	580	4.5	inc	con	bref	997	15.8	546	9.1
con	con	porche	743	1.8	591	4.5	inc	con	singe	894	0.0	555	0.0
con	con	porte	687	0.0	521	0.0	inc	con	timbre	766	0.0	523	0.0
con	con	poste	765	3.5	514	0.0	inc	con	tombe	796	1.8	621	9.1
con	con	prune	718	1.8	551	4.5	inc	con	tranche	944	12.5	541	0.0
con	con	trouble	777	0.0	562	4.5	inc	con	zeste	1002	17.5	649	31.8
con	inc	cuir	779	0.0	509	0.0	inc	inc	bûche	866	3.5	558	0.0
con	inc	gland	738	7.1	598	9.1	inc	inc	câble	857	3.5	514	4.5
con	inc	spot	998	1.8	633	9.1	inc	inc	chlore	884	3.5	588	9.1
con	inc	noce	892	14.0	646	4.5	inc	inc	crâne	827	1.8	510	4.5
con	inc	gang	1056	28.6	653	13.6	inc	inc	flair	1062	15.8	635	22.7
con	inc	front	834	5.3	550	0.0	inc	inc	graine	939	1.8	560	0.0
con	inc	scout	1029	17.5	582	22.7	inc	inc	treize	888	0.0	545	0.0
con	inc	gag	943	8.9	669	31.8	inc	inc	greffe	965	17.9	563	0.0
con	inc	grippe	839	3.5	529	4.5	inc	inc	grêle	834	7.0	599	13.6
con	inc	feuille	770	0.0	572	0.0	inc	inc	guêpe	798	7.3	516	0.0
con	inc	flamme	878	0.0	542	0.0	inc	inc	kyste	959	35.1	-	-
con	inc	stock	911	1.8	543	0.0	inc	inc	moelle	937	8.8	580	0.0
con	inc	sphère	920	0.0	552	4.5	inc	inc	moeurs	902	7.1	610	4.5
con	inc	chute	857	0.0	512	0.0	inc	inc	phoque	841	8.8	621	0.0
con	inc	griffe	883	5.4	525	4.5	inc	inc	plaie	956	43.9	571	0.0
con	inc	date	1023	33.3	536	0.0	inc	inc	poêle	810	5.5	556	0.0
con	inc	nord	934	15.8	563	9.1	inc	inc	seigle	1051	16.4	677	18.2
con	inc	zone	1004	7.0	525	9.1	inc	inc	soeur	940	5.5	514	0.0
con	inc	juin	908	1.8	542	9.1	inc	inc	style	993	5.4	513	0.0
con	inc	dièse	969	0.0	707	22.7	inc	inc	quatre	864	3.5	543	0.0
con	inc	sport	870	1.8	504	4.5	inc	inc	thèse	837	12.5	577	0.0
con	inc	science	1051	1.8	548	0.0	inc	inc	hall	1073	42.6	578	4.5
con	inc	square	1017	3.5	605	0.0	inc	inc	trône	886	7.0	540	0.0
con	inc	score	975	14.0	536	0.0	inc	inc	mètre	953	0.0	547	4.5

