Investigating the locus of transposed-phoneme effects using cross-modal priming

Sophie Dufour, Jonathan Mirault, Jonathan Grainger

Investigating the locus of transposed-phoneme effects using cross-modal priming

Sophie Dufour[1,3] Jonathan Mirault[2,] and Jonathan Grainger[2,3]

1. Aix-Marseille Université, CNRS, LPL, UMR 7309, 13100 Aix-en-Provence, France

2. Laboratoire de Psychologie Cognitive, Aix-Marseille Université & CNRS

3. Institute for Language, Communication, and the Brain, Aix-Marseille Université, Aix-en-Provence, France

Short title: Cross-modal transposed-phoneme priming

Sophie Dufour
Laboratoire Parole et Langage
Aix-Marseille Université, CNRS
5, avenue Pasteur
13604 Aix-en-Provence
France Email: sophie.dufour@univ-amu.fr

Abstract: In this study, we examined whether the facilitatory priming effect found when auditory primes and targets are related by a phoneme transposition (e.g., /ʀɔb/-/bɔʀ/: Dufour & Grainger, 2019, 2020) is also observed under cross-modal presentation. In two experiments using the same materials as in the previous studies, we found no evidence for a facilitatory priming effect when the targets were presented visually rather than auditorily. On the contrary, an inhibitory priming effect was found when both unrelated words (Experiment 1; e.g., /mas/-/bɔR/) and vowel overlap words (Experiment 2; e.g., /vɔl/-/bɔʀ/) were used as control conditions. In Experiment 2, this inhibitory effect was found to be equivalent in size whether the target words were of higher or lower frequency than the prime words (e.g. /ʀɔb/-/bɔʀ/ vs. /bɔʀ/-/ʀɔb/). We interpret this pattern of effects as reflecting the greater impact of word-level inhibition in cross-modal priming, and the parallel influence of prime-target relative frequency on bottom-up phoneme-to-word facilitation and word-level inhibition. Therefore, the facilitatory priming effect previously observed with auditory primes and targets would mainly reflect bottom-up activation of the target word representation during prime word processing.

Key words: Transposed-phoneme effects; Cross-modal priming; Speech processing

One key assumption in the literature on spoken word recognition is that as listeners attend to a word, not only the lexical representation of the word itself is activated but also the representations of phonologically similar words. Since the first phonemes of words are heard and begin to be processed before later phonemes, the first accounts of spoken-word recognition logically assumed that the set of activated lexical candidates consists of words that share their initial phonemes with targets (e.g., Marslen-Wilson & Welsh, 1978). As a result, upon hearing the French word BUT /byt/ “goal” words like /bys/, /byl/, /bytɛ̃/ etc. are activated, and thus constitute potential candidates for recognition. Later models ascribed less weight to the initial portion of the speech signal, and the set of activated lexical candidates was extended to include words that align with the speech input other than at the onset (McClelland & Elman, 1986; Norris, 1994). As a result, a word like /byt/ can also activate words like /lyt/ because the two words share their final phonemes and are thus aligned with the final portion of the input.

These two definitions of what constitutes a lexical candidate are in direct connection with the sequential nature of the speech signal, and originate from the dominant view of spoken word recognition according to which the sounds extracted from the speech signal are encoded according to their position in the input so as to be successfully mapped onto an ordered sequence of sounds stored in long-term memory. However, a growing body of research suggests that a word like /tyb/ that contains the same phonemes as /byt/ in a different order can also be part of the set of activated lexical candidates. The first demonstration was provided by Toscano et al. (2013). Using the visual world paradigm, these authors 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.

Such a finding thus suggests that CAT and TACK are confusable words even if the consonants that they shared are not in the same positions. Also, Toscano et al. (2013) showed that the probability of fixating transposed words was higher than the probability of fixating words sharing the same vowels at the same position plus one consonant in a different position (e.g., SUN-BUS). This finding suggests that the transposed-phoneme effect is due to more than just vowel position overlap in the transposed words, and that complete phonemic overlap is a necessary condition in order to obtain transposed-phoneme effects. Using the phonological priming paradigm and the lexical decision task, two other studies conducted in French (Dufour & Grainger, 2019; 2020) have also shown that speech input like [byt] facilitates not only the subsequent processing of an identical target word /byt/ BUT "goal" but also that of a target word /tyb/ TUBE "tube" that contains the same phonemes in a different order. This transposed-phoneme priming effect was found when unrelated words (*MOULE* /mul/ "mussel" – *TUBE* /tyb/ "tube"), vowel overlap words (*PUCE* /pys/ "flea" - *TUBE* /tyb/ "tube") and vowel plus one consonant in a different position overlap words (*BULLE* /byl/ bubble – *TUBE* /tyb/ "tube") were used as control conditions, thus providing further support to prior observations of transposed-phoneme effects. Dufour and Grainger (2020) also reported that the transposed-phoneme priming effect occurs when targets have a higher frequency than primes, but not when they have a lower frequency than primes.

