Phoneme-Order Encoding During Spoken Word Recognition: A Priming Investigation

Sophie Dufour, Jonathan Grainger

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12785

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12785

Phoneme-Order Encoding During Spoken Word Recognition: A Priming Investigation

Sophie Dufour,[a,b] Jonathan Grainger[b,c]

[a]*Aix-Marseille Université, CNRS, LPL, UMR 7309*
[b]*Institute for Language, Communication, and the Brain, Aix-Marseille Université*
[c]*Laboratoire de Psychologie Cognitive, Aix-Marseille Université & CNRS*

Received 24 January 2019; received in revised form 11 June 2019; accepted 26 July 2019

Abstract

In three experiments, we examined priming effects where primes were formed by transposing the first and last phoneme of tri-phonemic target words (e.g., /byt/ as a prime for /tyb/). Auditory lexical decisions were found not to be sensitive to this transposed-phoneme priming manipulation in long-term priming (Experiment 1), with primes and targets presented in two separated blocks of stimuli and with unrelated primes used as control condition (/mul/-/tyb/), while a long-term repetition priming effect was observed (/tyb/-/tyb/). However, a clear transposed-phoneme priming effect was found in two short-term priming experiments (Experiments 2 and 3), with primes and targets presented in close temporal succession. The transposed-phoneme priming effect was found when unrelated prime-target pairs (/mul/-/tyb/) were used as control and more important when prime-target pairs sharing the medial vowel (/pys/-/tyb/) served as control condition, thus indicating that the effect is not due to vocalic overlap. Finally, in Experiment 3, a transposed-phoneme priming effect was found when primes sharing the medial vowel plus one consonant in an incorrect position with the targets (/byl/-/tyb/) served as control condition, and this condition did not differ significantly from the vowel-only condition. Altogether, these results provide further evidence for a role for position-independent phonemes in spoken word recognition, such that a phoneme at a given position in a word also provides evidence for the presence of words that contain that phoneme at a different position.

Keywords: Spoken word recognition; Phoneme order; Priming

1. Introduction

Most models of spoken word recognition (Gaskell & Marslen-Wilson, 1997; Marslen-Wilson, 1990; Marslen-Wilson & Warren, 1994; Marslen-Wilson & Welsh, 1978;

Correspondence should be sent to Sophie Dufour, Laboratoire Parole et Langage, Aix-Marseille Université, CNRS, 5, avenue Pasteur, 13604 Aix-en-Provence, France. E-mail: sophie.dufour@lpl-aix.fr

McClelland & Elman, 1986; Norris, 1994) assume that listeners first conduct an acoustic-phonetic analysis of the speech signal, and then map the output of this analysis onto the phonological representations of words stored in the mental lexicon. Some of these models assume that words are represented as strings of phonemes (Marslen-Wilson & Welsh, 1978; McClelland & Elman, 1986; Norris, 1994). Consequently, in these models the phonemes are extracted from the speech signal and are used to make contact with whole-word phonological representations. Alternative accounts of spoken word recognition (Gaskell & Marslen-Wilson, 1997; Grossberg, 2003; Lahiri & Marslen-Wilson, 1991; Marslen-Wilson & Warren, 1994) assume that words are specified in terms of acoustic features. Consequently, in these models, the featural information extracted from the speech signal is mapped directly onto lexical representations, without the involvement of an intermediate phonemic level of representations.

Although the above-mentioned models differ in terms of how the speech signal makes contact with lexical representations, they all share the key assumption that the phonological form of words is an ordered sequence of information. As a result, within such a view the information extracted from the speech signal is coded according to their position in the speech input in order to be successfully mapped onto an ordered sequence of information. In line with this theorizing, there is plenty of evidence that words that share the same phoneme at the same position in a given target word influence target word processing (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Dufour & Peereman, 2003; Marslen-Wilson, Moss, & van Halen, 1996; Zwitserlood, 1989). There is, however, some evidence in favor of a more flexible phoneme-order encoding from studies showing that words that can be generated by adding or deleting a phoneme in the target word also influence target word recognition (Dufour & Frauenfelder, 2010; Luce & Pisoni, 1998; Vitevitch & Luce, 1999; see also Allopenna et al., 1998; Connine, Blasko, & Titone, 1993). Evidence for such flexibility is in line with more recent accounts of spoken word recognition, such as the TISK model of Hannagan, Magnuson, and Grainger (2013; see You & Magnuson, 2018, for a more recent implementation). Such flexibility is achieved in TISK via open diphone units that represent ordered sequences of contiguous and non-contiguous phonemes.

