



Electrophysiological evidence for benefits of imitation during the processing of spoken words embedded in sentential contexts

Angèle Brunellière¹, Sophie Dufour²

¹ Unité de Recherche en Sciences cognitive et affectives, Université de Lille3, France

² Aix Marseille Université, CNRS, LPL, UMR 7309, 13100 Aix-en-Provence, France

angele.brunelliere@univ-lille3.fr, sophie.dufour@lpl-aix.fr

Abstract

This event-related potential study examined the impact of imitating an unfamiliar accent on the processing of spoken words embedded in sentential contexts produced in that accent. The cloze probability effect in two groups of southern French speakers after they had to either listen to or imitate sentences spoken by a Belgian French speaker was tested. Speakers who did not imitate the unfamiliar accent showed a cloze probability effect on the phonological N200 wave, while those who did imitate the accent showed no effect on this component. Over a later time window, both groups showed a cloze probability effect on the N400, which is associated with lexical and semantic processing. Taken together, these results give clear evidence for processing benefits from the imitation of speech patterns, particularly at an acoustic/phonological level of processing.

Index Terms: speech imitation, cloze probability effect, event-related potentials.

1. Introduction

Talkers are known to imitate one another in many different ways during conversational exchanges, including body postures [1], head movements, and facial expressions [2]. Likewise, at the phonetic level, increases in similarity between talkers in vocal intensity [3], speech rate [4] and the phonetic characteristics of spoken words [5] have been observed. At a higher linguistic level, it has also been shown that talkers tend to repeat words and grammatical constructions produced by their interlocutors [6, 7]. Interestingly, talkers have also been found to spontaneously imitate the phonetic properties of spoken words even if they are presented in a non-interactive setting [8, 9]. In the present study, we investigated whether imitation of speech patterns improves the processing of spoken words embedded in sentential contexts.

This investigation is motivated by theoretical frameworks that postulate a close coupling between sensory/perceptual and motor systems [10, 11]. In particular, Wilson and Knoblich [11] suggested that imitative motor activation generates predictions of the unfolding action, thereafter making perceptual processing of the same action easier. Following this theoretical framework, Pickering and Garrod [12] suggested that linguistic imitation facilitates the interpretation of the interlocutor's message, particularly in the case of ambiguous or distorted input. According to their argument, this facilitation should occur because during language processing, the comprehension system uses the production system to predict what is coming next. In particular, on this view the language production system acts as an "emulator" during language comprehension, making predictions about the next

input at different linguistic levels (e.g. phonology, syntax and semantics). A recent study by Adank, Hagoort and Bekkering [13] provided supporting evidence for this view, showing that imitating a novel accent improves language comprehension under adverse listening conditions. The authors exposed Dutch-speaking participants to a novel accent under different conditions during a training phase, and assessed comprehension of the accent before and after training by measuring the signal-to-noise ratio at which listeners were able to repeat 50 % of the key words in a sentence. The novel accent was a variant of Dutch wherein particular vowels in words were switched around, such that the accent was unfamiliar to the native Dutch participants. Comprehension of accented speech improved only in participants who imitated the accent. In participants who only had to listen to the accented sentences, transcribe them as they listened, or repeat the accented sentences in their own accent, no improvement in comprehension was observed. Taken together, these results suggest that improvement in language comprehension results from vocal imitation *per se*, and not from other factors such as speaking out loud or paying attention to the phonetic/phonological properties of accented speech, since no improvement was observed in either the repetition group or the transcription group.

The aim of this study is to provide new evidence of the facilitatory role of imitation in spoken language comprehension. In contrast to Adank et al. [13], we examined the impact of an imitative behavior on language comprehension in the absence of noise. The motivation for this choice was the possibility that hearing degraded sentences forces participants to use subtle cues related to the voice of the speakers they are asked to imitate. In particular, we examined the processing of words embedded in a sentential context and spoken in an unfamiliar accent after participants have been asked to imitate this accent. As Adank et al. [13] pointed out, imitation of sentences spoken in an unfamiliar accent is a good way to test for the impact of imitation on spoken language comprehension, because accented speech contains phonetic and phonological variations which create ambiguities that listeners must resolve. Following the framework of Pickering and Garrod [12], speakers may be expected to be better at anticipating the phonetic and phonological cues of the unfamiliar accent after imitation. Facilitation during phonological decoding should be observed as a result. Moreover, because of this early-stage facilitation of spoken word processing, facilitation at a later lexico-semantic level of processing should also occur. The recording of event-related potentials (ERPs) seems particularly suitable to characterize the precise level of processing at which imitation exercises its facilitatory effect. Because of its high millisecond-level temporal resolution, this technique provides a sensitive real-