Appendix C: Items, latencies (RTs) and errors (%) in Experiment 3

FF	FB	Item	Auditory		Visual		FF	FB	Item	Auditory		Visual	
			RT	Errors	RT	Errors				RT	Errors	RT	Errors
inc	con	pint	821	0.0	648	0.0	inc	inc	swamp	948	2.6	664	13.0
inc	con	aunt	999	15.8	727	8.7	inc	inc	beard	913	0.0	594	0.0
inc	con	wasp	896	0.0	633	8.7	inc	inc	gross	1050	18.4	723	4.3
inc	con	puss	916	36.8	789	17.4	inc	inc	swan	924	5.3	643	4.3
inc	con	math	864	13.2	729	4.3	inc	inc	foul	945	10.5	767	0.0
inc	con	gloves	945	0.0	654	0.0	inc	inc	sweat	916	0.0	611	0.0
inc	con	gown	1125	31.6	665	13.0	inc	inc	vase	894	2.6	702	0.0
inc	con	shove	911	5.3	704	4.3	inc	inc	scone	995	7.9	685	4.3
inc	con	golf	864	2.6	642	0.0	inc	inc	bomb	893	31.6	563	0.0
inc	con	salve	1092	36.8	-	-	inc	inc	swear	1002	0.0	664	0.0
inc	con	wolf	798	2.6	689	4.3	inc	inc	warp	1044	31.6	827	21.7
inc	con	gouge	1039	18.4	872	43.5	inc	inc	deaf	754	2.6	612	0.0
inc	con	casque	1007	39.5	-	-	inc	inc	frost	934	7.9	713	0.0
inc	con	moth	834	2.6	659	4.3	inc	inc	monk	852	2.6	613	0.0
inc	con	bull	797	2.6	601	0.0	inc	inc	wool	856	10.5	668	8.7
inc	con	bush	790	2.6	621	0.0	inc	inc	squash	950	2.6	623	0.0
inc	con	dose	989	39.5	735	4.3	inc	inc	cough	779	5.3	634	0.0
inc	con	dove	728	0.0	579	4.3	inc	inc	squad	1037	39.5	731	4.3
inc	con	soot	952	15.8	775	47.8	inc	inc	caste	990	7.9	-	-
inc	con	wounds	918	5.3	661	0.0	inc	inc	clerk	1050	23.7	732	13.0
con	inc	mute	930	5.3	664	4.3	con	con	clash	837	5.3	613	4.3
con	inc	lamb	916	7.9	615	0.0	con	con	shave	929	7.9	607	0.0
con	inc	grade	974	5.3	662	0.0	con	con	crab	893	2.6	683	0.0
con	inc	crowd	919	7.9	664	0.0	con	con	coins	895	13.2	633	4.3
con	inc	bald	943	13.2	657	4.3	con	con	loaf	870	2.6	675	4.3
con	inc	squaw	-	-	-	-	con	con	blink	932	2.6	651	4.3
con	inc	cox	1028	18.4	784	8.7	con	con	dive	938	5.3	677	4.3
con	inc	freak	1034	15.8	645	0.0	con	con	junk	820	5.3	730	4.3
con	inc	smear	977	0.0	710	0.0	con	con	crisp	905	0.0	595	8.7
con	inc	tact	918	5.3	806	26.1	con	con	smart	1039	2.6	634	0.0
con	inc	dune	987	5.3	858	39.1	con	con	plug	797	0.0	659	0.0
con	inc	soap	868	2.6	610	4.3	con	con	snob	1010	39.5	683	4.3
con	inc	mule	961	10.5	686	4.3	con	con	silk	815	0.0	616	0.0
con	inc	spur	1136	31.6	712	13.0	con	con	tape	858	2.6	602	0.0
con	inc	myth	821	0.0	664	4.3	con	con	tusk	866	15.8	792	26.1
con	inc	muse	858	10.5	704	8.7	con	con	malt	1071	26.3	714	4.3
con	inc	plea	915	13.2	755	4.3	con	con	dish	833	0.0	641	4.3
con	inc	truce	993	21.1	853	30.4	con	con	clip	791	2.6	660	4.3
con	inc	blur	989	15.8	642	4.3	con	con	thrill	869	5.3	740	0.0
con	inc	priest	866	5.3	669	4.3	con	con	vague	902	13.2	736	0.0

Acknowledgments

Johannes Ziegler, CNRS and Université de Provence, Marseille (France). Ana Petrova, Université René Descartes, Paris (France). Ludovic Ferrand, CNRS and Université Blaise Pascal, Clermont-Ferrand (France).

This work was supported by an ANR grant n°JC05_44765 to J. Ziegler (Agence Nationale de la Recherche, France).

We wish to thank Gareth Gaskell for hosting us at the University of York while testing the English part of this study and Melvin Yap for providing the semantic measures of Exp. 3. We thank numerous reviewers for providing critical and helpful feedback on various versions of this manuscript. Apologies to my dear friends Guy Van Orden and Greg Stone if they somehow feel betrayed by this work.

Correspondence concerning this article should be addressed to Johannes Ziegler, Laboratoire de Psychologie Cognitive, CNRS UMR 6146, Université de Provence, 3 place Victor Hugo, 13331 Marseille, France. E-mail: Johannes.Ziegler@univ-provence.fr

Tables

Table 1

Stimulus characteristics of Experiment 1 manipulating feedback consistency of the onset and the rime. All statistics are corrected for excluded items.