Furthermore, a study by Toscano, Anderson, and McMurray (2013) points to a more extreme form of flexibility that was anticipated in TISK via the implementation of a set of position-independent phoneme units (Hannagan et al., 2013, p. 5). Using the visual world paradigm, Toscano et al. (2013) examined the eye movements of participants who followed spoken instructions to manipulate objects pictured on a computer screen. They found more fixations on the picture representing a CAT than on a control picture (e.g., the picture of a MILL) when the spoken target was TACK /tak/, thus suggesting that CAT and TACK are confusable words. Importantly, the activation of CAT when the target is TACK cannot be attributed to shared features between the phonemes /t/ and /k/, since exactly the same pattern of results was found for pairs of words such as BUS and SUB whose phonemes /s/ and /b/ are maximally distinct. All together, these findings suggest that position-independent phonemes play a role in spoken word recognition, and thus constitute a major challenge for the dominant accounts of spoken word recognition that code for the precise order of segments, while supporting the predictions of the TISK model (Hannagan et al., 2013).

In this study, we provide a further test of a role for position-independent phonemes in spoken word recognition by using a priming paradigm. This is an important endeavor because, to the best of our knowledge, they are only two studies to date that have examined this important question (Gregg, Inhoff, & Connine, in press; Toscano et al., 2013), and also because Toscano et al.'s findings were recently questioned regarding their statistical reliability (Mitterer, Reinisch, & McQueen, 2018). It is also important because, to our knowledge, there is only one model of spoken word recognition, the TISK model (Hannagan et al., 2013), that can account for Toscano et al.'s findings. Furthermore, this study provides a test of the time-course of the influence of position-independent phonemes by comparing effects obtained with a long-term priming procedure (Experiment 1) and a short-term priming procedure (Experiment 2).

2. Experiment 1

Experiment 1 used the long-term priming paradigm in which auditory primes and targets were presented in separated blocks of stimuli. The logic behind this paradigm is that the presentation of primes in a first block of stimuli should influence processing of related target words presented in a second block of stimuli. Although the long-term priming is not as widely used as the classic short-term priming paradigm in which primes and targets are presented in close temporal succession, it is a well-established procedure to probe processes involved in spoken word recognition, and despite the long lag, robust priming effects have been reported. For example, word repetition effects have been obtained with this paradigm (Dufour & Nguyen, 2014; McLennan & Luce, 2005; Sumner & Samuel, 2009), as well as more subtle priming effects obtained across phonologically similar words (Dufour & Nguyen, 2017; Monsell & Hirsh, 1998; Sumner & Samuel, 2009). In the present experiment, prime and target words shared all phonemes in the same order (/tyb/-/tyb/; repetition condition), all phonemes in a different order (/byt/-/tyb/; transposed-phoneme condition), or were unrelated (/mul/-/tyb/; control condition). The prediction was straightforward. If, as suggested by the results of Toscano et al. (2013), a phoneme at a given position can provide evidence for the presence of words that contain that phoneme at a different position, then the target word /tyb/ that shares all of its phonemes with the prime word /byt/ should receive support during prime processing, and thus facilitate its subsequent processing as a target, in comparison to the control condition.

2.1. Method

2.1.1. Participants

Forty-five French speakers from Aix-Marseille University participated in the experiment. All participants reported having no hearing or speech disorders.

2.1.2. Materials

Forty-five monosyllabic target words with a CVC syllabic structure were selected from Vocolex, a lexical database for French (Dufour, Peereman, Pallier, & Radeau, 2002). All target words had their uniqueness point—the phonemic position at which a word can be reliably identified—after their last phoneme. The 45 target words also served as repeated primes. For each target word, a CVC word was created by transposing the two consonants (LOBE /lɔb/ "*lobe*" – BOL /bɔl/ "*bowl*") and served as transposed-phoneme primes. Fifteen other CVC words that shared no phonemes with the target words served as control primes. For the purpose of the lexical decision task, 75 CVC non-words serving as both primes and targets were created by changing the last phoneme of words not used in the experiment (e.g., the word bague /bag/ "ring" became /ban/). This allowed us to have wordlike non-words, and to force participants to listen to the stimuli up to the end prior to giving their response. So that the non-words followed the same criteria as the words, 30 of them consisted of pairs sharing the same phonemes but in different order (e.g., /nab/-/ban/). Note that the non-word prime-target pairs were not further analyzed, and only the critical 45 target words preceded by their respective repeated, transposed, and control prime words were submitted to statistical analyses. All of the stimuli were recorded by a female native speaker of French, in a sound attenuated room, and digitized at a sampling rate of 44 kHz with 16-bit analog to digital recording. The main characteristics of the prime and the target words are given in Table 1. The prime and target words are given in Appendix A.