time measure of word processing. Indeed, one value of the ERP technique is that specific electrophysiological components have been associated with distinct stages of processing. For example, the N400 is thought to reflect lexical and semantic processing [e.g. 14], whereas the N200 is considered to be associated with acoustic and phonological processing [e.g. 15].

We chose to measure spoken-language comprehension using the cloze probability effect, which reflects the ease with which a word embedded in a sentential context is recognized. In a classical cloze procedure, participants are instructed to complete a given sentence fragment with the first word that comes to mind. A word's cloze probability is related to the proportion of individuals who provide it to complete a given sentence fragment, and is used as a measure of the level of predictability of the word. The first behavioral studies on this effect showed that written words with a high cloze probability are recognized faster than those with a low cloze probability [16, 17]. With the ERP technique, the cloze probability effect has been shown to affect the N400 component, a centro-parietal negative wave, peaking at 400 ms after word onset, often used to study lexical access and sense-making processes during sentence comprehension (see [14] for a review). In particular, Kutas and Hillyard [18] found that the amplitude of the N400 was higher for written words with low cloze probability than for ones with high cloze probability. The authors interpreted the reduced N400 amplitude after high cloze probability words as reflecting the meaning-building effect of the prior context, which makes the processing of later words that fit the context easier. In spoken sentence comprehension, cloze probability has often been reported to affect the N400, but also an earlier electrophysiological component, the N200 [15, 19]. Connolly and colleagues [15, 19] observed that the amplitude of N200 was higher for words with low cloze probability than for words with high cloze probability. This negative shift was interpreted as reflecting the ease with which a word is decoded at the phonological level.

In the present study, southern French speakers were exposed to an unfamiliar Belgian French accent. In the first of two phases, one group was instructed to listen carefully to sentences spoken by a Belgian French speaker, while the other was asked to imitate the same sentences produced by the same Belgian speaker. Because we were particularly interested in the precise level of processing at which imitation has a facilitatory effect, the participants were explicitly instructed to imitate the unfamiliar accent. Moreover, as mentioned above, Adank et al. [13] observed no improvement in language comprehension when participants repeated sentences spoken in an unfamiliar accent without imitating it. In the second phase, the two groups were instructed to listen to different sentences from the first phase but spoken by the same Belgian French speaker. In this second phase we recorded event-related potentials time-locked to the onset of the final words, which had either high or low cloze probability, for each group of participants.

Following the framework of Pickering and Garrod [12], we predicted that the more difficult it is to predict the occurrence of a word based on the preceding context (i.e. low cloze probability), the more benefits should result from imitation, due to improved analysis of acoustic/phonological cues in the unfamiliar accent. Because the N200 has been shown to be related to acoustic and phonological processing,

we predicted a smaller cloze probability effect on the N200 for participants who imitated the accent. Moreover, because the benefit from analysis of the acoustic/phonological cues in the unfamiliar accent should carry over into later stages of word processing, we expected to find better word processing at the lexico-semantic level as well. As a result, we predicted that the cloze probability effect on the N400 would be also affected by imitation.

2. Method

2.1. Participants

Twenty-four French-speaking students (four men) from the University of Provence, between the ages of 18 and 25 years, participated in the experiment. All were right-handed (handedness assessed by the Edinburgh Inventory) and had no auditory or language impairments. Participants declared that they had never been exposed to the Belgian French accent before the experiment. They gave written informed consent before participating in the experiment.

