When facilitation becomes inhibition: effects of modality and lexicality on transposed-phoneme priming

Sophie Dufour, Jonathan Mirault, Jonathan Grainger

When facilitation becomes inhibition:

Effects of modality and lexicality on transposed-phoneme 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: Unimodal and cross-modal 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: We examined the contributions of phoneme-to-word facilitation and word-to-word inhibition to transposed-phoneme priming effects under unimodal and cross-modal presentations. Experiments 1A and 1B showed that the presentation of an auditory prime formed by transposing two phonemes in a given target word facilitated lexical decisions to auditory targets. This facilitation was independent of the lexicality of the primes. In Experiment 2 the targets were presented visually rather than auditorily. We found an inhibitory priming effect, which, in contrast to Experiment 1, was influenced by the lexicality of the primes, with an effect emerging only with word primes. These findings point to a greater impact of phoneme-to-word facilitation under unimodal presentation and a greater role for word-to-word inhibition under cross-modal presentation. Hence, by simply manipulating the modality of target presentation, it is possible to separately probe two central mechanisms postulated in models of spoken word recognition, namely phoneme-to-word activation and lexical competition.

Keywords: Unimodal auditory priming; Cross-modal priming; Transposed-phoneme effects; Lexical competition.

How do speaker-hearers of a given language recognize the spoken words of that language? Most current models of spoken word recognition assume that several stages of processing are involved in, via which listeners succeed in mapping the speech signal onto the phonological representations of words stored in long-term memory. First, listeners extract from the speech signal units smaller than the words which may be, for instance, features (Gaskell & Marslen-Wilson, 1997), phonemes (Marslen-Wilson and Welsh, 1978; Norris, 1994), syllables (Mehler, 1981), features and phonemes (McClelland & Elman, 1986), or phonemes and diphones (Hannagan et al., 2013). These pre-lexical units are then mapped onto whole-word phonological representations that best match the unfolding speech signal, and it is further hypothesized that not only exact matches, but also partial matches constitute potential candidates for word recognition. In Interactive-Activation models like TRACE (McClelland & Elman, 1986), all co-activated lexical representations compete for selection as the best match for a given portion of the speech signal via a process of lateral inhibition. Another key characteristic of this class of model is the cascaded nature of processing. Thus, for example, as soon as phonemes become activated this activity is fed-forward to word representations in a continuous manner. The combination of cascaded processing, interactivity, and lateral inhibitory processes, means that different words can dominate processing at different time windows, and thus word recognition may be best characterized as a succession of activation peaks rather than a succession of discrete word identifications.

The phonological priming paradigm has been often used to investigate the various stages of spoken word recognition. In this paradigm, two stimuli, the prime and target, that share phonemes in the related prime condition, are presented in close temporal succession, and participants have to perform a task (e.g., lexical decision or shadowing) on the target. Of interest is the effect related auditory primes have relative to unrelated primes on the processing of the auditory target. In the vast literature on phonological priming, two opposing effects have been reported (see McQueen & Sereno, 2005; Norris et al., 2002 for a review). The first one is a facilitatory priming effect, that is faster responses on target words when preceded by a phonologically related prime than by an unrelated control prime, and the second one is an inhibitory priming effect, that is slower responses on target words when preceded by a phonologically related prime than by an unrelated control prime. Facilitatory effects have been interpreted as reflecting bottom-up facilitation operating between sublexical and lexical phonological representations, while inhibitory effects have been explained as reflecting competition/inhibition processes operating between co-activated lexical candidates.

One key issue in current research on spoken word recognition is how to tease apart the contributions of pre-lexical facilitation and lexical competition on phonological priming effects? One well-established method is to examine whether the phonological priming effects found with auditory primes and targets transfer under cross-modal presentation. The logic behind cross-modal presentation is that a priming effect which involves lexical representations should be modality independent and should be of similar magnitude when tested in the same (auditory) modality as across different modalities (see Dumay et al., 2001; Spinelli et al., 2001). In contrast, an effect which involves pre-lexical representations should not be observed when primes and targets are presented in different modalities, since in this case the sublexical units activated during prime processing are not the same as those used during target processing. Such a manipulation has already proved useful in examining the precise locus of phonological priming effects when primes and targets share either their initial or final phonemes. For example, the facilitation observed when primes and targets share some of their initial phonemes (Spinelli et al., 2001) and the inhibition found when primes and targets share all of their phonemes except the last one (e.g. DRESS-DREAD; see Dufour &