2.1.3. Design

For the purpose of the long-lag priming procedure, the targets and their respective primes were dispatched in two blocks of stimuli, the prime block and the target block. The target block consisted of the 45 target words along with 45 non-words. Among the 45 target words 15 were used in the repetition condition, 15 in the transposed-phoneme condition and 15 in the control condition. The prime block was constructed from the target block so that each target was preceded by their respective primes. The prime block thus consisted of 45 prime words. Among the 45 prime words, 15 consisted of the target words (i.e., the repetition primes), 15 consisted in the transposed-phoneme primes, and 15 were the unrelated primes. The prime block also included 45 non-words. Among these 45 non-words, 15 were also used in the target block, 15 were the transposed-phoneme version of non-words used in the target block, and 15 were unrelated to any of the non-words used in the target block. A Latin-square design was used such that the critical 45 target words were preceded by the three types of prime (repeated, transposed, control) across different participants, and participants were tested in all three conditions but were presented with each target word only once. This required creating three experimental lists each composed of two blocks.

2.1.4. Procedure

The participants were tested in a sound-attenuated booth, and stimuli were presented over headphones at a comfortable sound level. Stimulus presentation and recording of the

Table 1
Characteristics of the prime and target words (mean values) for the main analyses in Experiments 1–3, and for the vocalic overlap analyses of Experiment 2

Experiments 1 and 2	Target Words	Repeated Primes	Transposed Primes	Control Primes
Frequency (in number of occurrences per million)	92	92	92	94
Number of phonemes	3	3	3	3
Duration (in ms) in Experiment 1	585	585	581	589
Duration (in ms) in Experiment 2	629	629	624	621

Vocalic Overlap (Experiment 2)	Target Words	Related Primes	Control Primes
Frequency (in number of occurrences per million)	121	115	121
Number of phonemes	3	3	3
Duration (in ms)	637	640	638

Experiment 3	Target Words	Transposed Primes	Vowel + One Consonant Primes	Vowel Primes
Frequency (in number of occurrences per million)	92	92	92	97
Number of phonemes	3	3	3	3
Duration (in ms)	606	598	601	602

data were controlled by a PC running E-Prime software (version 2.0, Psychology Software Tools). In both the prime and target blocks, participants were asked to make a lexical decision as quickly and accurately as possible with "word" responses being made using their dominant hand on an E-Prime response box that was placed in front of them. Reaction times (RTs) were recorded from the onset of stimuli. Within each block, the stimuli were presented randomly. An inter-trial interval of 2,000 ms elapsed between the participant's response and the presentation of the next stimulus. The repeated and transposed targets were separated from their corresponding primes by 2–178 items, with a mean of 89 items. A short break of 1 min separated the prime and target blocks. Participants were tested on only one experimental list and began the experiment with 12 practice trials.

2.2. *Results*

Statistical analyses were performed on the results from the target blocks. The mean RT and percentage of correct responses on target words in each priming condition are presented in Fig. 1. Error rates were overall quite low, and no significant effects were found in an analysis of these data.

Reaction times on target words were analyzed using linear mixed effects models with participants and target words as crossed random factors, using R sotware (R Development

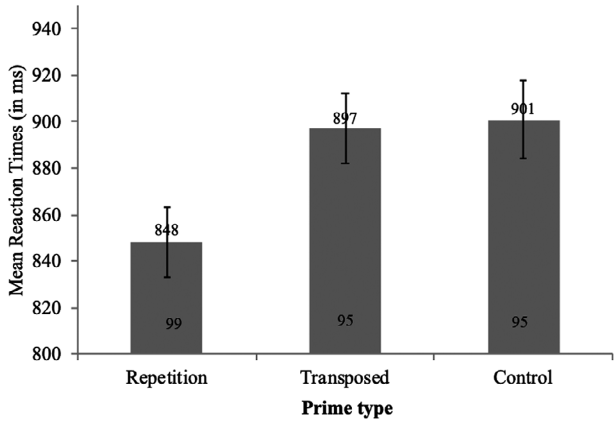

Fig. 1. Mean reaction times (in ms) and standard errors as a function of each type of prime in Experiment 1. Percentages of correct responses are shown below the bar for each condition.