2.2. Materials

To construct sentences ending in words of differing cloze probability, a completion task on written sentence fragments was conducted with 17 native French speakers. This task allowed us to estimate the cloze probability of the sentence-ending words. None of the speakers who performed this task participated in the experiment. Two hundred sentence fragments were divided into two lists, each of which was completed by all the participants. Participants were instructed to read each sentence fragment and write the first word that they expect to complete the sentence fragment. Sentences with a final word whose cloze value was greater than 50% were classified as high-cloze sentences, those whose most frequent final word had a cloze value between 25 and 50% as middle-cloze sentences, and those whose most frequent final word had a cloze value lower than 25% as low-cloze sentences. For the experimental stimuli in the second phase we selected fifty high-cloze sentences and 50 low-cloze sentences. High- and low-cloze sentences were matched for length (mean: 9.1 words for high-cloze, 8.5 for low-cloze; $F(1,98)=1.62$, $p>0.20$). The final words in the high- and low-cloze sentences were also matched for frequency (taken from Lexique [20] mean: 80 for high-cloze, 60 for low-cloze; $F(1,98)=0.64$, $p>0.20$), and number of phonemes (mean: 5 phonemes for high-cloze, 5.2 for low-cloze; $F(1,98)=0.19$, $p>0.20$). Fifty middle-cloze sentences were also selected and were used as fillers in the second phase. The remaining sentences were used for the first session of the experiment and the practice stimuli. An example of the experimental sentences (high-cloze and low-cloze) presented during the second phase is given in Table 1.

All sentences were produced by a Belgian French-speaking man and were digitized at a sampling rate of 44 kHz with 16-bit analog-to-digital conversion. The speaker was asked to pronounce the sentences with natural prosody. The mean total duration of the high- and low-cloze sentences used in the second phase was 2182 and 2148 ms respectively. The mean duration of the final words was 476 ms for the high-cloze sentences and 474 ms for the low-cloze ones. To calculate word duration, a trained French native speaker identified the onset and the offset of each final word.

Table 1. Examples of experimental stimuli

| <i>Cloze probability</i> | <i>Example sentences</i> |
|--------------------------|---|
| High | Le nouveau journal télévisé a récolté une grande audience. <i>The new television news program has gained a large audience.</i> |
| Low | Tu vas prendre le commandement de notre équipe. <i>You are going to become the head of our team.</i> |

2.3. Procedure

The experiment was divided into two phases, in the course of which sentences spoken in an unfamiliar accent were presented binaurally at a comfortable sound level via headphones. In the first phase, half of the participants were asked to imitate the sentences (imitation), while the other half were asked to carefully listen to them (perceptual exposure). To promote and optimize imitative behavior, the participants were given visual feedback after each sentence indicating whether they had spoken more slowly or more quickly than the Belgian speaker. The first phase began with four practice sentences. Fifteen auditory sentences were then repeated twice, in two separate blocks. The sentences were separated by an inter-trial interval of 3000 ms. In the second phase, the two groups of participants were asked to listen to 150 sentences separated into five blocks. Each experimental block was composed of thirty sentences (10 high cloze, 10 middle cloze, 10 low cloze) presented in random order. To minimize blinks and eye and muscle movements, each trial began with the presentation of a fixation point at the middle of the screen. It remained there for 500 ms before the onset of the auditory sentence and stayed on the screen until the offset of the sentence. At the end of each block, participants were asked to answer true-false comprehension questions. The questions were presented for one third of sentences. This allowed us to ensure that the participants had paid attention to the sentences. The participants began the second phase with 6 practice sentences, including two with high, two with middle, and two with low cloze probability.

2.4. EEG recording

The EEG signal was recorded in a silent room during the second phase of the experiment with a 64-channel BioSemi ActiveTwo AD-box. The signal was sampled at 1024 Hz and referenced to the CMS-DRL ground during recording. Individual electrodes were adjusted to yield a stable offset (running average voltage at each electrode relative to the common-mode voltage) below 20 mV. EEG epochs, starting 100 ms before the onset of the final word of the sentence and lasting for 800 ms afterward, were averaged for each participant and each level of cloze probability. An artefact rejection criterion of $\pm 100 \mu V$ was applied to determine acceptance of EEG epochs at all electrodes, and to exclude any muscular and ocular artefacts. The number of accepted EEG epochs did not differ across the four experimental conditions (Imitation - Low cloze: 46.1, High cloze: 45; Perceptual exposure - Low cloze: 47.8, High cloze: 46.4). A bandpass filter of 1-30 Hz was applied and a 100-ms pre-stimulus

baseline was subtracted. Bad channels were interpolated for each participant [21] and the average reference was computed offline and applied to the EEG signal.

