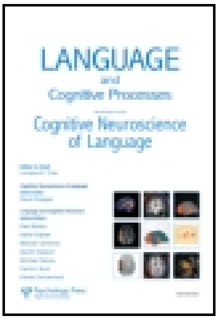

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/plcp20

# Seeing the initial articulatory gestures of a word triggers lexical access

Mathilde Fort [a], Sonia Kandel [a b c], Justine Chipot [a], Christophe Savariaux [c], Lionel Granjon [c] & Elsa Spinelli [a b d]

[a] Psychology, Laboratoire de Psychologie et NeuroCognition (CNRS UMR 5105), Université Pierre Mendès France, Grenoble Cedex 9, France

[b] Institut Universitaire de France, Paris, France

[c] GIPSA-lab, Dpt. Parole et Cognition (CNRS UMR 5216), Université Stendhal, Grenoble Cedex 9, France

[d] University of California, Berkeley, CA, USA

Version of record first published: 26 Sep 2012.

To cite this article: Mathilde Fort, Sonia Kandel, Justine Chipot, Christophe Savariaux, Lionel Granjon & Elsa Spinelli (2012): Seeing the initial articulatory gestures of a word triggers lexical access, Language and Cognitive Processes, DOI:10.1080/01690965.2012.701758

To link to this article: http://dx.doi.org/10.1080/01690965.2012.701758

Psychology Press
Taylor & Francis Group

# Seeing the initial articulatory gestures of a word triggers lexical access

**Mathilde Fort[1], Sonia Kandel[1,2,3], Justine Chipot[1], Christophe Savariaux[3], Lionel Granjon[3], and Elsa Spinelli[1,2,4]**

[1]Psychology, Laboratoire de Psychologie et NeuroCognition (CNRS UMR 5105), Université Pierre Mendès France, Grenoble Cedex 9, France
[2]Institut Universitaire de France, Paris, France
[3]GIPSA-lab, Dpt. Parole et Cognition (CNRS UMR 5216), Université Stendhal, Grenoble Cedex 9, France
[4]University of California, Berkeley, CA, USA



When the auditory information is deteriorated by noise in a conversation, watching the face of a speaker enhances speech intelligibility. Recent findings indicate that decoding the facial movements of a speaker accelerates word recognition. The objective of this study was to provide evidence that the mere presentation of the first two phonemes—that is, the articulatory gestures of the initial syllable—is enough visual information to activate a lexical unit and initiate the lexical access process. We used a priming paradigm combined with a lexical decision task. The primes were syllables that either shared the initial syllable with an auditory target or not. In Experiment 1, the primes were displayed in audiovisual, auditory-only or visual-only conditions. There was a priming effect in all conditions. Experiment 2 investigated the locus (prelexical vs. lexical or postlexical) of the facilitation effect observed in the visual-only condition by manipulating the target's word frequency. The facilitation produced by the visual prime was significant for low-frequency words but not for high-frequency words, indicating that the locus of the effect is not prelexical. This suggests that visual speech mostly contributes to the word recognition process when lexical access is difficult.

***Keywords:*** Visual speech; Lexical access; Phonological priming; Lexical frequency.

It is now widely admitted that visual information plays a significant role in speech perception. For example, seeing the articulatory gestures of a speaker's face enhances speech intelligibility in noisy environments (Sumby & Pollack, 1954). A study conducted in French with five noise conditions showed that the contribution of visual information when identifying consonants and vowels in an utterance increases as the signal-to-noise ratio decreases (Benoît, Mohamadi, & Kandel, 1994).

Correspondence should be addressed to Mathilde Fort, Laboratoire de Psychologie et NeuroCognition, Université Pierre Mendès France, BP 48, 38040 Grenoble Cedex 9, France. E-mail: mathilde.fort@upmf-grenoble.fr
We would like to thank Ronald Peereman for his very helpful comments on previous versions of the manuscript


http://www.psypress.com/lcp http://dx.doi.org/10.1080/01690965.2012.701758

Thus, decoding articulatory gestures in the presence of auditory information facilitates phoneme identification. However, there is scarce evidence regarding the contribution of visual information in lexical access. The purpose of the present study was to examine the mapping process of visual speech information to lexical units.