Core Team, 2016) and the *lme4* package (Baayen, Davidson, & Bates, 2008; Bates & Sarkar, 2007). All *p*-values were computed using the Satterwaite approximation for degrees of freedom using lmerTest (Kuznetsova, Brockhoff, & Christensen, 2013). The RT analysis was performed on correct responses, thus removing 74 (3.65%) data points out of 2025. For the model to meet the assumptions of normally-distributed residuals and homogeneity of variance, a log transformation was applied to the RTs (Baayen & Milin, 2010) prior to running the model. The model was run on the totality of the correct RTs, namely 1,951 data points. We tested a model with the variable prime type (repetition, transposed-phoneme, control) entered as fixed effect. The model also included participants and items as random intercepts, plus random participant and item slopes (see Barr, Levy, Scheepers, & Tily, 2013). The reference was the control prime condition. The model revealed a significant repetition priming effect with RTs on target words being 53 ms shorter when preceded by repetition primes in comparison to control primes ($t(404.90) = -7.30$; $p < .001$). However, no significant difference was found between the transposed-phoneme and control primes ($t(401) = -0.81$; $p = .42$).

3. Experiment 2

While a clear repetition priming effect was observed in Experiment 1, we found no evidence for priming across words sharing the same phonemes with the target words in different positions. Because we used the long-term priming paradigm with primes and targets presented in different blocks of stimuli, a possibility, however, is that prime stimuli did lead to partial activation of the target word, but that this activation had time to dissipate before the target was presented. In Experiment 2, we tested the same stimuli as in Experiment 1 using a short-term priming paradigm. Prime and target words were presented in close temporal succession, separated by a 20 ms ISI.

3.1. Method

3.1.1. Participants

Fifty-four French speakers from Aix-Marseille University participated in the experiment. All participants reported having no hearing or speech disorders.

3.1.2. Materials

The same word and non-word pairs and the same design were reused in this experiment except that the primes and the targets were presented immediately one after the other in a single block (as opposed to different blocks in Experiment 1). A further difference with respect to Experiment 1 is that in order to ensure that an eventual transposed-phoneme priming effect was not merely due to the fact that the transposed prime and target pairs have in common their vowel in medial position (LOBE /lɔb/– BOL/bɔl/), we also tested 30 target words that were associated with prime words that shared the medial vowel only (JUPE /jyp/ "skirt" – DUNE /dyn/ "dune") and a set of corresponding unrelated control primes (RICHE /ʀiʃ/ "rich" – DUNE /dyn/ "dune").[1] Their characteristics are given in Table 1. For the purpose of the lexical decision task, and so that the nonwords mimicked words, 30 non-word pairs, 15 with the medial vowel in common and 15 unrelated, were added to the lists. Finally, 150 fillers consisting in prime and target pairs without any relation were added to each list. Again, for the purpose of the lexical decision task, half of the filler targets were words and the other half were non-words. So that participants cannot anticipate the lexical status of the targets on the basis of the primes, the filler target words were preceded by non-word primes, and the filler target non-words were preceded by word primes. As in Experiment 1, the stimuli were recorded by a female native speaker of French, in a sound-attenuated room, and digitized at a sampling rate of 44 kHz with 16-bit analog to digital recording.

3.1.3. Procedure

The primes and the targets were presented sequentially. An interval of 20 ms (ISI) separated the offset of the primes and the onset of the targets. All other aspects of the procedure were identical to Experiment 1.

3.2. Results and discussion

The mean RT and percentage of correct responses on target words in each priming condition are presented in Fig. 2. The same analyses as in Experiment 1 were performed. Error rates were overall quite low, and no significant effects were found in an analysis of error rates.

Analysis was performed on correct responses, thus removing 109 data points out of 2,430 (4.49%). Seven outliers (one RT = 64 ms and the six other RTs > 2,500 ms) were also excluded from the analysis. The model was run on 2,314 data points. The model revealed a significant repetition priming effect with RTs on target words 127 ms shorter when preceded by repeated primes in comparison to control primes ($t(57.33) = -11.28$;