Taken together, these findings suggest that position-independent phonemes play a role during spoken word recognition, thus challenging the dominant view according to which the precise order of phonemes must be encoded (see Bowers et al., 2016, for further evidence). The transposed-phoneme effect found in phonological priming studies (Dufour & Grainger, 2019, 2020) has been accounted for within the framework of the TISK model of spoken word recognition (Hannagan et al., 2013; You & Magnuson, 2018). TISK is an interactive-activation model similar to the TRACE model (McClelland & Elman, 1986), but it replaces the position-dependent units in TRACE by both a set of position-independent phoneme units and a set of open-diphone units that represent ordered sequences of contiguous and non-contiguous phonemes. Within such a framework, both the position-independent phoneme units and the open-diphone representations can contribute to transposed-phoneme effects. Dufour and Grainger (2019, 2020) mainly interpreted transposed-phoneme priming effects as resulting from the bottom-up activation of transposed-phoneme words during prime processing, thus facilitating their processing when they are subsequently presented as target words (a phoneme-to-word excitatory effect). In this perspective, due to faster increase in activation, the higher the frequency of a transposed-phoneme target word the more it should be activated during prime processing, thus accounting for the greater facilitation effect observed when targets were of higher frequency than the primes in comparison to when the targets were of lower frequency than the primes.

The most obvious explanation of the results of Dufour and Grainger (2020) is in terms of phoneme-to-word activation which would be greater for high-frequency target words due to greater connection strengths between phonemes and words for high-frequency words. However, the above explanation of the transposed priming effect on the one hand and of the impact of prime-target relative frequency on the other hand, couched within the general framework of interactive-activation models (McClelland & Elman, 1986; Hannagan et al., 2013), ignores one key principle of such models – that of word-level inhibition. Indeed, according to interactive-activation models the transposed-phoneme word that is activated upon presentation of a target word should enter in a competition process (lateral inhibition) whereby the target word receives inhibitory input from the co-activated transposed-phoneme word. In the present study, we attempted to find evidence for this competitive influence of transposed-phoneme words, by using a cross-modal priming procedure. This was motivated by the findings of prior comparisons of priming effects obtained within and across modalities. When primes and targets are phonological neighbors with initial phoneme overlap, then robust inhibitory priming effects are observed when primes and targets are presented auditorily and also under cross-modal presentation (Radeau, 1995; Slowiaczek et al. 1992). On the other hand, facilitatory phonological priming effects, obtained when primes and targets share their final phonemes, and in particular the rime, only arise when both primes and targets are presented auditorily (Dumay et al., 2001; Radeau, 1995; Spinelli et al., 2001). This pattern of results suggests that the cross-modal priming procedure gives more weight to inhibitory lexical influences relative to facilitatory sublexical influences in determining net priming effects. Thus, in Experiment 1 we examined whether the transposed-phoneme priming effects reported by Dufour and Grainger (2019) would be affected by a shift from within-modality to cross-modal priming. We hypothesized that the cross-modal procedure would increase the impact of within-level lexical competition at the expense of bottom-up phoneme-to-word facilitation, hence leading to the emergence of inhibitory transposed-phoneme priming effects.

Experiment 1

In Experiment 1 we tested the same stimuli as in Dufour and Grainger's (2019) study but under conditions of cross-modal presentation, such that primes were presented auditorily and targets presented visually.[1]

Method

Participants: Thirty-six native French speakers from Aix-Marseille University participated in the experiment. All participants reported having no hearing or speech disorders.

Materials: The prime-target pairs were taken from Dufour and Grainger (2019)'s study (Experiment 2). They consisted of forty-four pairs of CVC French words sharing all of the phonemes but with the two consonants in different order (e.g. prime LOBE /lɔb/ "*lobe*"; target BOL /bɔl/ "*bowl*"). Each target word was also associated with a control prime word sharing no phoneme with the target (e.g. COMTE /kɔ̃t/ "count"-/bɔl/). Moreover, 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/), the materials also included 30 target words that were associated with prime

[1] Note that as we were interested in processes occurring during auditory word recognition, and as priming effects are generally interpreted in reference to processes occurring during prime presentation, either in terms of activation and/or inhibition of the target word during priming processing, it is the primes that were presented auditorily.