Peereman, 2003a; 2003b; Radeau et al., 1995; Slowiaczek and Hamburger, 1992) transfers under cross-modal presentation (Radeau, 1995; Slowiaczek and Hamburger, 1992; Spinelli et al., 2001), and these two effects have been respectively interpreted as reflecting lexical activation and competition between co-activated lexical candidates (e.g. Dufour & Peereman, 2003a; 2003b; Radeau et al., 1995; Slowiaczek and Hamburger, 1992; Spinelli et al., 2001). In contrast, the facilitation found when the primes and the targets share their final phonemes (e.g. LAMP-RAMP) does not transfer under cross-modal presentation (Dumay et al., 2001; Radeau, 1995; Spinelli et al., 2001) and has been largely taken as reflecting processes that occur before lexical access (e.g., Radeau et al., 1995; Spinelli et al., 2001), and in particular the repeated activation of the same pre-lexical units during prime and target processing.

Although the unimodal procedure is more widespread than the cross-modal procedure, the presentation of the primes and the targets in different modalities appears well suited to probe competition between lexical candidates. In fact, in a recent study, Dufour, Mirault and Grainger (2022) suggested that the cross-modal priming paradigm could be even better than the unimodal priming paradigm for capturing the competition process operating between co-activated lexical candidates (i.e., lexical inhibition). In several studies using the phonological priming paradigm, Dufour and Grainger (2019; 2020; 2022) observed that the prior presentation of an auditory prime word (e.g., /ʀɔb/ *ROBE "dress"*) facilitated the subsequent processing of an auditory transposed-phoneme target word (e.g., /bɔʀ/ *BORD* "edge"). This unimodal facilitatory transposed-phoneme priming effect was observed when both words (e.g., /ʀɔb/-/bɔʀ/; Dufour & Grainger, 2019, 2020) and nonwords (e.g., /biksɔt/-/biskɔt/ *BISCOTTE* "toasted bread"; Dufour & Grainger, 2022) were used as primes, and was found to be stronger when the target words had a higher frequency than the prime words (e.g., /ʀɔb/-/bɔʀ/) compared with when targets words had a lower frequency than primes (/bɔʀ/-/ʀɔb/;

Dufour & Grainger, 2020). This pattern of results was interpreted as reflecting partial activation of the transposed-phoneme target word during nonword or word prime processing, with more activation generated when words increase in frequency. In a follow-up study, Dufour et al. (2022) changed the modality of target presentation (from auditory to visual) and observed clear transposed priming effects, but in the opposite direction. Thus, the prior presentation of an auditory word prime slowed-down the subsequent processing of a visual transposed-phoneme target word.

At a theoretical level, Dufour et al. (2022) concluded that in accordance with interactive-activation models of spoken word recognition (Hannagan et al., 2013; McClelland & Elman, 1986), transposed-phoneme target words are not only activated during prime processing, but when the auditory prime is a word, a competition also occurs between the lexical representations of the prime and the transposed-phoneme target. As a result, the transposed-phoneme target word receives inhibition from the prime word, thus slowing-down its subsequent identification as a visual target. At a methodological level, the authors concluded that the cross-modal priming procedure could give more weight to word-to-word inhibition than to phoneme-to-word facilitation, and thus inhibitory influences are more likely to dominate than facilitatory influences under cross-modal presentation of the primes and targets. Of course, the corollary of this is that the unimodal priming procedure would give more weight to phoneme-to-word facilitation than to word-to-word inhibition, and thus facilitatory influences would be more likely to dominate than inhibitory influences under unimodal presentation of primes and targets.