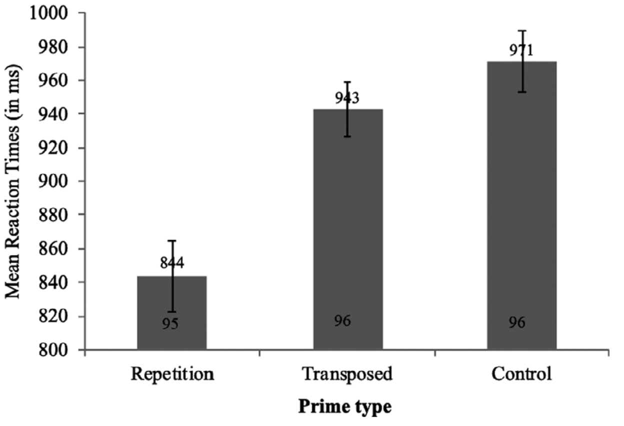

Fig. 2. Mean reaction times (in ms) and standard errors as a function of each type of prime in Experiment 2. Percentages of correct responses are shown below the bar for each condition.

p < .001). Crucially here, the model also revealed a significant transposed-phoneme priming effect with RTs on target words being 28 ms shorter when preceded by transposed-phoneme primes in comparison to control primes (t(45.34) = −2.07; p = .045).

A second model was run on prime-target pairs overlapping in the vowel. The model was run on the correct RTs, thus removing 58 data points out of 1,080 (5.37%). One outlier (RT > 2,500 ms) was also excluded from the analysis. The model was run on 1,021 data points and was tested with the variable prime type (related, control) entered as fixed effect. The model also included participants and items as random intercepts, plus random participant and item slopes (Barr et al., 2013). The model revealed no significant difference (t(26.98) = −1.01; p = .32) between the control (Mean RT = 970 ms) and the related primes (Mean RT = 964 ms).

To summarize, using the short-term priming paradigm, we were able to observe a significant priming effect with prime and target pairs sharing all of their phonemes but in a different order. Such a finding thus constitutes further evidence that phonemes are processed and encoded independently of their position in the speech signal. Since our transposed-phoneme primes did share their central vowel at the same position as in targets, we also tested for effects of vowel overlap alone, and this was found not to be significant. We are therefore confident that the transposed-phoneme priming effect is not being driven by vocalic overlap alone. However, a more powerful demonstration of the transposed-phoneme priming effect would be to observe it using vowel overlap primes as control condition instead of unrelated primes. This was the aim of Experiment 3.

4. Experiment 3

The aim of Experiment 3 was twofold. First, we wanted to replicate the transposed-phoneme priming effect with primes sharing the medial vowel with the targets as control condition (/pys/-/tyb/). Second, primes and targets sharing the medial vowel and only one consonant in an incorrect position (/byl/-/tyb/) were also used. Such a condition allowed us to examine the impact of the degree of phoneme overlap (2/3 vs. 3/3) on position-independent priming effects.

4.1. Method

4.1.1. Participants

Fifty-four French speakers from Aix-Marseille University participated in the experiment. All participants reported having no hearing or speech disorders.

4.1.2. Materials

The Forty-five transposed prime-targets pairs (/byt/-/tyb/) used in Experiments 1 and 2 were re-used. For each target word, two other prime words were also selected. One shared only the medial vowel (/pys/-/tyb/) with the target word, and the other shared both the medial vowel and one consonant in the incorrect position (/byl/-/tyb/). The main characteristics of the prime and the target words are given in Table 1. The prime and target words are given in Appendix B. All the stimuli were recorded by a female native speaker of French, in a sound-attenuated room, and digitized at a sampling rate of 44 kHz with 16-bit analog to digital recording.

Three experimental lists were also created so that each of the 45 target words was preceded by the three types of prime (transposed, vowel overlap, vowel plus one consonant overlap), and participants were presented with each target word only once. One hundred fifty fillers consisting in prime and target word pairs without any relation were added to each list. All of the filler primes were words, and for the purpose of the lexical decision task, half of the filler targets were words and the other half were non-words.

4.1.3. Procedure

This was the same as in Experiment 2.

4.2. Results and discussion

The mean RT and percentage of correct responses on target words in each priming condition are presented in Fig. 3. The same analyses as in Experiments 1 and 2 were performed. Error rates were overall quite low, and no significant effects were found in an analysis of error rates.