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”). The main characteristics of the primes and the targets are given in Table 1. The prime and target words are given in Appendix 1.

<Insert Table 1 about here>

Two experimental lists were created so that each of the 74 target words were preceded by the two types of prime (related, control), and participants were presented with each target word only once. Each list included the 44 target words used in the transposition condition, half preceded by the control primes and the other half preceded by the related primes, and the 30 target words used in the vocalic condition, again half preceded by the control primes and the other half preceded by the related primes. For the purpose of the lexical decision task, 74 CVC nonwords were added to each list. So that the non-words mimicked words 22 of them consisted of pairs sharing the same phonemes but in different order (e.g., /nab/ - /ban/), 15 of them consisted of pairs sharing the medial vowel e.g., /sip/-/lif/) and the remaining 37 consisted of unrelated pairs (/ʃãd/-/sub/). Finally, 148 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.

Procedure: The audio files used in Dufour and Grainger (2019), with acoustic recording made by a female native speaker of French, were re-used here. The participants were tested in a sound-attenuated booth. Stimulus presentation and recording of the data were controlled by a PC running E-Prime software. The primes were presented over headphones at a comfortable sound level. At the end of the auditory primes, the targets were visually displayed in upper case at the center of the screen and remained until the participant's response. 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. The computer clock was triggered by the presentation of the target and was stopped by the response. The prime-target pairs were presented randomly, and an inter-trial interval of 2000 ms elapsed between the participant's response and the presentation of the next pair. Participants were tested on only one experimental list and began the experiment with 12 practice trials.

Results and Discussion

RTs to target words (available at https://osf.io/4spq3/; Open Science Framework; Foster & Deardorff, 2017) in the transposition condition were first analyzed using linear mixed effects models with participants and items as crossed random factors, using R software (R Development Core Team, 2016) and the lme4 package (Baayen et al., 2008; Bates & Sarkar, 2007). The RT analysis was performed on correct responses, thus removing 126 data points out of 1584 (7.95%). 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 1458 data points. We tested a model with the variable Prime Type (transposed, control) as fixed effect. The model failed to converge when random participant and item slopes were included (see Barr et al., 2013). Therefore the final model included only random intercepts for participants and items.

The effect of prime type was significant (*b* =0.0367, *SE* = 0.0128, *t* = 2.86, *p*<.01). As expected, RTs on target words were 41 ms slower when preceded by related primes (Mean RT = 700 ms) in comparison to control primes (Mean RT = 659 ms).

In order to examine whether a vocalic overlap is sufficient to cause a facilitation effect, a second model was run on prime-target pairs overlapping in the vowel following exactly the same procedure as previously. The model was run on the correct RTs, thus removing 48 data points out of 1080 (4.44%). The model was run on 1032 data points and was tested with the variable prime type (related, control) entered as fixed effect. The model revealed no significant difference (*b* =0.0103, *SE* = 0.0134, *t* = 0.77, *p*>.20) between the control (Mean RT = 659 ms) and the related primes (Mean RT = 667 ms). Hence, the transposed-phoneme priming effect was likely not caused by the fact the primes and the targets shared the medial vowel.

The percentage of correct responses was analyzed using a mixed-effects logit model (Jaeger, 2008) following the same procedure as for RTs. No significant effects were found (Mean = 92 % for both control and related primes of the transposition condition; Mean = 95 % for both control and related primes of the vocalic condition).

The results are clear-cut. In contrast to what we observed under unimodal presentation (Dufour & Grainger, 2019; 2020), here we obtained no sign of a facilitation effect under cross-modal presentation. On the contrary, the prior presentation of an auditory prime word that shares all its phonemes with a target word but in a different order inhibited identification of the visually presented target. Hence, using exactly the same materials as in Dufour and Grainger (2019), the 28 ms facilitation effect previously found (i.e., 943 and 971ms for the transposed and control primes, respectively) became an inhibitory effect of 41 ms (i.e., 700 and 659 ms for the transposed and control primes, respectively) simply by changing the modality of target presentation. This is perfectly in line with our prediction that the cross-modal procedure would give more weight to lexical (word-level inhibition) relative to sublexical (phoneme-to-word facilitation) influences on transposed-phoneme priming effects. Given the importance of this result regarding the involvement of lexical vs. sublexical influences in driving cross-modal transposed-phoneme priming, we deemed it important to replicate the inhibition effect found in Experiment 1 using another set of items.

