Does a change in talker identity help listeners resolve lexical competition? Evidence from phonological priming.

Sophie Dufour, Amélie Dumon, Noël Nguyen

Aix Marseille Université, CNRS, LPL, UMR 7309, 13100 Aix-en-Provence, France
sophie.dufour@lpl-aix.fr, noel.nguyen@lpl-aix.fr

ABSTRACT

In this study, we examined whether the lexical competition process embraced by most of models of spoken word recognition is sensitive to talker variation. We used a long lag priming experiment in which primes and targets sharing all except the last phoneme (e.g. /bagaR/ "fight" vs /bagaʒ/ "luggage") were presented in two separate blocks of stimuli. Our results showed clear inhibitory priming effects with slower response times when target words were preceded by a phonologically related prime in comparison to a control prime. However, we also observed that the magnitude of the inhibitory priming effect was of the same magnitude whether the prime competitor and the target word were spoken by different talkers or by the same talker. The results are discussed in reference to other studies showing talker specificity-effects and in which access to talker-specific representations was demonstrated.

Keywords: Spoken word recognition, abstract representations, detailed representations.

INTRODUCTION

The traditional view of spoken word recognition assumes that as listeners attend to a spoken word, similar sounding words are activated and compete for recognition [6, 7]. For instance, the auditory input corresponding to the word “CAT” could lead to the partial activation of words such as “CAB”, “CAP” or “BAT” because they overlap with the target word “CAT” in some of their phonemes. Such a competition process is generally adopted in abstractionist models of spoken word recognition [7,10] which posit that each word in the mental lexicon is associated with an abstract phonological representation consisting of string of discrete units, for instance phonemes, which do not include details about how these words are pronounced. In this view, the listener first converts the speech signal into a sequence of discrete segments, removing all acoustic details that are deemed irrelevant to identification.

There is, however, a growing body of evidence showing that spoken word recognition is sensitive to fine grained acoustic details, and in particular to both within- and between- speaker variations in how a word is pronounced [5, 8]. A striking demonstration comes from studies using the long-term repetition priming paradigm [e.g. 8]. These studies showed that under some circumstances, the repetition priming effect – i.e. decrease in Reaction Times (RTs) when a word is encountered for a second time– is smaller when two different talkers are used between the first and second presentation. This talker specificity effect was taken as evidence in favor of exemplar-based models of the lexicon [5] which assume that all instances or exemplars of words are stored in memory with their acoustic details. As a result, an imperfect acoustic match between the first and the second presentation diminishes the repetition priming effect, because it is not the same form-based representation that is activated in memory.

In the present study, we examined whether the competition process embraced by most of models of spoken word recognition is sensitive to talker variation. In particular, we asked whether differences in talker identity help listeners disambiguate competing lexical representations. If this is indeed the case, such an observation should constitute a strong challenge for abstractionist models of spoken word recognition, since they should revise their assumptions in such a way that the lexical entries encode specific talker information, and that activation of lexical candidates be modulated by fine-grained acoustic details.

Supporting evidence for a competition process between similar sounding words comes from studies using the phonological priming paradigm [3, 12]. A review of all these studies indicates that the time taken to identify a target word is lengthened when this target word is preceded by a prime that has a high degree of phonemic overlap with the target word on its initial phonemes, compared with a phonologically-unrelated prime. Dufour and Peereman [3] showed that inhibition effects occurred provided that the primes and targets diverged on the last phoneme only. Thus, with French bisyllabic target words, they reported strong inhibition effects with primes and targets such as bagarre “fight” /bagaR/ - bagage “luggage” /bagaʒ/ sharing all phonemes except the last one. According to

Slowiaczek and Hamburger [12], inhibitory priming effects arise during the presentation of the target itself, rather than that of the prime. In case of a high phonemic overlap, the prime is re-activated until the last phonemes of the target are processed, and thus acts as a strong competitor of the target word, thereby slowing-down its recognition.

In the present study, we conducted a long lag priming experiment in which the prime competitors and the target words occurred in two separated blocks of stimuli [9]. Talker identity was manipulated in such a way that the competitor prime and the target were pronounced either by the same talker or by different talkers. We predicted that if talker characteristics are used by listeners to disambiguate the target word from its lexical competitor, smaller inhibitory priming effect should be observed when the prime and the target words are pronounced by different talkers. In such a case, the prime competitor starts to acoustically diverge from the target words as soon as the first phonemes of the target word are processed, and thus should not be strongly re-activated during target word processing. To maximize the involvement of talker-specific representations, we used a lexical-decision task in which nonwords were wordlike [8].

2. METHOD

2.1. Participants

Seventy-eight participants from Aix-Marseille University took part in the experiment for course credits. All were native speakers of French, and reported having no hearing or speech disorders.