Analysis was performed on correct responses, thus removing 185 data points out of 2,430 (7.61%). Seventeen outliers (1 RT = 56 ms and the 16 other RTs > 2,500 ms) were also excluded from the analysis. The model was run on 2,228 data points. A model with vowel overlap primes as reference revealed a significant transposed-phoneme priming effect with RTs on target words being 26 ms shorter when preceded by transposed primes in comparison to vowel overlap primes ($t(173.63) = -2.34$; $p = .021$). No significant difference was found between vowel plus one consonant primes and vowel primes

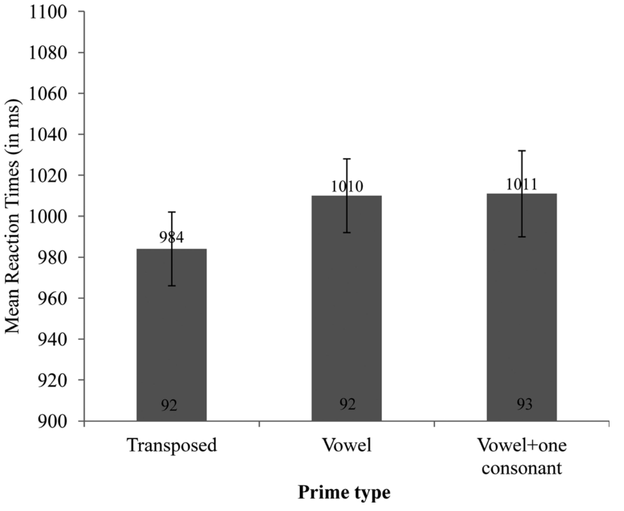

Fig. 3. Mean reaction times (in ms) and standard errors as a function of each type of prime in Experiment 3. Percentages of correct responses are shown below the bar for each condition.

($t(37.65) = 0.13$; $p = .90$). Finally, the model was re-referenced to the vowel plus one consonant overlap condition. A significant difference was found with RTs on target words being 27 ms shorter when preceded by transposed primes in comparison to vowel plus one consonant overlap primes ($t(51.20) = -2.17$; $p = .035$). In sum, the results of Experiment 3 replicate the transposed-phoneme priming effect found in Experiment 2 when measured against a vowel overlap prime condition. Moreover, when vowel overlap was combined with a shared consonant in a different position, there was no significant difference with the vowel-only condition, and the transposed-phoneme prime condition produced significantly faster responses than the vowel plus one consonant condition.

5. General discussion

Toscano et al. (2013) reported evidence that a speech input like [kat] not only provides support for the corresponding lexical representation *cat* but also for the lexical representation that contains the same phonemes in a different order *tack*. This finding counters the predictions of all current models of spoken word recognition except one—the TISK model of Hannagan et al. (2013). Therefore, given the theoretical importance of Toscano et al.'s finding, in this study, we sought further evidence for a role for position-independent phonemes in spoken word recognition. We used a long-term priming paradigm in Experiment 1, with primes and targets presented in two separated blocks, and a short-term priming paradigm in Experiments 2 and 3 with targets immediately following primes. In Experiments 1 and 2 the primes could be the same word as targets (e.g., /tyb/-/tyb/), or words formed of the same phonemes as targets but in a different order (e.g., /byt/-/tyb/), or words formed of different phonemes (e.g., /mul/-/tyb/). There was no evidence for transposed-phoneme priming in Experiment 1, but Experiment 2 revealed a robust priming effect. The short-term priming effect observed in Experiment 2 with transposed-phoneme primes was replicated in Experiment 3 when both primes sharing the medial vowel with the targets (e.g., /pys/-/tyb/) and primes sharing the medial vowel plus one consonant in an incorrect position with the targets (/byl/-/tyb/) served as control condition. Hence, our results strengthen Toscano et al.'s (2013) observation and provide further evidence that words that share all their phonemes with a given target word but in a different order receive a certain amount of bottom-up support during spoken word recognition. It is the partial activation of such words that is thought to have driven the eye-movement patterns in Toscano et al.'s study and the short-term priming effects of this study. Our results also confirm the eye-tracking observations that the transposed-phoneme effect is due to more than vowel position overlap in the transposed words (Gregg et al., in press; Toscano et al., 2013). They also confirm that two shared phonemes with one phoneme in a different position cause lesser activation of the target words (Toscano et al., 2013), and that complete phonemic overlap is a necessary condition in order to obtain transposed-phoneme effects. Our study therefore provides further evidence for a role for position-independent phonemes during spoken word recognition, as postulated in the TISK model of Hannagan et al. (2013).