	ON con		ON inc		ANOVA (F1, 90)		
	RI con	RI inc	RI con	RI inc	ON	RI	ON x RI
Manipulated Variables							
FB Consistency Onset	1.00	1.00	0.18	0.18	3771.00**	0.00	0.03
FB Consistency Rime	0.99	0.17	0.97	0.19	0.01	3441.17**	2.68
Controlled Variables							
Nb Letters	5.46	4.92	5.20	5.20	0.00	2.17	2.30
Frequency (Lexique)	35.19	31.39	29.66	32.60	0.02	0.00	0.05
Frequency (Lexop)	42.67	44.32	42.87	42.86	0.00	0.00	0.00
Orthographic N	4.75	3.64	3.26	3.14	2.95	1.13	0.72
FF Consistency	0.93	0.89	0.94	0.92	0.37	1.04	0.04
Familiarity Ratings	6.05	6.14	5.92	5.68	3.57	0.25	1.10

Notes. ON, onset; RI, rime; con, consistent; inc, inconsistent; FF, feeforward; FB, feedback; *

p< .05; **, p<.001; N, Neighborhood

Table 2. Mean latencies and errors for an orthogonal manipulation of feedback consistency at the onset and the rime position in the visual lexical decision task of Exp 1. Standard errors in parentheses.

ONSET	RIME		avg
	CON	INC	
Latency			
CON	634 (17.1)	628 (15.3)	631
INC	629 (15.6)	639 (17.2)	634
Avg	631	634	
Errors			
CON	3.4 (0.8)	5.7 (1.2)	4.5
INC	4.7 (0.8)	6.5 (0.9)	5.6
Avg	4.0	6.1	

Notes. CON, consistent; INC, inconsistent, avg, average

Table 3. Results from a post-hoc analysis of Exp. 1
 using only items with a frequency < 20 per million
 (16 items per cell).

ONSET	RIME		avg
	CON	INC	
	Latency		
CON	654	661	657
INC	655	670	663
Avg	654	665	
	Errors		
CON	4.7	8.3	6.5
INC	6.1	8.5	7.3
Avg	5.4	8.4	

Table 4

Stimulus characteristics of Experiment 2 manipulating feedback consistency of the onset and the rime in the visual and the auditory modality.

	ON con		ON inc		ANOVA (F1, 92)		
	RI con	RI inc	RI con	RI inc	ON	RI	ON x RI
Manipulated Variables							
FB Consistency Onset	0.99	1.00	0.18	0.17	4010**	0.01	0.16
FB Consistency Rime	0.99	0.18	0.96	0.19	0.27	3509**	1.68
Controlled Variables							
NB Letters	5.33	4.96	5.13	5.33	0.25	0.25	3.12
NB Phonemes	4.08	3.58	3.88	3.88	0.06	2.06	2.06
Frequency (Lexique)	34.53	27.68	30.47	27.65	0.03	0.15	0.03
Frequency (Lexop)	41.29	40.21	43.67	32.88	0.03	0.17	0.12
Orthographic N	5.96	3.29	3.33	2.46	10.77*	11.29*	2.89
Phonological N	10.00	10.54	9.04	12.50	0.20	3.18	1.69
Uniqueness point	4.08	3.58	3.75	3.83	0.06	1.51	2.96
FF Consistency	0.96	0.92	0.97	0.97	3.49	1.13	1.37
Familiarity Ratings	6.05	6.19	5.81	5.91	2.66	0.61	0.02
Duration	618.00	647.00	652.00	637.00	0.43	0.18	1.51

Notes. ON, onset; RI, rime; con, consistent; inc, inconsistent; FF, feeforward; FB, feedback;

N, Neighborhood; * $p < .05$; **, $p < .001$.

Table 5. Mean reaction times and errors in the visual and the auditory lexical decision task of Experiment 2 using the same items.

ONSET	VISUAL			AUDITORY		
	RIME			RIME		
	CON	INC	avg	CON	INC	avg
Latencies						
CON	585	571	578	852	916	884
INC	567	574	571	915	907	911
avg	576	573	-	883	911	-
Errors						
CON	5.0	6.3	5.6	4.6	7.2	5.9
INC	5.7	3.7	4.7	9.8	10.9	10.4
avg	5.4	5.0	-	7.2	9.1	-

Table 6. Comparison between the 10 fastest participants in the auditory lexical decision task of Exp 2 and the 10 slowest participants in the visual lexical decision task of Exp 1 and Exp 2 collapsed.

ONSET	VISUAL			AUDITORY		
	RIME			RIME		
	CON	INC	Avg	CON	INC	Avg
Latencies						
CON	773	752	763	740	810	775
INC	761	764	762	799	800	799
avg	767	758	-	769	805	-
Errors						
CON	2.9	8.1	5.5	4.6	9.6	7.1
INC	7.5	8.3	7.9	10.5	15.0	12.8
avg	5.2	8.2	-	7.5	12.3	-

Table 7

Stimulus characteristics of Experiment 3 manipulating feedforward and feedback consistency in the visual and the auditory modality.