Experiment 2

In Experiment 2, the pairs of prime and target words were taken from Dufour and Grainger (2020)'s study. In this prior study, the relative frequency of the prime and target words was manipulated so that the targets were either of lower or of higher frequency than the primes, and this manipulation was maintained in the present study. In the case where the inhibitory priming effect under cross-modal presentation is replicated, this would allow us to examine whether lexical frequency influences the magnitude of the inhibition effect, in the same way that lexical frequency was found to modulate the facilitation effect under unimodal presentation. Furthermore, in order to strengthen the claim that the transposed-phoneme priming effect is not driven by vocalic overlap alone, control primes had in common the central vowel with targets.

Method

Participants: A total of 141 participants were recruited on-line for the experiment. All participants indicated that French was their native language. Because on-line experimentation facilitates both the recruitment of participants and running the experiment, we decided to increase the number of participants to provide a stronger test of the crucial inhibitory effect seen in Experiment 1.

Materials: Exactly the same materials as in Dufour and Grainger (2020) were re-used. They consisted of 26 pairs of French transposed-phoneme words with a CVC syllabic structure (e.g. *ROBE* /ʀɔb/ *"*dress*"* – *BORD* /bɔʀ/ "edge"). In each pair, the words differed in frequency, and thus each of the words of a pair was used as a prime or as a target depending of the frequency condition. In each frequency condition, each target word was also associated with a control prime word sharing with the target only the medial vowel (e.g. *VOL* /vɔl/ "flight" – *BORD* /bɔʀ/ "edge" for the higher-frequency target condition; *SOMME* /sɔm/ "sum" – *ROBE* /ʀɔb/ "dress" for the lower-frequency target condition). Note that in each frequency condition, the transposed and control primes were matched in terms of positional similarity with the targets. Initial and final consonant phonetic similarity was calculated taking into account the four phonetic features generally used in French phonology (place, voice, manner and nasality). As an illustration, consider the transposed prime /Rɔb/ and the control prime /vɔl/ for the target word /bɔR/. The phonemes /R/ and /b/ of the transposed-phoneme prime-target pair are both voiced and oral consonants, and thus the transposed-phoneme prime and the target share two phonetic features out of 4 for the initial and final consonants. Also, the /v/ of the control prime and the /b/ of the target word are both voiced and oral consonants, and the /l/ of the control prime and the /R/ of the target word are both liquids, voiced and oral consonants. Thus, the control prime and the target word share two phonetic features out of 4 for the initial consonant and three phonetic features out of 4 for the final consonant. It should be noted that for the phonetic feature “manner”, we made the distinction between fricative and liquid consonants. Also, given existing evidence for activation of orthographic codes during speech processing, we calculated the orthographic similarity between the primes and targets. We used the Levenshtein distance which consists in counting the minimum number of single-letter changes (replacement, addition, deletion) required to change one word into the other. For instance the Levenshtein distance between the transposed-phoneme prime ROBE /Rɔb/ and the target word BORD is 3, namely three letter substitutions. Also the distance between the control prime VOL /vɔl/ and the target word BORD is 3, namely two letter substitutions and one letter addition. As shown in Table 2, using this metric, the transposed-phoneme primes are not more orthographically similar to target words than the control primes. Note also that none of the transposed-phoneme prime words were single letter substitution neighbors of the target words (i.e., Coltheart et al.’s, 1977, definition of orthographic neighborhood). Finally, given the evidence for transposed-letter effects (e.g., Perea & Lupker, 2003; 2004; Perea et al., 2008), we also checked that the transposed phoneme primes were not transposed-letter neighbors of the target words. Among the 26 prime-target pairs only two was a transposed-letter pair. The main characteristics of the prime and the target words are given in Table 2. The prime and target words are given in Appendix 2.