We also found that repetition priming effects were significantly greater than transposed-phoneme priming effects in both experiments. Indeed, transposed-phoneme priming effects differed from repetition priming effects both by their magnitude and their time-course, with only repetition effects being significant in a long-term priming paradigm. The partial activation interpretation that we propose easily accounts for this pattern by assuming that activation dissipates over time, such that only fully activated lexical representations can influence long-term priming. Within the same activation framework, an alternative interpretation would be that long-term priming effects require that lexical representations reach a criterion level of activation during prime processing in order to trigger an adjustment in connection strengths, and it is this adjustment in connection strengths, a learning mechanism, that subtends long-term priming (see Monsell & Hirsh, 1998). In either case, while our results provide support for the existence of position-independent phonemes, they also demonstrate that phoneme order information is a crucial ingredient of spoken word recognition. This is the case in the TISK model (Hannagan et al., 2013), where a set of bi-phone representations, that co-exist with position-independent phonemes, encodes the order of phonemes. Furthermore, as discussed by Toscano et al. (2013), the duration of phonemes can vary according to their position, with, for example, phonemes that occur at the beginning of words having a shorter duration than when they occur at the end of words (see Shatzman & McQueen, 2006; Spinelli, McQueen, & Cutler, 2003). Duration information could therefore be combined with position-independent phoneme identities in order to obtain some information about phoneme position.

However, the results of a recent study (Gregg et al., in press) impose further constraints on how positional information is encoded during spoken word recognition. Using the eye-tracking paradigm, the authors replicated Toscano et al.'s (2013) findings that transposed-phoneme competitors (GUM-MUG) were fixated more than unrelated words. At the same time, they showed that competitors without vowel position overlap (LEAF-FLEA) were not fixated more that unrelated words, thus indicating that positional vowel match is critical in the observation of the transposed-phoneme effect. As discussed by Gregg et al. (in press), such a finding could argue for a special status for vowels that would be processed more rapidly than consonants. Within the framework of the TISK model, the findings of Gregg et al. suggest that the position-independence of phoneme identities might be constrained by the abstract CV-structure of the stimulus being processing. That is, when processing a CVC word, then consonants can migrate across their respective positions, but not to the central vowel position.

To sum up, the key result of this study is the finding that, in a short-term priming paradigm with target words immediately following prime words, auditory lexical decisions to target words were facilitated when the prime word was formed of the same phonemes but with the first and last phonemes transposed (e.g., /byt/-/tyb/)—a transposed-phoneme priming effect. Building on the findings of Toscano et al. (2013) obtained with the visual world paradigm, the present findings demonstrate that such transposed-phoneme effects can be generalized beyond a specific task or a particular experimental procedure. Taken together, these results clearly suggest that at some point during the process of spoken word recognition, phoneme identities are encoded independently of their position in the speech signal. The key question for future research is how such position-independent encoding combines with the type of encoding that enables an accurate representation of phoneme order. The TISK model (Hannagan et al., 2013) provides one possible solution, and this model could therefore be usefully applied to generate predictions to be tested in future experimentation. Finally, because the present study, as well as that of Toscano et al. (2013), used short words (three phonemes), future research should examine whether the effect generalizes to longer words and examine its relation with other factors involved in spoken word recognition, such as phonological uniqueness point, and the CV-structure of the stimulus word and its competitors.

Note

1. For the analysis of effects of vocalic overlap, the counterbalancing between related and control primes required two lists only. In order that all participants be tested in exactly the same experimental conditions, the 15 targets words preceded by the related primes and the 15 other targets words preceded by the control primes and used in the first list were also presented in the third list. Nonetheless, for the vocalic overlap, the analyses were performed on the two first lists only, and thus on 36 rather than 54 participants. Note also that it would have been necessary to test vocalic overlap primes in a long-term priming experiment if a transposed-phoneme priming effect had emerged in Experiment 1. This was not the case.

References

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

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*, 12–28.

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

Bates, D. M., & Sarkar, D. (2007). lme4: Linear mixed-effects models using S4 classes. R package version 2.6, available at http://lme4.r-forge.r-project.org/. Accessed November 20, 2017.

Connine, C. M., Blasko, D. M., & Titone, D. A. (1993). Do the beginning of words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*, 193–210.

Dufour, S., & Frauenfelder, U. H. (2010). Phonological neighbourhood effects in French spoken word recognition. *Quarterly Journal of Experimental Psychology*, *63*, 226–238.

Dufour, S., & Nguyen, N. (2014). Access to talker-specific representations is dependent on word frequency. *Journal of Cognitive Psychology*, *26*, 256–262.

Dufour, S., & Nguyen, N. (2017). Does talker-specific information influence lexical competition? Evidence from phonological priming. *Cognitive Science*, *41*, 2221–2233.

Dufour, S., & Peereman, R. (2003). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory and Cognition*, *31*, 1271–1283.

Dufour, S., Peereman, R., Pallier, C., & Radeau, M. (2002). VoCoLex: A lexical database on phonological similarity between French words. *L'Année Psychologique*, *102*, 725–746.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: a distributed model of speech perception. *Language and Cognitive Processes*, *12*, 613–656.