Here, we tested this claim by directly comparing the influence of phoneme-to-word facilitation and word-to-word inhibition in transposed-phoneme priming under unimodal presentation (Experiments 1A & 1B) and under cross-modal presentation (Experiment 2). More specifically, in each experiment we manipulated the lexical status of the primes, which permitted us for the first time to compare the amount of priming caused by word and nonword primes on exactly the same target words. We reasoned as follows: If as suggested by Dufour et al. (2022) the unimodal priming procedure gives more weight to phoneme-to-word facilitation than to word-to-word inhibition, then we expect to observe a facilitatory transposed-phoneme priming effect when both words and nonwords are used as primes, since in both cases the presentation of the primes cause activation of the transposed-phoneme target words. In contrast, if as suggested by Dufour et al. (2022), the cross-modal priming procedure gives more weight to word-to-word inhibition than to phoneme-to-word facilitation, then we expect to observe an inhibitory transposed-phoneme priming effect only when words are used as primes since, because by definition, nonwords having no lexical representations, they cannot directly inhibit the lexical representation of the transposed-phoneme target. Also, we predicted that the facilitatory effect which is expected with nonword primes under unimodal presentation should be reduced under cross-modal presentation, since the latter procedure is hypothesized to give less weight to phoneme-to-word facilitation.

Experiment 1A

Experiment 1A used the unimodal priming procedure in which primes and targets were presented auditorily. As auditory-auditory presentation is hypothesized to give more weight to phoneme-to-word facilitation, we predicted that a facilitatory transposed-phoneme effect would be observed when both words and nonwords are used as primes.

Method

Participants: A total of 100 participants were recruited on-line via the Prolific platform (Palan & Schitter, 2018) and were paid 8 Euros for their participation. All participants self-reported to be native speakers of French and their reported age was between 18 and 61 years. 16 participants were over 40 years old, and only one was over 60.

Materials: Sixty monosyllabic target words with a CVC syllabic structure were selected from Vocolex, a lexical database for French (Dufour et al., 2002). Half of the target words were preceded by word primes, and the other half by nonword primes. In each condition of prime lexicality (words, nonwords), two primes were associated with each of target word. Transposed prime shared the same CVC phonemes as the target word, but with the two consonants in a different order (e.g. DIGUE /dig/ “sea wall” – GUIDE /gid/ “guide” for the word priming condition; PUGE /pyʒ/ - JUPE /ʒyp/ “skirt” for the nonword priming condition), and the corresponding control prime shared only the medial vowel with the target (e.g. fiche /fiʃ/ “sheet” - GUIDE /gid/ “guide” for the word priming condition; DUVE /dyv/ - JUPE /ʒyp/ “skirt” for the nonword priming condition). The target words in each of the prime lexicality conditions, as well as the transposed and control primes were matched in frequency (i.e., for word primes), number of phonemes and durations. The main characteristics of the primes and the targets are given in Table 1. The complete set of primes and targets are given in the Appendix.

<Insert Table 1 about here>

For each condition of prime lexicality, two experimental lists were created so that each of the 60 target words was preceded by the two types of prime (transposed, control), and participants were presented with each target word only once. For the purpose of the lexical decision task, 60 target nonwords were added to each list. So that the nonwords mimicked the words, 30 of them were paired with a word prime and the 30 other were paired with a nonword prime. Among the 30 word primes, half shared the same phonemes as the target nonword but in a different order (e.g. the word prime *BANQUE* /bãk/ "bank" and the target nonword /kãb/), and the other half shared only the medial vowel with the target nonword (e.g. the word prime *TIR* /tiʀ/ "shot" and the target nonword FIPE /fip/). Also, among the 30 nonword primes, half shared the same phonemes as the target nonword but in a different order (e.g. the nonword prime *VIME* /vim/ and the target nonword MIVE /miv/), and the other half shared only the medial vowel with the target nonword (e.g. the nonword prime *SUPE*/syp/ and the target nonword BUDE/byd/). In addition, 120 unrelated prime-target pairs having no phoneme in common 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 nonwords. Among the 60 filler target words, 30 were paired with a word prime and the 30 other target words were paired with a nonword prime. Also, among the 60 filler target nonwords, 30 were paired with a word prime and the 30 other target nonwords were paired with a nonword prime. All the targets with their primes 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 durations of the primes and the targets are given in Table 1.