<Insert Table 2 about here>

Exactly the same experimental lists as in Dufour and Grainger (2020) were re-used. They consist of two experimental lists by frequency condition created in such way that each of the 26 target words were preceded by the two types of prime (transposed, control), and participants were presented with each target word only once. Note that for the same reasons as in Dufour and Grainger (2020), a between-participants design for the factor relative prime-target frequency was used. This permitted to have a sufficient number of trials for each type of prime across lists (i.e. 13 in this present case), while avoiding stimuli repetition within participants, such that a given prime or target was never heard twice by a same participant. For the purpose of the lexical decision task, the list also included 26 target nonwords. The target non-words mimicked words, and thus 13 of them were paired with a prime word sharing the same phonemes but in a different order (e.g. the prime word *JAMBE* /ʒãb/ "leg" and the nonword target /bãʒ/), and the 13 other nonwords were paired with a prime word sharing only the medial phoneme (e.g. the prime word *FOUR* /fuʀ/ "oven" and the nonword target /mup/). Finally, the lists also included 78 unrelated prime-target pairs having no phoneme in common were added to each list. Half of the unrelated pairs consisted of a prime word and a target word (e.g. *GUERRE* /gɛʀ/ "war" – *DANSE* /dãs/ "dance"), and the other half consisted of a prime word and a target nonword (e.g. LUGE /lyʒ/ "luge" – /bif/).

Procedure: The audio files used in Dufour and Grainger (2020), with acoustic recording made by a female native speaker of French, were re-used here. Exactly the same procedure as in Experiment 1 was used except that the experiment was programmed using LabVanced software (Finger et al., 2017), and participants gave their responses with the left and right arrows of their personal computer keyboard.

Results and Discussion

Eight participants were excluded from the analyses. Among them five participants had an error rate above 50%, and three others declared that they were more than sixty years old. The mean RT and percentage of correct responses to target words in each condition are presented in Table 3.

<Insert Table 3 about here>

RTs to target words (available at https://osf.io/4spq3/) were analyzed using linear mixed effects models following the same procedure as in Experiment 1. The RT analysis was performed on correct responses, thus removing 406 data points out of 3458 (11.74%). Two RTs <350 ms and 13 RTs > 4,000 ms were considered as outliers (less than 1%) and were excluded from the analysis. 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 3037 data points. We tested a model with the variable Prime Type (transposed, control), Target Frequency (lower, higher) and their interaction entered as fixed effect. The model failed to converge when random participant and item slopes were included (see Barr et al., 2013). Therefore the final model included only random intercepts for participants and items. We applied orthogonal contrast coding for the independent variables, namely 0.5 for one condition and -0.5 for the other condition, which allows an estimation of main effects.

The main effect of Target Frequency was significant (*b* =-0.1192, *SE* = 0.0444, *t* = -2.68, *p*<.01) with RTs on target words shorter for high-frequency words than low-frequency words. The main effect of Prime Type was significant (*b* =0.0245, *SE* = 0.0104, *t* = 2.36, *p*<.05). RTs on target words were 27 ms slower when preceded by transposed primes in comparison to control primes. The interaction between Prime Type and Target Frequency was not significant (*b* =-0.0039, *SE* = 0.0207, *t* = -0.19, *p*>.20). To evaluate the non-significance of this critical interaction, a Bayes Factor analysis was conducted using the lmBF function from the BayesFactor package (Morey, 2015). The BF_{01} value was 24 thus indicating strong evidence in favor of H0.

The percentage of correct responses was analyzed using a mixed-effects logit model (Jaeger, 2008) following the same procedure as for RTs. Only the main effect of Target Frequency was significant (*b* =2.3196, *SE* = 0.5889, *z* = 3.94, *p*<.001) with more correct responses for high-frequency targets than low-frequency targets.

To sum-up, Experiment 2 replicated the inhibitory priming effect found in Experiment 1 with a larger sample size and a different set of items that were better controlled in terms of vocalic overlap, and phonetic and orthographic similarity across primes and targets. We are thus confident that the prior presentation of an auditory prime word that shares all its phonemes with the target word but in a different order inhibits the subsequent processing of the visually presented target. Moreover, in contrast to what was previously observed for the facilitation effect found under unimodal presentation, the inhibition effect found with transposed-phoneme prime-target pairs with cross-modal presentation did not vary as a function of relative prime-target frequency. Hence the size of the inhibition effect observed under cross-modal presentation did not differ significantly across the two levels of prime-target relative frequency (31 and 23 ms for the lower-frequency and the higher-frequency target conditions, respectively; see Table 3), whereas the 45 ms facilitation effect found under unimodal presentation in Dufour and Grainger (2020) when the targets were of higher frequency than the primes (i.e., 989 and 1034 ms for the transposed and control primes, respectively) became a null effect when the targets were of lower frequency than the primes (i.e., 1112 and 1103 ms for the transposed and control primes, respectively). The observation that relative prime-target frequency does not influence the magnitude of inhibitory priming in Experiment 2 can be tentatively interpreted as due to cross-modal transposed-phoneme priming being the result of a trade-off between facilitatory phoneme-to-word excitatory influences and inhibitory word-level influences. We provide more details concerning this explanation in the General Discussion.