In the field of spoken word recognition, there is a debate concerning the processes underlying the mapping of sensory information from the acoustic input to the stored entries in the lexicon. Psycholinguistic models describing lexical access such as the Cohort model (Cohort I; Marslen-Wilson & Welsh, 1978; Cohort II; Marslen-Wilson, 1987, 1990), TRACE (McClelland & Elman, 1986), or Shortlist (Norris, 1994) posit different mechanisms to account for the mapping process. The timing of spoken word recognition is one of the differences that contrast these models. For instance, the Cohort model emphasises the "left to right" nature of lexical access and makes strong predictions on how the process of recognition unfolds over time. As in other models, a speech event (e.g., "log") activates a large set of lexical candidates. The Cohort model posits however that only the units that share the same onset[1] (e.g., "**log**", "**log**in", etc.) can be activated and be part of this pool of candidates (i.e., the initial cohort). In contrast, TRACE and Shortlist do not attribute such importance to the beginning of words. They postulate that the acoustic input will briefly activate the lexical units that *contain* the input, regardless of its location in the stimulus (e.g., "**log**", "**log**in" but also "cata**logue**"). Beyond those differences, note that these models have restricted their focus to *auditory* word recognition. There is scarce information about the *visual* contribution in the lexical access process.

Furthermore, the few studies that investigated word recognition in an audiovisual context presented contradictory results. Sams, Manninen, Surakka, Helin, and Kättö (1998) conducted a study in Finnish and used the McGurk effect to examine this issue. The McGurk effect is a perceptual illusion resulting from the integration of auditory and visual information (McGurk & MacDonald, 1976). An auditory /ba/ dubbed onto a visual /ga/ is perceived as /da/. Sams et al. compared the strength of audiovisual integration across situations in which the integration would result in the perception of a nonword from two real words, or vice versa. For example, pairing an auditory-presented word (e.g., /panu/, "**p**annu", stove) with another visually presented word (e.g., /kanu/, "**k**annu", pitcher) resulted in the perception of a pseudo-word (e.g., /**t**anu/). In another condition, an auditory-presented pseudo-word (e.g., /**p**iili/) paired with visual presentation of another pseudo-word (e.g., /kiili/) resulted in perception of a word (e.g., /tiili/, "**t**iili", brick). They expected a stronger McGurk effect for word responses than for pseudo-word responses, but their results did not support this prediction. The McGurk effect was similar for words and pseudo-words. The authors concluded that lexical knowledge did not mediate audiovisual integration, at least at the stage of phonetic processing.

Nonetheless, more recent studies showed a benefit of visual speech information in lexical access. Brancazio (2004) conducted a study in English that combined the McGurk and Ganong (Ganong, 1980) paradigms. In the latter, participants have to identify a phoneme (e.g., /t/ or /d/) that varies along a synthesised continuum (e.g., t↔d). Typically, when stimuli in the continuum form words and nonwords

[1] Note that the second version of the Cohort Model (Marslen-Wilson, 1987, 1990) does not posit an absolute match with the input for a word to be included in the initial cohort. During this bottom-up process, it tolerates a certain phonetic featural mismatch, since the change does not lead to the perception of another word (see e.g., Marslen-Wilson, Moss, & Van Halen, 1996; Marslen-Wilson & Zwitserlood, 1989, for further information about this issue).