3. Results

We performed two-way analyses of variance (ANOVA) on the amplitude of ERP responses with group (perceptual exposure vs. imitation) and cloze probability (Low vs. High cloze) as factors. Because the two ERP components of interest (N200 and N400) have a centroparietal scalp distribution [see 14, 15], and because the scalp distribution of the two ERP components was shown to be similar [22], statistical analyses on the mean amplitude were performed for centroparietal recording sites using 6 electrodes (CP1, CPz, CP2, P1, Pz, P2)¹. Grand-average ERP responses for low and high cloze probability words and for each group at Pz are displayed in Figure 1A. Pz was chosen for display because it is the centroparietal electrode where the cloze probability effect is usually the largest. Visual inspection of grand-average ERP responses at Pz revealed the two expected components associated with the cloze probability effect: one early negative wave occurring around 200 ms, the N200, and a later large negative wave, the N400 (see Figure 1A). The amplitude of each component at centroparietal recording sites was measured for two time windows as follows: 200-245 ms (N200) and 250-400 ms (N400). In keeping with common practice, analysis for the N400 component was conducted for a large time window [23]. Additional analyses with *t*-tests revealed no later N400 time windows with significant differences. Mean average values on centroparietal recording sites in each condition are shown in Figure 1B.

3.1. Comprehension test

All participants paid careful attention to the spoken sentences, as attested by their good performance on the comprehension questions (85% after perceptual exposure and 83% after imitation). The two groups were not found to differ significantly ($F(1,22)=0.20$, $p>0.2$).

3.2. ERP results

As seen in Figures 1A and 1B, the group with perceptual exposure to the unfamiliar accent showed greater amplitudes for the low-cloze words in comparison to the high-cloze words on both the N200 and N400 components. On the contrary, the imitator group showed greater amplitudes for the low-cloze words in comparison to the high-cloze words only on the N400 component. Note that because the onset of final words was not preceded by a pause, the N100 wave could not be clearly distinguished.

3.2.1. The N200 time window: 200-245 ms

In this time window, a significant cloze probability \times group interaction ($F(1,22)=4.88$, $p<0.05$) was observed. While a cloze probability effect was found for the group with perceptual exposure to the unfamiliar accent ($F(1,22)=5.36$, $p<0.05$), there was no effect of cloze probability in the imitator group ($F(1,22)=0.65$, $p>0.2$). As seen in Figures 1A and 1B, only the perceptual-exposure group showed a higher-

¹ Note that no significant difference was found at any other sites.

amplitude N200 component for low cloze probability words than for high cloze probability words. A significant difference between the imitation and perceptual groups was also found on the low cloze probability words ($F(1,22)=6.46$, $p<0.05$) but not on the high cloze probability words ($F(1,22)=0.55$, $p>0.2$).

3.2.2. The N400 time window: 250-400 ms

For the N400 component, only a main effect of cloze probability ($F(1,22)=14.55$, $p<0.001$) was observed. The amplitude of the N400 was greater for the low cloze probability words than for the high cloze probability words (see Figures 1A and 1B). Contrary to what was observed for the N200 component, the effect of cloze probability did not interact with group ($F(1,22)=2.24$, $p=0.18$).