General Discussion

In two experiments using exactly the same materials as in two previous studies reporting facilitatory transposed-phoneme within-modality (auditory-auditory) priming effects (Dufour & Grainger, 2019; 2020), we found no evidence for a facilitatory effect when the targets were presented visually rather than auditorily. On the contrary, we found that the prior presentation of an auditory prime word, that shared all its phonemes with the following target word but in a different order, inhibited identification of visually presented target words. This inhibitory priming effect was found when unrelated words (Experiment 1; e.g., /kɔ̃t/-/bɔl/) and vowel overlap words (Experiment 2; e.g., /sɔm/-/bɔʀ/) were used as control conditions, and was found to be equivalent in size (Experiment 2) whether the target words were of higher or lower frequency than the prime words (e.g. /ʀɔb/-/bɔʀ/ vs. /bɔʀ/-/ʀɔb/).

The motivation for the present work was not simply to provide an investigation of transposed-phoneme priming effects under cross-modal presentation, but more precisely to use cross-modal presentation as a means to examine the relative contribution of bottom-up phoneme-to-word excitatory influences and word-level inhibition on transposed-phoneme priming effects. Within the general framework of interactive-activation models of spoken word recognition (McClelland & Elman, 1986; Hannagan et al., 2013) there are two main components to phonological priming effects: phoneme-to-word facilitation driven by phonemes shared by prime and target, and word-level inhibition between phonologically similar words. We therefore reasoned that the fact that priming was only found with low-frequency primes and high-frequency targets in the Dufour and Grainger (2020) study was due to the fact that the main contribution to such effects was phoneme-to-word facilitation. We then noted that prior research suggests that lexical influences on priming are stronger than sublexical influences under conditions of cross-modal presentation (Radeau, 1995; Slowiaczek et al. 1992), and therefore reasoned that the overall direction of priming effects could change in cross-modal compared with intra-modal priming. This prediction was confirmed in Experiment 1 where we found inhibitory cross-modal transposed-phoneme priming effects with the same stimuli that had given rise to facilitatory effects under within-modality presentation.

We further examined the influence of prime-target relative frequency on the inhibition effect found under cross-modal presentation. Contrary to the facilitation found under unimodal presentation, the inhibition found under cross-modal presentation was not influenced by the relative frequency of primes and the targets. Here we provide a tentative explanation for the different impact of prime-target relative frequency on unimodal and cross-modal transposed-phoneme priming. Prior research using unimodal presentation of primes and targets has provided evidence that inhibitory priming effects occur when the targets are of higher frequency than the primes, but not when they are of lower frequency. For example, although only significant in the by-participant analysis, Radeau et al. (1995) reported an inhibitory effect when primes and targets shared their initial phonemes in a lower/higher frequency condition but no effect at all in a higher/lower frequency condition. Also, in an orthographic priming study in the visual modality with unmasked primes, Segui and Grainger (1990) reported a significant inhibitory priming effect when the targets were of higher frequency than the primes, but not when the targets were of lower frequency than the primes. Moreover, the study of Dufour and Grainger (2020) showed that unimodal faciliatory priming is greater when the targets are of higher frequency than the primes. Thus, inhibitory priming effects are enhanced with higher frequency targets (Segui and Grainger, 1990; Radeau et al., 1995), and this is also the case for facilitatory effects (Dufour & Grainger, 2020). One can then derive a tentative explanation of the null effect of prime-target relative frequency in the current Experiment 2 by considering: 1) that prime-target relative frequency has the same impact on phoneme-to-word facilitatory effects and word-level inhibitory effects, and 2) that cross-modal priming, although being dominated by inhibitory effects, is also partly driven by phoneme-to-word facilitation. Therefore, when both factors are at play, the effects of relative prime-target frequency are cancelled out, and the observed effects are determined by the relative weight of lexical vs. sublexical influences. For example, a priming effect of +10ms and -30ms (net effect = -20ms) with high frequency primes and low frequency targets becomes an effect of +20ms and -40ms (net effect = -20ms) with low frequency primes and high frequency targets. So, with word-level inhibition having a greater impact on cross-modal priming than phoneme-to-word facilitation, the result is equivalent inhibitory effects independently of relative prime-target frequency.