<u>Procedure</u>: The experiment was programmed using LabVanced software (Finger et al., 2017). Participants were instructed to put on their headphones and adjust the volume to a comfortable sound level. The primes and the targets were presented auditorily, and an interval

(ISI) of 20 ms separated the offset of the primes and the onset of the targets. Participants were asked to make a lexical decision as quickly and accurately as possible by pressing the right arrow of their keyboard for the word response and the left arrow for the nonword response. Reaction Time (RT) recording 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

Eight participants were excluded from the analyses. Among them five participants had an error rate above 50%, and the three others had RTs greater than 2,000 ms on average. One target word used in the word prime condition that gave rise to an error rate of more than 40% was also removed. The mean RT and percentage of correct responses to target words in each condition are presented in Table 2.

<Insert Table 2 about here>

RTs to target words (available at https://osf.io/c589d/; Open Science Framework; Foster & Deardorff, 2017) were 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 473 data points out of 5428 (8.71%). Three extremely long RTs > 6000 ms and one short RT <300 ms were considered as outliers, and were excluded from the analysis. Moreover, in each condition RTs greater than three standard deviations above the mean were excluded from the analysis (1.90%). 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 4858 data points. We tested a model with the variables Prime Lexicality (words, nonwords), Prime Type (transposed, control) and their interaction entered as fixed effects. 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 Prime Lexicality failed to reach significance (b =-0.0424, SE = 0.0222, t = -1.92, p=.06). The main effect of Prime Type was significant (b =-0.0237, SE = 0.0050, t = -4.74, p<.001) with RTs on target words being faster when preceded by transposed primes in comparison to control primes. The interaction between Prime Type and Prime Lexicality was not significant (b =-0.0013, SE = 0.0010, t = -0.13, p>.20).

The percentage of correct responses was analyzed using a mixed-effects logit model (Jaeger, 2008) following the same procedure as for RTs. The main effect of Prime type was significant (b =0.2768, SE = 0.1018, z = 2.72, p<.01) with more correct responses in the transposed priming condition than in the control priming condition. No other effect was significant.

To sum-up, in line with our predictions, both nonword and word primes generated a facilitatory transposed-phoneme priming effect that was not influenced by the lexicality of the primes.

Experiment 1B

Because RTs in Experiment 1A were somewhat longer than those generally observed in laboratory experiments, we decided to conduct a replication of Experiment 1A using a different system for programming online experiments.

Method

Participants: A total of 100 participants were recruited on-line for the experiment via the Prolific platform (Palan & Schitter, 2018) and were paid 8 Euros for their participation. All participants self-reported to be native speakers of French. Their reported age was between 18 and 65 years. 8 participants were over 40 years old, and only one was over 60.

Materials and Procedure: The materials were the same as in Experiment 1A. The procedure was identical to Experiment 1A except that the experiment was programmed using html and php protocols and then uploaded onto the lab server.

Results and Discussion

Six participants were excluded from the analyses. Among them three participants had an error rate above 50% and the three others had RTs greater than 2000 ms on average. Three target words, one used in the nonword prime condition and the two others used in the word prime condition, that gave rise to an error rate of more than 40% were also removed. 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/c589d/) were analyzed using linear mixed effects models following the same procedure as in Experiment 1A. The RT analysis was performed on correct responses, thus removing 499 data points out of 5358 (9.31%). Seven extremely long RTs > 6000 ms and one short RT <500 ms were considered as outliers, and were excluded from the analysis. Moreover, in each condition RTs greater than three standard deviations above the mean were excluded from the analysis (2.68%). 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 4721 data points. We tested a model with the variables Prime Lexicality (words, nonwords), Prime Type (transposed, control) and their interaction entered as fixed effects. 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 Prime Lexicality was significant (*b* =-0.0477, *SE* = 0.0198, *t* = -2.40, *p*<.05) with RTs on target words being shorter when preceded by nonword primes in comparison to word primes. The main effect of Prime Type was significant (*b* =-0.0286, *SE* = 0.0054, *t* = -5.34, *p*<.001) with RTs on target words being faster when preceded by transposed primes in comparison to control primes. The interaction between Prime Type and Prime Lexicality was not significant (*b* =0.0109, *SE* = 0.0107, *t* = 1.01, *p*>.20).

The percentage of correct responses was analyzed using a mixed-effects logit model (Jaeger, 2008) following the same procedure as for RTs. The main effect of Prime Lexicality was significant (*b* =0.5624, *SE* = 0.2675, *z* = 2.10, *p*<.05) with more correct responses in the nonword prime condition than in the word prime condition. The main effect of Prime type was also significant (*b* =0.2812, *SE* = 0.0996, *z* = 2.82, *p*<.01) with more correct responses in the transposed condition than in the control condition. The interaction between Prime Type and Prime Lexicality was not significant (*b* =0.1389, *SE* = 0.1991, *z* = 0.70, *p*>.20).