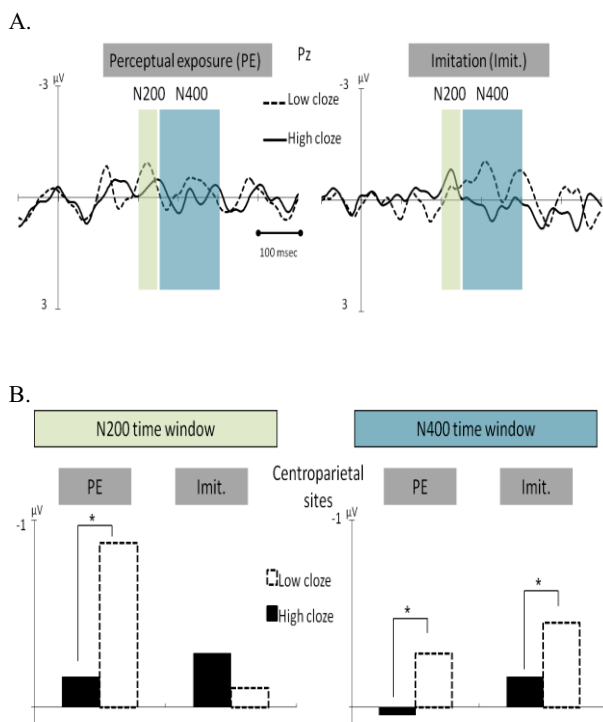


Figure 1. A: Grand-average ERP time-locked to the onset of final words at Pz for each level of cloze probability (dashed line: low cloze; solid line: high cloze) within each group of participants (Left side: perceptual exposure group; Right side: imitation group) for a time window running from -100 to 600 ms relative to final word onset. Scalp voltages are in μV .

B: Values of grand-average ERP time-locked to the onset of final words at centroparietal sites for each level of cloze probability (black: low cloze; white: high cloze) within each group of participants (Left side: perceptual exposure group; Right side: imitation group) for the time windows of the components of interest (200-245 ms for N200; 250-400 ms for N400). Scalp voltages are in μV .

4. Discussion

In this study we examined the impact of imitating an unfamiliar accent on the processing of spoken words embedded in a sentence produced in that accent. To this end, the changes in neuronal activity associated with the cloze probability effect were compared in two groups of listeners

after they either listened to or imitated the unfamiliar accent. ERP differences between the high- and low-cloze words were found on both the N200 and the N400 components in the group that did not imitate. In the group that did imitation, the word's cloze probability affected the N400 but not the N200 component.

What was the impact of imitation? In contrast to the group with only perceptual exposure to an unfamiliar accent, in the group who imitated the accent no difference in the N200 component between high- and low-cloze words was observed. As the N200 is generally interpreted as reflecting acoustic/phonological processes [15, 19], the absence of a cloze probability effect in the imitator group seems to indicate that imitation facilitates the processing of the acoustic/phonological properties of words spoken in an unfamiliar accent. In line with our predictions, the cloze probability \times group interaction also showed that the benefit of an imitative behavior was stronger for low-cloze words that is, for words that are more difficult to predict from the preceding context. One possible way of accounting for the null effect of imitation on high cloze probability words is a ceiling effect. These words may already be so highly expected by the preceding context that imitative behavior creates no detectable additional benefit at the acoustic/phonological level of processing.

The observed facilitation of acoustic/phonological processing with low-cloze words was not followed, however, by a benefit at the following stage of processing. Indeed, as for the group with perceptual exposure only, the N400 component in the imitation group was affected by the word's cloze probability, with greater amplitude for low-cloze than for high-cloze words. It thus appears that the improvement in acoustic/phonological processing of low cloze probability words due to imitation did not overcome the fact that these words do not fit well with the preceding context. Indeed, no facilitation of the N400 component was observed with imitation.

In conclusion, our findings provide evidence that imitation leads to more efficient processing of the acoustic/phonological properties of words produced in an unfamiliar accent. Following the theoretical perspective of Pickering and Garrod [12], we interpret this facilitation as being due to better prediction of the acoustic/phonological properties of words by the production system. Finally, in accordance with our predictions, imitation appears to specifically facilitate the processing of words that are not easily predictable from the preceding sentential context. Hence, the facilitatory influence of imitation reported here appears to fit with models that postulate a strong coupling between sensory/perceptual and motor systems [10, 11, see also, 12].

5. Acknowledgements

This work was supported by grant ANR-08-BLAN-0276-01 from the Agence Nationale de la Recherche (France). ERP analyses were performed with the Cartool software (supported by the Center for Biomedical Imaging of Geneva and Lausanne). We would like to thank Noël Nguyen for interesting and relevant discussions on this research, and Thierry Legou for his help in programming the "feed-back" of the Experiment.