2.2. Materials

Thirty monosyllabic target words with a CVCVC syllabic structure were selected from Vocolex, a lexical database of the French language [4]. For each target word, a CVCVC prime sharing all phonemes except the last one with the target was also selected (e.g. /bagaR/-/bagaʒ/). Because it has been shown that primes produce more inhibition when they are of lower frequency than the target word [11], the primes were selected in such a way that they were less frequent than the targets. Ten other CVCVC primes used as control and having no overlapping phoneme with the 30 target words were also selected. For the purpose of the lexical decision task, 70 CVCVC non-words were created by changing only the last phoneme of a real word (e.g. the word séjour “stay” /seʒuR/ became /seʒum/). This allowed us to have wordlike nonwords, and also to constrain the participants to listen to the stimuli up to the end prior to giving their response. The non-words followed the same criteria as the words, and thus consisted in 60 pairs that diverged only on the last phoneme (e.g. /seʒum/-/seʒud/). All the stimuli were recorded by both a male and 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 characteristics of the prime and the target words are given in Table1.

Table 1: Characteristics of the primes and the target words (mean values).

	Targets	Related primes	Control primes
Frequency[1]	41	4	6
Number of phonemes	5	5	5
Uniqueness point	6	6	6
Female speaker duration (in ms)	655	669	687
Male speaker duration (in ms)	707	712	736

1: in occurrence per million

2.3. Design

Two blocks of stimuli were presented. The first consisted of the primes and the second of the targets. Within each block, half of the stimuli were produced by the female speaker, and the other half by the male speaker. The primes block consisted of 60 stimuli: 30 words, 10 serving as talker match primes, 10 serving as talker mismatch primes, 10 serving as control primes and 30 nonwords. The targets block also consisted of 60 stimuli: 30 words and 30 nonwords. Among the 30 target words, 10 were phonologically related to the primes and were pronounced by the same talker (matched condition), 10 were phonologically related to the primes and were pronounced by the other talker (mismatched condition), and 10 were unrelated to the primes (control condition). Also for the nonword targets, 10 were phonologically related to the primes and were pronounced by the same talker, 10 were phonologically related to the nonword primes and were pronounced by the other talker, and 10 were unrelated to the primes.

Because each target was paired with three different primes (match, mismatch, control), and no participant was presented with the same target twice, three experimental lists were created. The three lists were then all divided in two sublists so that each stimulus was heard as produced by both the female and the male speaker.

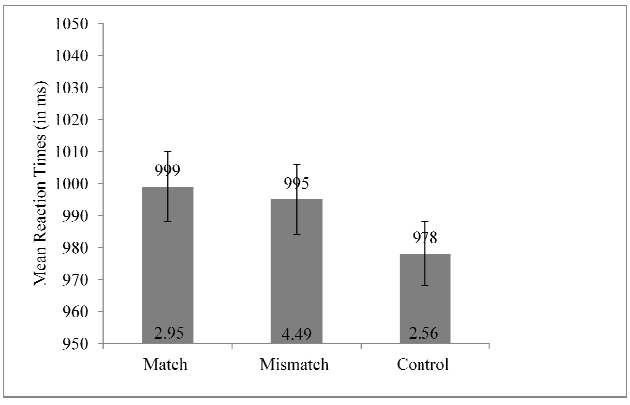
2.4. Procedure

The participants were tested in a sound-attenuated booth, and stimuli were presented over headphones at a comfortable sound level. Stimulus presentation and recording of the data were controlled by a PC running the E-Prime software (version 2.0, Psychology Software Tools). In both the primes and the targets blocks, participants were asked to make a lexical decision as quickly and accurately as possible with "word" responses using their dominant hand on an E-Prime response box that was placed in front of them. RTs were recorded from the onset of stimuli. Within each block, the stimuli were presented randomly. An intertrial interval of 2000 ms elapsed between the participant's response and the presentation of the next stimulus. A short break of 1 min separated the primes and the targets blocks. The participants were tested on only one experimental list and began the experiment with 12 practice trials.

3. RESULTS

Statistical analyses were performed on the target block. RTs greater than 1500 ms were excluded from the analyses. Also, for each participant and for each condition, RTs higher or lower than 2.5 standard deviations from the participant's mean were also removed. Adopting these criteria 2.92% of the data was rejected. Incorrect responses were also removed from the RT analyses. Analyses of variance (ANOVA) by participants ($F1$) and by items ($F2$) were conducted with prime type (talker match, talker mismatch, control) as variable. The mean RTs and error rates in each condition are presented in Figure 1.

Figure 1: Mean Reaction Times (in ms) as a function of prime type. Error rates (in %) are shown below the bar for each condition.



The main effect of prime type was significant both by participants ($F1(2,154) = 5.53$, $p < .01$) and by items ($F2(2, 58) = 3.63$, $p < .05$). Following our predictions, planned comparisons were performed to assess both competitor priming effects (match vs. control primes, and mismatch vs. control primes) and talker specificity effects (match vs. mismatch primes). Target words were responded to 21 ms slower when they were preceded by the talker match primes in comparison with the control primes ($F1(1,77) = 10.44$, $p < .01$; $F2(1,29) = 9.50$, $p < .01$). Also target words were responded to 17 ms slower when they were preceded by the talker mismatch primes in comparison with the control primes ($F1(1,77) = 6.16$, $p < .05$; $F2(1,29) = 4.68$, $p < .05$). Crucially, the comparison between talker match and mismatch primes revealed no specificity effect ($F1(1,77) = 0.32$, $p > .20$; $F2(1,29) = 0.12$, $p > .20$). No significant effect was found on Error Rates.