To sum-up, transposed-phoneme priming effects fit well with one of the core mechanisms of spoken word recognition, namely activation of multiple lexical candidates and subsequent competition between them via inhibitory mechanisms (see Allopenna et al., 1998). We found that simply changing the modality of presentation of target words changes the direction of transposed-priming effects. The facilitatory priming effect observed under unimodal presentation became an inhibitory priming effect under cross-modal presentation. We interpret the unimodal facilitation effect as mainly reflecting bottom-up activation of target words during prime word processing, and the cross-modal inhibition effect as mainly reflecting lexical competition. In a more general way, both the unimodal facilitation effect and the cross-modal inhibition effect suggest that position-independent phonemes are extracted pre-lexically, and are then used to contact lexical representations. In this way, words sharing all their phonemes with a given target word but in a different order are part of the set of activated lexical candidates, and thus are potential competitors of the target word. Finally, our study has important implications at a methodological level, since it suggests that a cross-modal procedure assigns more weight to lexical influences (word-level inhibition) relative to sublexical influences (here, phoneme-to-word facilitation) in determining net priming effects.

Acknowledgements

This research was supported by ERC grant 742141.

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., and 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., and 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. and Sarkar, D. (2007). lme4: Linear mixed-effects models using S4 classes, R package version 2.6, retrieved 20 November 2017 from http://lme4.r-forge.r-project.org/.

Bowers, J. S., Kazanina, N., & Andermane, N. (2016). Spoken word identification involves accessing position invariant phoneme representations. *Journal of Memory and Language*, *87*, 71–83.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Dufour, S., & Grainger, J. (2020). The influence of word frequency on the transposed-phoneme priming effect. *Attention, Perception, and Pyschophysics***,** *2060-2069*.

Dufour, S., & Grainger, J. (2019). Phoneme-order encoding during spoken word recognition: A priming investigation. *Cognitive Science*, *43,* 1-16.

Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience, 13,* 121–143.

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & K¨onig, P. (2017). LabVanced: A Unified JavaScript Framework for Online Studies. In *International Conference on Computational Social Science IC2S2, Cologne*.

Foster, E. D., & Deardorff, A. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association, 105*(2), 203–206.

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

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language,* 59, 434-446.

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.

Morey, R. D. (2015). Using the BayesFactor package version 0.9.2+. http://bayesfactorpcl.r-forge.r-project.org/.

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

Perea, M., & Duñabeitia, J. A., & Carreiras, M. (2008). Transposed-letter priming effects for close versus distant transpositions. *Experimental Psychology, 55*,397– 406.

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. Memory & Cognition, *31*, 829-841.

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*, 231–246.

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

Radeau, M. (1995). *Facilitatory and inhibitory phonological priming effects as a function of overlap position.* Abstracts of the 36th Annual Meeting of the Psychonomic Society, p. 59, Los Angeles, CA.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1297–1311.

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception & Performance, 16,* 65-76.

Slowiaczek, L. M., & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 1239-1250.

Spinelli, E., Segui, J., & Radeau, M. (2001). Phonological priming in spoken word recognition with bisyllabic targets. *Language and Cognitive Processes, 16,* 367–392.

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.

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.

Table 1. Characteristics of the prime and target words (mean values) in Experiment 1.

	Frequency[1]	Number of phonemes/letters[2]	Uniqueness point[3]	Duration[4]
Transposition condition				
Target words (/bɔl/)	94	4.36	-	-
Related primes (/lɔb/)	94	3	4	623
Control primes (/kõt/)	101	3	4	622
Vocalic condition				
Target words (/dyn/)	121	4.67	-	-
Related primes (/jyp/)	115	3	4	640
Control primes (/ʀiʃ/ /)	121	3	4	638

Note: [1] In number of occurrences per million. [2] phonemes for the primes, letters for the targets. [3] The phonemic position at which the auditory word primes can be reliably identified. [4]In milliseconds.

Table 2. Characteristics of the prime and target words (mean values) in Experiment 2.