In sum, we successfully replicated the results of Experiment 1A using a different procedure for programming online experiments. We again found a facilitatory transposed-phoneme priming effect under unimodal (auditory-auditory) presentation, that did not depend on prime lexicality[1]. Again, RTs are on average greater than 1100ms, and thus are

[1] Note that Dufour and Grainger (2022) did not observe transposed-phoneme effects with nonwords created by transposing nonadjacent phonemes in multisyllabic words (e.g. /ʃo**l**o**k**a/-/ ʃo**k**o**l**a/) in both a primed and unprimed

substantially longer than have been observed in previous auditory-auditory priming studies with monosyllabic words where average RTs typically lie between 700 and 1100 ms (see for example, Dufour & Grainger, 2019; 2020; Radeau et al., 1995; Luce et al., 2000). This is likely due to differences between in-lab experimentation and experiments run on-line. One advantage of on-line experimentation is that it enables the testing of participants from various backgrounds (not just psychology students for example) as well as being able to rapidly obtain sample sizes much larger than those typical of laboratory experiments. Thus, the greater variability in the participants that are tested, and perhaps the lower measurement precision associated with on-line experimentation might well be the source of the longer RTs.

Experiment 2

In Experiment 2, we used exactly the same materials as in Experiments 1A and 1B, but the target words were presented visually rather than auditorily. Because cross-modal presentation of the primes and targets is hypothesized to give more weight to word-to-word inhibition, we predicted that an inhibitory transposed-phoneme effect would be observed when words are used as primes. Also, because cross-modal presentation is hypothesized to give less weight to phoneme-to-word facilitation, we expected that the facilitation found in Experiment 1 with nonword primes would be reduced when the transposed-phoneme target words are visually presented.

---

lexical decision task. However, in a different study (Dufour et al., 2021) we reported clear transposed-phoneme effects in an unprimed lexical decision task with nonadjacent transpositions both when the transposed phonemes belonged to a different syllable (/ʃo**l**o**k**a/-/ ʃo**k**o**l**a/) and, as in the present study, to the same syllable (/bis.**t**ɔ**k**/-/bis.**k**ɔ**t**/). Post-hoc and correlation analysis in Dufour et al.'s (2021) study revealed that the discrepancy between the two studies for the /ʃo**l**o**k**a/ type of stimuli was due to differences in the overall speed of responding of participants, with the effects becoming greater with slower responses.

Participants: As in Experiments 1A and 1B, a total of 100 participants were recruited on-line for the experiment via the Prolific platform (Palan & Schitter, 2018) and were paid 8 Euros for their participation. All participants self-reported being native speakers of French and their reported ages were between 18 and 61 years. 10 participants were over 40 years old, and only one was over 60.

Materials: Exactly the same materials and the same experimental lists as in Experiments 1A & 1B were used here. To ensure that the transposed-phoneme effect observed with visual targets is driven by phonological overlap and not orthographic overlap, we calculated the orthographic similarity between our word primes and targets using the Orthographic Levenshtein Distance metric. The Levenshtein distance was 3.5 and 3.4 for the control and transposed-phoneme primes, respectively. Hence, using this metric, the transposed-phoneme primes are not more orthographically similar to target words than the control primes, and this was rendered possible because French has a highly irregular mapping from sound-to-spelling (e.g., the same sound /o/ can be spelt O, Ô, OT, AU, AUX, AUT, AUD, EAU). French is therefore an ideal language for operating this specific manipulation as well as other manipulations that require that orthography and phonology be varied independently (e.g., Ferrand & Grainger, 1992). 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 checked that the transposed phoneme primes were not transposed-letter neighbors of the target words. Among the 30 prime-target pairs only two were transposed-letter pairs.

Procedure: The same as in Experiments 1A and 1B except that the targets were now visually displayed at the center of the screen.

Results and Discussion

Six participants were excluded from the analyses. Among them three participants had an error rate above 50%, one reported to be more than seventy years old, and the other two had average RTs greater than 2000 ms. The mean RT and percentage of correct responses to target words in each condition are presented in Table 4.