	FF con		FF inc		F(1,75)		
	FB con	FB inc	FB con	FB inc	FF	FB	FF*FB
Manipulated Variables							
FF consistency ratio	0.95	0.95	0.30	0.21	761.86**	2.83	3.59
FB consistency ratio	0.94	0.23	0.89	0.19	2.78	687.99**	0.05
Controlled Variables							
Frequency (CELEX)	11.57	12.45	12.38	14.02	0.19	0.21	0.02
NB letters	4.45	4.32	4.45	4.65	1.27	0.05	1.27
NB phonemes	3.80	3.68	3.50	3.75	0.60	0.20	1.45
Duration	558.44	540.40	534.94	562.07	0.00	0.07	1.64
ON	5.65	5.42	7.60	4.50	0.36	3.81	2.83
HFN	2.45	2.32	3.45	2.40	0.98	1.17	0.70
BN	3.80	3.68	3.50	3.75	0.60	0.20	1.45
PN	8.20	10.21	7.35	10.20	0.11	3.65	0.11
C & F imageability	5.03	4.42	5.32	4.94	1.76	2.68	0.15
WordNet connectivity	1.67	1.25	1.25	1.54	0.13	0.14	4.22*
Word association strengths	3.11	2.86	2.86	3.00	0.09	0.10	1.31

Notes. FF = feedforward; FB = feedback; con = consistent; inc = inconsistent; ON = number of orthographic neighbors; HFN = number of higher frequency neighbors; BN = number of body neighbors; PN = number of phonological neighbors; C&F imageability = Cortese & Fugett (2004) imageability measure; *, $p < .05$; **, $p < .001$.

Table 8. Mean reaction times and errors for feedforward and feedback consistent and inconsistent words in the visual and the auditory lexical decision task of Experiment 3.

FF	VISUAL			AUDITORY		
	FB		avg	FB		avg
	CON	INC		CON	INC	
Latencies						
CON	667	693	680	887	943	915
INC	673	669	671	900	934	917
avg	670	681	-	894	939	-
Errors						
CON	3.9	8.0	6.0	7.6	9.7	8.7
INC	8.7	3.7	6.2	14.3	10.7	12.5
avg	6.3	5.9	-	10.9	10.2	-

Table 9. Comparison between the 10 fastest participants in the auditory lexical decision task and the 10 slowest participants in the visual lexical decision task.

FF	VISUAL			AUDITORY		
	FB		Avg	FB		Avg
	CON	INC		CON	INC	
Latencies						
CON	775	778	776	793	833	813
INC	777	784	780	802	841	822
avg	776	781	-	797	837	-
Errors						
CON	3	4	3.5	11	14.5	12.8
INC	4.7	2.5	3.6	13.7	11.5	12.6
avg	3.9	3.3	-	12.3	13	-

Figure Captions

Figure 1. Schematic description of the resonance model (adapted from Van Orden & Goldinger, 1994).

Figure 2. Simulations of the feedback consistency effect reported by Lacruz and Folk (2004) and Perry (2003) using the connectionist dual process model (CDP+, Perry et al. 2007).

Figure 1

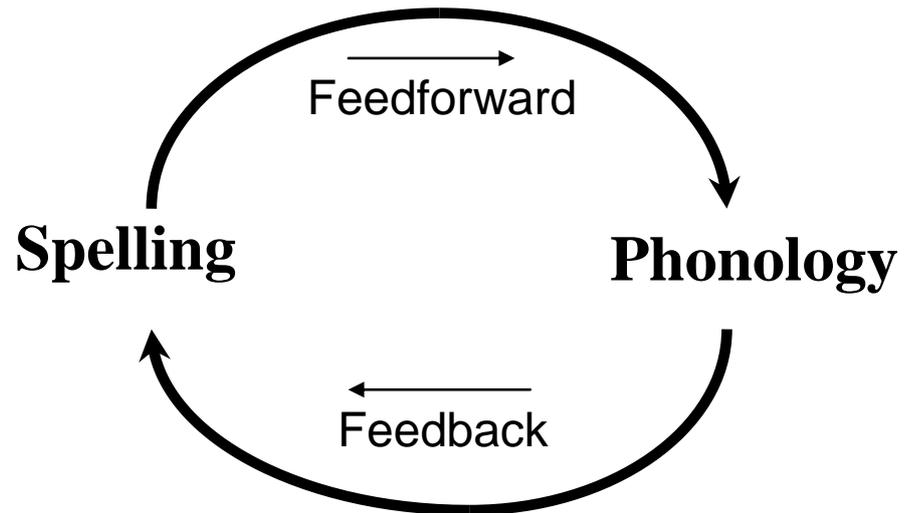


Figure 2