	Frequency[1]	Number of phonemes/letters[2]	Uniqueness point[3]	Duration[4]	Phonetic similarity Initial/Final	Levenshtein distance
Higher-frequency targets						
Target words (/bɔʀ/)	678	4.19	-	-	-	-
Transposed primes (/ʀɔb/)	34	3	4	634	1.46/1.46	3.46
Control primes (/vɔl/)	33	3	4	632	1.69/1.69	3.38
Lower-frequency targets						
Target words (/ʀɔb/)	34	4.54	-	-		
Transposed primes (/bɔʀ/)	678	3	4	630	1.46/1.46	3.46
Control primes (/sɔm/)	631	3	4	629	1.50/1.35	3.23

Note: [1] In number of occurrences per million. [2] phonemes for the primes, letters for the targets. [3] The phonemic position at which the auditory word primes can be reliably identified. [4]In milliseconds. None of the differences between transposed primes and control primes in phonetic similarity or Levenshtein distance were significant (all *p*s > .20).

Table 3: Mean Reaction Times (in ms) and percentages of correct responses for the control and transposed primes in each relative prime-target frequency condition in Experiment 2.

	Control	Transposed	Priming effect
Lower-frequency targets			
RT	843	874	-31
Correct Responses	80	80	0
Higher-frequency targets			
RT	766	789	-23
Correct Responses	97	97	0

Appendix 1: Prime and target words used in Experiment 1

Transposition Condition

Control primes	**Related Primes**	**Target words**
rampe	lobe	bol
masse	robe	bord
jeune	cache	chaque
veille	chatte	tâche
juge	sec	caisse
singe	cale	lac
date	rèche	chair
pompe	chic	quiche
vigne	chope	poche
cure	lisse	cil
fiche	loque	col
bille	loupe	poule
jambe	toque	cotte
lampe	rade	dard
souche	digue	guide
comte	dire	ride
balle	douce	soude
mine	pouce	soupe
gaffe	rousse	sourd
base	dur	rude
vache	geine	neige
moule	jarre	rage
monde	jour	rouge
cours	laisse	sel
bonne	lame	mal
puce	râle	lard
tombe	mille	lime
chance	nulle	lune
linge	mare	rame
couche	tonne	note
page	rhume	mur
mouche	panne	nappe
seuil	nerf	reine
loge	niche	chine
banque	tape	patte
gamme	pente	tempe
pêche	quitte	tique
galle	top	pote
type	verre	rêve
bac	gîte	tige
fille	route	tour
mode	rate	tard
gare	butte	tube
bulle	coche	choc

Vocalic Condition

Control primes	**Related Primes**	**Target words**
mousse	laine	pelle
gaz	tank	bande
nord	cube	mule
chasse	nuque	lutte
riche	jupe	dune
pile	quinze	dinde
rite	terre	chèque
chef	salle	car
vase	chaîne	guerre
coq	cèpe	veine
sol	longue	bombe
bâche	vente	sens
fête	boeuf	soeur
pire	gomme	vote
fils	vol	tort
dame	port	botte
chute	ronce	sonde
sainte	duc	fugue
bec	four	louche
fouille	tasse	dalle
face	somme	mort
chaise	bouc	goutte
père	doute	bouche
ville	tête	mer
passe	dix	vite
femme	seul	peur
douche	natte	cape
touche	peigne	messe
coupe	pâle	sac
bave	lotte	roche

Appendix 2: Prime and target words used in Experiment 2

Higher-frequency target condition			Lower-frequency target condition		
Control primes	**Transposed primes**	**Target words**	**Control primes**	**Transposed primes**	**Target words**
vol	robe	bord	somme	bord	robe
vase	cache	chaque	femme	chaque	cache
pelle	rêche	chair	belle	chair	rêche
pile	ride	dire	ville	dire	ride
tasse	lame	mal	car	mal	lame
fiche	lime	mille	type	mille	lime
puce	lune	nul	dur	nul	lune
fugue	rhume	mur	sud	mur	rhume
bave	rate	tard	face	tard	rate
vache	casse	sac	dame	sac	casse
loge	roc	corps	bonne	corps	roc
natte	lasse	salle	date	salle	lasse
botte	rhum	mort	sol	mort	rhum
gaffe	râpe	part	bal	part	râpe
vote	rosse	sort	comme	sort	rosse
bulle	russe	sur	lutte	sur	russe
bouc	soute	tous	pour	tous	soute
dalle	chatte	tache	basse	tache	chatte
galle	rabe	barre	passe	barre	rabe
loque	chope	poche	nord	poche	chope
loupe	soude	douce	cours	douce	soude
bouche	rouge	jour	doute	jour	rouge
fête	laisse	sel	mer	sel	laisse
gaz	tape	patte	rare	patte	tape
neige	rêve	verre	même	verre	rêve
quiche	rite	tir	vide	tir	rite