<Insert Table 4 about here>

RTs to target words (available at https://osf.io/c589d/) were analyzed using linear mixed effects models following the same procedure as in Experiments 1A and 1B. The RT analysis was performed on correct responses, thus removing 373 data points out of 5640 (6.61%). Three extremely long RTs > 8000 ms were considered as outliers, and were excluded from the analysis. Moreover, in each condition RTs greater than three standard deviations above the mean were excluded from the analysis (1.96%). 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 5161 data points. We tested a model with the variables Prime Lexicality (words, nonwords), Prime Type (transposed, control) and their interaction entered as fixed effects. 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 Prime Lexicality was significant (b =-0.0624, SE = 0.0251, t = -2.49, p<.05) with RTs on target words being shorter when preceded by nonword primes in comparison to word primes. The main effect of Prime Type was significant (b =0.0171, SE = 0.0064, t = 2.66, p<.01) with RTs on target words being slower when preceded by transposed primes in comparison to control primes. The interaction between Prime Type and Prime Lexicality was significant (b =-0.0284, SE = 0.0128, t = -2.21, p<.05). This interaction was due to a significant priming effect emerging only in the word prime condition (b =0.0316, SE = 0.0097, t = 3.26, p<.01) but not in the nonword prime condition (b =0.0029, SE = 0.0084, t = 0.34, p>.20).

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 Prime Lexicality was significant (b =1.1815, SE = 0.1216, z =9.71; p<.001) with more correct responses in the nonword prime condition than in the word prime condition.

To sum-up, a clear inhibition effect was observed when the target words were presented visually rather than auditorily. Hence, by simply changing the modality of target presentation, we were able to reveal the hypothesized effects of lexical competition operating between prime words and transposed-phoneme target words. Moreover, as predicted, prime lexicality influenced cross-modal priming effects, with an inhibitory effect emerging with word primes but not with nonword primes. It should be noted that the inhibitory priming effect found with transposed-phoneme words does not modulate as a function of prime-target relative frequency (Dufour et al., 2022). In this sense, it clearly differs from the inhibitory priming effect observed in visual studies with unmasked primes that has been found to be stronger when the visually presented target words were of higher frequency that the visually presented prime words (Colombo, 1986; Lupker & Colombo, 1994; Segui & Grainger, 1990). Moreover, as indicated in the materials section, our related primes and targets i) do not involve transposed-letter words, ii) are not orthographic neighbors, iii) and perhaps more crucially, our control and related primes are well-matched in Orthographic Levenshtein Distance. Hence, we are confident that the inhibitory priming effect found with auditory primes and visual targets reflect processes involved during spoken word recognition, and in particular the inhibition of the transposed-phoneme target word during the auditory processing of the prime words.

General Discussion

On the basis of previous observations that auditory primes facilitate the processing of transposed-phoneme target words when presented auditorily (Dufour & Grainger, 2019, 2020, 2022), but inhibit target word processing when presented visually (Dufour et al., 2022), we hypothesized that the reason for these differences is that the cross-modal priming procedure gives more weight to word-to-word inhibition (lexical competition) while the unimodal priming procedure gives more weight to phoneme-to-word facilitation (lexical activation). In order to test this hypothesis, in the present study we manipulated the lexicality of primes and compared transposed-phoneme priming effects under unimodal and cross-modal priming procedures. In the present Experiments 1A and 1B primes and targets were presented auditorily (unimodal presentation) and we found a facilitatory transposed-phoneme priming effect that was independent of prime lexicality. In Experiment 2, in which primes were presented auditorily and the target visually, we found an inhibitory priming effect which, in contrast to Experiment 1, was found to be influenced by the lexicality of the primes, with an effect emerging only when primes were words.

The results of the present study therefore provide strong support for our explanation of the different pattern of transposed-word priming effects found in our prior work as a function of the modality of prime and target presentation (auditory-auditory vs. auditory-visual). That is, that unimodal presentation strengthens the influence of bottom-up phoneme-to-word facilitation effects relative to the effects of lexical competition, whereas cross-modal presentation has the opposite impact. The combined influence of modality of prime and target presentation plus the influence of prime lexicality (nonword vs. word) provides a complete account of the present findings. Under unimodal (auditory) presentation, bottom-up facilitatory influences dominate priming effects and prime lexicality therefore has little impact. Under cross-modal presentation, lexical competition dominates priming effects, with inhibition only emerging when primes are words.