4. DISCUSSION

In this study, we asked whether or not listeners use talker–specific information to disambiguate between a target word and its lexical competitor. We predicted that if the competition process embraced by most models of spoken word recognition is sensitive to change in talker identity, smaller inhibitory priming effects would be observed when the target word and its competitor prime are spoken by different talkers. In contrast to our hypothesis, we showed that the magnitude of the inhibitory priming effect was not attenuated by talker differences. Hence, it appeared the competition process between similar sounding words is not sensitive to talker variation, at least in the specific setting of our experiment.

Our findings are not in accordance with those obtained by Creel, Aslin and Tanenhaus [1]. Using the head-mounted eye-tracking methodology, the authors examined the time-course of lexical activation in the face of talker variation. In their experiment, English competitor word pairs (e.g. couch-cows) were spoken by either a single or different talkers. Creel et al. [1] observed that after repeated instances of the word pairs, different-talker pairs showed smaller proportions of competitor fixations than same-talker pairs. Crucially, a closer look at the results of Creel et al. [1] revealed an attenuation of the competition effect for different-talker pairs in the last 20% of trials but not in the first 20% of trials. Note that in the Creel et al.' s [1] experiment, 10 blocks of stimuli were used and within each block, each competitor word pairs was presented twice. More specifically, 20 occurrences of a given word pronounced by the same talker were

presented over the course of the experiment. Hence the study of Creel et al. also indicates that several repetitions of the same word by the same voice may be necessary to encode talker-specific information in the lexical representations. In our experiment each prime and each target were presented only once. Thus, it could be that our failure to find a modulation of the competition effect when a change in talker identity was introduced results from the small number of repetitions, which has not permitted a full integration of the specific characteristics of each talker in the listeners' mental lexicon.

What the present study reveals is that talker-specificity effects are not so easy to capture. As we have just discussed, Creel et al.'s study [1] revealed that several repetitions of the same word may be required, at least when we study lexical competition processes. In respect to repetition priming effects, McClennan and Luce [8] showed that a change in talker identity diminishes the magnitude of the repetition effect in the lexical decision task when the nonwords were wordlike, but not when the nonwords were unwordlike and, in the shadowing task, when participants had to wait for a response cue to repeat words (delayed shadowing task), but not when participants had to repeat words immediately upon hearing them. In a follow-up study using a lexical decision task with wordlike nonwords, Dufour and Nguyen [2] manipulated the frequency of the words and showed that talker-specificity effects emerged when words were of low frequency but not when they were of high frequency. Altogether, repetition priming studies indicate that access to detailed representations occurs when processing is slowed down, either by using wordlike nonwords [8], or by using a long delay between word presentation and the participant's response [8]or yet by using low-frequency words [2]. As for the repetition priming effect, the precise circumstances under which inhibitory phonological priming effect could be affected by a change in talker identity should be more precisely examined.

To sum up, our study points out the difficulty in demonstrating talker-specificity effects during on-line spoken word recognition. Such a difficulty is corroborated by others studies [1, 2, 8] that showed that talker specificity effects emerge under some precise manipulations. As we have just discussed, Creel et al.'s [1] study indicated that talker-specific information affects the competition process between phonologically similar words, only after repeated instances of the same words by the same voice. Thus, we have identified one potential factor – the number of repetitions of a particular instance– which could be responsible for our failure in the emergence of a talker-specificity effect. What nonetheless shows our study is that abstract representations exist and are preferentially accessed during spoken word recognition. Hence, the precise circumstances under which detailed specific representations are susceptible to be accessed and to affect lexical competition process remain to be determined.

5. REFERENCES

[1] Creel, S. C., Aslin, R. N., Tanenhaus, M. K. 2008. Heeding the voice of experience: The role of talker variation in lexical access. *Cognition* 106, 633–664.

[2] Dufour, S., Nguyen, N. 2014. Access to talker-specific representations is dependent on word frequency. *Journal of Cognitive Psychology* 26, 256-262.

[3] Dufour, S., Peereman, R. 2003. Lexical Competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory and Cognition*, 31, 1271-1283.

[4] 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.

[5] Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251–279.

[6] Luce, P. A., Pisoni, D. B. 1998. Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing* 19, 1–36.

[7] McClelland, J. L., Elman, J. L. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18, 1–86.

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

[9] Monsell, S., Hirsh, K.W. 1998. Competitor priming in spoken word recognition. *Journal of Experimental Psychology*: *Learning, Memory & Cognition* 24, 1495-1520.

[10] Norris, D. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52, 189–234.

[11] Radeau, M., Morais, J., Segui, J. 1995. Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance* 21, 1297-1311.

[12] 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.