Gregg, J., Inhoff, A. W., & Connine, C. M. (in press). Re (Re) re-considering the role of temporal order in spoken word recognition. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/1747021819849512

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, *31*, 423–445.

Hannagan, T., Magnuson, J. S., & Grainger, J. (2013). Spoken word recognition without a TRACE. *Frontiers in Psychology*, *4*, 563.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests in linear mixed effects models (R Package Version 2.0-11).

Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of a lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*, 245–294.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*, 1–36.

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1376–1392.

Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*, 653–675.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interaction and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

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

McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 306–321.

Mitterer, H., Reinisch, E., & McQueen, J. M. (2018). Allophones, not phonemes in spoken-word recognition. *Journal of Memory and Language*, *98*, 77–92.

Monsell, S., & Hirsh, K. W. (1998). Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 1495–1520.

Norris, D. (1994). SHORTLIST: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.

R Development Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. Available at: http://www.R-project.org.

Shatzman, K. B., & McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception and Psychophysics*, *68*, 1–16.

Spinelli, E., McQueen, J., & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, *48*, 233–254.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, *60*(4), 487–501.

Toscano, J. C., Anderson, N. D., & McMurray, B. (2013). Reconsidering the role of temporal order in spoken word recognition. *Psychonomic Bulletin & Review*, *20*, 981–987.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374–408.

You, H., & Magnuson, J. (2018). TISK 1.0: An easy-to-use Python implementation of the time-invariant string kernel model of spoken word recognition. *Behavior Research Methods*, *50*, 871–889.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25–64.

Appendix A: Prime and target words used in Experiments 1 and 2

Control Primes	Transposed Primes	Repeated Primes/Targets
comte	lobe	bol
page	robe	bord
vache	cache	chaque
monde	chatte	tâche
moule	sec	caisse
singe	cale	lac
puce	rêche	chair
date	lâche	châle
vigne	chic	quiche
pince	chope	poche
cure	lisse	cil
fiche	loque	col
couche	loupe	poule
galle	toque	cotte
lampe	rade	dard
	digue	guide

(continued)

Appendix A. (continued)

Control Primes	Transposed Primes	Repeated Primes/Targets
	dire	ride
	douce	soude
	pouce	soupe
	rousse	sourd
	dur	rude
	geine	neige
	jarre	rage
	jour	rouge
	laisse	sel
	lame	mal
	râle	lard
	mille	lime
	nulle	lune
	mare	rame
	tonne	note
	rhume	mur
	panne	nappe
	nerf	reine
	niche	chine
	tape	patte
	pente	tempe
	quitte	tique
	top	pote
	verre	rêve
	gîte	tige
	route	tour
	rate	tard
	butte	tube
	coche	choc

Appendix B: Prime and target words used in Experiment 3

Vowel Primes	Vowel + One Consonant Primes	Transposed Primes	Targets
vote	lotte	lobe	bol
vol	rosse	robe	bord
vase	casse	cache	chaque
bar	date	chatte	tâche
veine	bec	sec	caisse
natte	galle	cale	lac
pelle	raide	rêche	chair
gaffe	vache	lâche	châle

(continued)

Appendix B. (continued)

Vowel Primes	Vowel + One Consonant Primes	Transposed Primes	Targets
bise	pic	chic	quiche
fort	top	chope	poche
rive	lire	lisse	cil
roche	phoque	loque	col
douche	louche	loupe	poule
loge	tord	toque	cotte
cape	fade	rade	dard
pile	ligue	digue	guide
fille	dix	dire	ride
touche	mousse	douce	soude
goutte	tous	pouce	soupe
bouc	rouille	rousse	sourd
jupe	pure	dur	rude
messe	laine	geine	neige
bague	phare	jarre	rage
doute	court	jour	rouge
fer	baisse	laisse	sel
face	latte	lame	mal
baffe	rabe	râle	lard
type	ville	mille	lime
sud	nuque	nulle	lune
dalle	masse	mare	rame
corps	bonne	tonne	note
fugue	ruche	rhume	mur
bac	pack	panne	nappe
chèque	net	nerf	reine
rime	fiche	niche	chine
gaz	tasse	tape	patte
chance	lente	pente	tempe
riche	vite	quitte	tique
bosse	tome	top	pote
fête	mer	verre	rêve
biche	rite	gîte	tige
bouche	soute	route	tour
bave	race	rate	tard
puce	bulle	butte	tube
botte	moche	coche	choc