Our account of the two distinct pattern of priming effects obtained under unimodal and cross-modal presentation is compatible with interactive-activation models according to which feed-forward activation (phoneme-to-word) and lateral inhibition (word-to-word) processes are two key mechanisms involved in spoken word recognition. Perhaps the most impressive aspect of the present results is that we have shown that these two key mechanisms can be probed separately with exactly the same stimuli simply by changing the modality of target presentation. As was predicted on the basis of our prior findings concerning transposed-phoneme priming effects, unimodal presentation accentuated the effects of feedforward facilitation, whereas cross-modal presentation enhanced the role of lateral inhibition. A possibility however is that at least part of the difference observed between unimodal and cross-modal priming was driven by differences in the ease of discrimination of the prime from the target in the two priming situations. Indeed, in the unimodal situation of the present study, the same speaker was used for the primes and the targets which could have encouraged participants to attend more carefully to the prime's phonemes in order to distinguish them from the targets. This might have promoted a reliance on phonological information more than in the cross-modal situation in which there is arguably less pressure to attend carefully to the prime's phonemes. Further unimodal priming studies could therefore investigate the impact of the use of same vs. different speakers for the primes and the targets on the magnitude and direction of the transposed-phoneme priming effect. The use of different speakers for the primes and the targets could indeed diminish the reliance on phonological information such that either a smaller facilitation effect than that observed in the present study or even an inhibitory effect might be observed. In the latter scenario, this would suggest that feedforward facilitation and lateral inhibition could also be probed separately under unimodal presentation of primes and the targets simply by changing the acoustic characteristics (i.e., speaker voice) of the primes. It should also be noted that inhibitory phonological priming effects have been observed in prior research using unimodal presentation and the same speaker for the primes and the targets. Such an inhibitory effect has been observed when related primes have a high initial phonemic overlap with targets, sharing all of their phonemes except the last one (e.g. DRESS-DREAD; Dufour & Peereman, 2003a, 2003b; Hamburger and Slowiaczek, 1996; Radeau et al., 1995). This suggests that initial phoneme overlap between prime words and target words creates stronger inhibition, most likely because the ambiguity between prime and target identity is only resolved at the very end of target processing. This is not the case for transposed-phoneme priming where the difference between prime and target words can be resolved earlier. The results of Toscano et al. (2013) are in accordance with this view. Using the visual world paradigm, these authors examined the eye movements of participants who followed spoken instructions to manipulate pictures of objects presented on a computer screen. They found more fixations on the picture representing a BUS than on a control picture (e.g., the picture of a WELL) when the spoken target was SUB, thus suggesting that SUB induces activation of BUS that shares all phonemes but in a different order. Importantly, they also observed that the picture representing the word SUN, sharing its initial phonemes with the target SUB, attracted even more fixations than the picture of a BUS, thus suggesting that initial phoneme overlap words interfere more with target word identification than transposed-phoneme words.

One particular model of spoken word recognition, the TISK model (Hannagan et al., 2013), provides a principled explanation for the lesser inhibitory impact of related transposed-phoneme words compared with initial phoneme overlap words. 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 combinations of contiguous and non-contiguous phonemes. Within such a framework, it is position-independent phoneme units that lead to activation of transposed words while diphone units that encode the order of phonemes allow the model to distinguish between transposed words. Now consider French short words like /lup/ LOUPE “magnifying glass” composed of the three phonemes units /l/, /u/, /p/ and of the three biphone units /lu/, /lp/, /up/, and its two competitors /pul/ POULE “chicken” and /luʃ/ LOUCHE “ladle”. During processing of /lup/, the transposed-phoneme competitor /pul/ receives activation from the position-independent phonemes (i.e. /l/, /u/ and

/p/) only, while the initial phoneme overlap competitor /luʃ/ receives activation from both the position-independent phoneme representations (i.e. /l/ and /u/) and the diphone representation (i.e. /lu/). Moreover, in TISK connection weights between diphone representations and whole-word representations are stronger than those between single phoneme representations and whole-word representations (see also You & Magnuson, 2018). Thus, initial phoneme overlap competitors receive more bottom-up support than transposed-phoneme competitors, hence the greater competitive impact of initial phoneme overlap words.