6. References

- [1] Shockley, K., Santana, M.V. and Fowler, C.A., "Mutual interpersonal postural constraints are involved in cooperative conversation", *Journal of Experimental Psychology: Human Perception and Performance*, 29: 326-332, 2003.
- [2] Chartrand, T.L. and Bargh, J.A., "The Chameleon Effect: The Perception-Behavior Link and Social Interaction", *Journal of Personality and Social Psychology*, 76: 893-910, 1999.
- [3] Natale, M., "Convergence of mean vocal intensity in dyadic communication as a function of social desirability", *Journal of Personality and Social Psychology*, 32: 790-804, 1975.
- [4] Giles, H., Coupland, N. and Coupland, J., "Accommodation theory: communication, context and consequences", in H. Giles, J. Coupland and N. Coupland [Eds.], *Contexts of accommodation: Development in applied sociolinguistics*, 1-68, Cambridge University Press, 1991.
- [5] Pardo, J.S., "On phonetic convergence during conversational interaction", *Journal of the Acoustical Society of America*, 119: 2382-2393, 2006.
- [6] Branigan, H.P., Pickering, M.J. and Cleland, A.A., "Syntactic coordination in dialogue", *Cognition*, 75: B13-B25, 2000.
- [7] Garrod, S. and Anderson, A., "Saying what you mean in dialogue: a study in conceptual and semantic co-ordination", *Cognition*, 27: 181-218, 1987.
- [8] Goldinger, S.D., "Echoes of echoes? An episodic theory of lexical access", *Psychological Review*, 105: 251-279, 1998.
- [9] Shockley, K., Sabadini, L. and Fowler, C.A., "Imitation in shadowing words", *Perception and Psychophysics*, 66: 422-429, 2004.
- [10] Liberman, A.M. and Mattinly, I.G., "The motor theory of speech perception revisited", *Cognition*, 21: 1-36, 1985.
- [11] Wilson, M. and Knoblich, G., "The case for motor involvement in perceiving conspecifics", *Psychological Bulletin*, 131: 460-473, 2005.
- [12] Pickering, M.J. and Garrod, S., "Do people use language production to make predictions during comprehension?" *Trends in Cognitive Sciences*, 11: 105-110, 2007.
- [13] Adank, P., Hagoort P. and Bekkering H., "Imitation improves language comprehension", *Psychological Science*, 21: 1903-1909, 2010.
- [14] Kutas, M. and Federmeier, K.D., "Electrophysiology reveals semantic memory use in language comprehension", *Trends in Cognitive Sciences*, 12: 463-470, 2000.
- [15] Connolly, J.F., Phillips, N.A., Stewart, S.H. and Brake, W.G., "Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences", *Brain and Language*, 43: 1-18, 1992.
- [16] Erlich, S. and Rayner, K., "Contextual effects on word perception and eye movements during reading", *Journal of Verbal Learning and Verbal Behavior*, 20: 641-655, 1981.
- [17] Schwanenflugel, P.J. and Shoben, E.J., "The influence of sentence constraint on the scope of facilitation for upcoming words", *Journal of Memory and Language*, 24: 232-252, 1985.
- [18] Kutas, M. and Hillyard, S.A., "Brain potentials during reading reflect word expectancy and semantic association", *Nature*, 307:161-163, 1984.
- [19] Connolly, J.F., Stewart, S.H. and Phillips, N.A., "The effects of processing requirements on neurophysiological responses to spoken sentences", *Brain and Language*, 39: 302-318, 1990.
- [20] New, B., Pallier, C., Ferrand, L. and Matos. R., "Une base de données lexicales du français contemporain sur internet: LEXIQUE", *L'Année Psychologique*, 101: 447-462, 2001.
- [21] Perrin, F., Pernier, J., Bertrand, O., Giard, M-H. and Echallier, J.F., "Mapping of scalp potentials by surface spline interpolation", *Electroencephalography and Clinical Neurophysiology*, 66: 75-81, 1987.
- [22] Van den Brink, D., Hagoort, P., "The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs", *Journal of Cognitive Neuroscience*, 16: 1068-1084, 2004.
- [23] DeLong, K.A., Urbach, T.P., & Kutas, M., "Probabilistic word pre-activation during language comprehension inferred from electrical brain activity", *Nature Neuroscience*, 8: 1117-1121, 2005.