In sum, our study shows that by simply changing the modality of target presentation with auditory primes, we were able to reverse the direction of priming effects. Although prior research has shown that an effect found under unimodal presentation can disappear under cross-modal presentation, to our knowledge, this is the first time in spoken word recognition research that a priming effect reverses as a function of the modality of target presentation. Perhaps more crucially, with exactly the same materials, we were able to capture separately two key mechanisms postulated in current models of spoken word recognition, namely the bottom-up activation of lexical candidates and the subsequent competition that arises between co-activated lexical representations. On the basis of our results, we conclude that the unimodal priming paradigm is better suited than the cross-modal priming paradigm to probe bottom-up lexical activation, while the cross-modal priming paradigm is better suited than the unimodal priming paradigm to probe lexical competition.

References

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/.

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 226-234.

Dufour, S., & Grainger, J. (2022): When you hear /baksɛt/ do you think /*baskɛt*/? Evidence for transposed-phoneme effects with multi-syllabic words. *Journal of Experimental Psychology, Learning, Memory and Cognition, 48,* 98-107.

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.

Dufour, S., Mirault, J. & Grainger, J. (2022). Investigating the locus of transposed-phoneme effects using cross-modal priming. *Acta Psychologica, 226,* 103578.

Dufour, S., Mirault.J., & Grainger, J. (2021): Do you want /ʃoloka/ on a /bistɔk/? On the scope of transposed-phoneme effects with non-adjacent phonemes. *Psychonomic Bulletin and Review*, 28, 1668-1678.

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.

Dufour, S., & Peereman, R. (2003a). 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. (2003b). Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition, 88,* B33–B44.

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.

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.

Ferrand, L. & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *The Quarterly Journal of Experimental Psychology*, 45A, 353-372.

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.

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.

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

Hamburger, M., & Slowiaczek, L. M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review, 3,* 520–525.

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.

Luce, P. A., Goldinger, S. D., Auer, E. T., Jr., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*, *62*, 615– 625.

Lupker, S. J., & Colombo, L. (1994). Inhibitory effects in form priming: Evaluating a phonological competition explanation. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 437-451.

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.

McQueen, J. M., & Sereno, J. (2005). Cleaving automatic processes from strategic biases in phonological priming. *Memory & Cognition, 33,* 1185–1209.

Mehler, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society, Series B, 295,* 333–352.

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

Norris, D., McQueen, J. M., & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory & Cognition, 30,* 399-411.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

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, pp. 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 primes and the targets used in Experiments 1A, 1B and 2.

	Frequency[1]	Number of phonemes and letters [2]	Duration[3]
Word Prime condition			
Control	101	3	565
Transposed	144	3	564
Target word	96	3/4.53	579
Nonword Prime condition			
Control	-	3	561
Transposed	-	3	560
Target word	93	3/4.50	582

Note: [1] In number of occurrences per million. [2] phonemes for the control and related primes; phonemes /letters in Experiment 2 for the targets. [3]In milliseconds.

Table 2: Mean Reaction Times (in ms) and percentages of correct responses for the control and transposed primes in each prime lexicality condition in Experiment 1A (auditory primes and targets).

	Control	Transposed	Priming effect
Word Prime			
RT	1151	1127	24
Correct Responses	88	91	
Nonword Prime			
RT	1104	1081	23
Correct Responses	93	94	

Table 3: Mean Reaction Times (in ms) and percentages of correct responses for the control and transposed primes in each prime lexicality condition in Experiment 1B (auditory primes and targets).

	Control	Transposed	Priming effect
Word Prime			
RT	1227	1179	48
Correct Responses	87	89	
Nonword Prime			
RT	1164	1132	32
Correct Responses	92	94	

Table 4: Mean Reaction Times (in ms) and percentages of correct responses for the control and transposed primes in each prime lexicality condition in Experiment 2 (auditory primes and visual targets).

	Control	Transposed	Priming effect
Word Prime			
RT	744	783	-39
Correct Responses	90	90	
Nonword Prime			
RT	720	718	+2
Correct Responses	96	97	