(e.g., "**d**ask" vs. "**t**ask") there is tendency to respond more towards the phoneme that forms a word (e.g., the proportion of /t/ response is larger than /d/). By dubbing an auditory word (e.g., "**b**eg") onto a visual nonword (e.g., "**d**eg"), or a visual word (e.g., "**d**esk") onto an auditory nonword (e.g., "**b**esk"), he compared the strength of lexical activation across the auditory and visual domains. The results revealed that the lexical bias was stronger in the visual word condition than in the auditory word condition. This suggests that lexical context not only influences auditory phoneme categorisation but also visual speech processing during audiovisual word recognition (see also Barutchu, Crewther, Kiely, & Murphy, 2008; Windmann, 2004). Recent data in French indicate that visual information on the articulatory gestures of the speaker not only facilitates phoneme detection but also contributes to the process of word recognition (Fort, Spinelli, Savariaux, & Kandel, 2010). Fort et al. conducted a phoneme monitoring task with words and nonwords presented in auditory-only and audiovisual contexts with noise masking the acoustic signal. The results indicated that consonant phonemes (e.g., the target /**p**/) were more quickly and more accurately detected when they were embedded in words (e.g., /ʃa**p**o/ "chapeau", hat) than in nonwords (/ʃa**p**y/). When the acoustic signal was strongly deteriorated (i.e., at −18 dB), this "word superiority effect" was greater in the audiovisual than the auditory-only condition. The fact that the lexical bias was stronger in the audiovisual condition suggests that visual information associated with phoneme identity contributes to lexical activation during word recognition.

In sum, these studies suggest that visual speech contributes to the activation of lexical units. But what role does visual information alone—that is, without any auditory information—play in word recognition? Does it accelerate the word recognition processes? If it does, how does the facilitation take place? Priming tasks are particularly well adapted to address these questions because the information activated by primes can be manipulated experimentally. To our knowledge, three studies carried out in English examined this issue using a priming repetition procedure (Buchwald, Winters, & Pisoni, 2009; Dodd, Oerlemens, & Robinson, 1989; Kim, Davis, & Krins, 2004). Dodd et al. (1989) were the first to show that auditory words were better categorised when they were preceded by the presentation of the articulatory gestures of the same words. In the priming phase, Dodd et al. presented 10 visual word primes. The participants had to indicate whether the speaker was naming a plant or an animal (i.e., semantic categorisation task). During the second phase, those 10 "familiar" words were presented auditorily, amongst 10 other "new" words. The participants were also instructed to categorise each word as a plant or animal. The results showed that the participants were faster at categorising the "familiar" words when they were preceded by the block of visual primes, compared to a control condition with no priming. This suggests that visual speech may activate the same lexico-semantic network than the auditory information.

Kim et al. (2004) displayed word or nonword primes (e.g., word "back" or nonword "scay") in visual only speech that were followed by a written or auditory target. The target could either repeat the prime (repeating condition) or just match the syllable-length of the prime (i.e., unrelated condition, e.g., "sharp" for word "back", or "nunth" for nonword "scay"). Using naming and lexical decision tasks, the authors found a facilitatory priming effect on response times in the repeated condition (compared to the unrelated condition) when the stimuli were words, but not when they were nonwords. With a similar repetition priming paradigm, Buchwald et al. (2009) reported that participants identified spoken words in noise more accurately when the words were preceded by a visual speech prime of the same word compared to a control condition. Taken together, Buchwald et al. and Kim et al. studies suggest that the visual information in the word prime contributes to lexical processing by activating the lexical units that match the visual information. Moreover, Kim et al. (2004) found that the whole visual-only presentation of a word *accelerates* the following subsequent recognition of the same word. Although these studies provide evidence of the contribution of visual information in lexical access, the repetition priming task does not tell us how much visual information mediates the recognition process because the prime and target matched completely. In our research, the primes were the *initial fragment* of the targets. We could thus determine whether the visual presentation of the initial portion of a target-word could activate its lexical representation.

In a French priming study, Spinelli, Segui, and Radeau (2001) showed that auditory primes consisting of the first syllable of a disyllabic word facilitated the recognition of the written word (e.g., auditory /kaʁ/ → written "CARTABLE" /kaʁtabl/, schoolbag) compared to auditory primes presenting another syllable (e.g., auditory /lwɛ̃/ → written "CARTABLE" /kaʁtabl/). Their results revealed that the auditory syllable prime /kaʁ/ was enough information to activate the word recognition process of the word "CARTABLE". In the present research, we used the same paradigm but with the visual modality. We examined whether the presentation of the articulatory gestures corresponding to the two initial phonemes of a word is enough visual information to activate its lexical unit. Experiment 1 compared the effect of auditory-only, audiovisual or visual-only primes on the processing of an auditory target. We expected a priming effect in the auditory-only and audiovisual conditions but also in visual-only condition. The objective of Experiment 2 was to specify the locus of the priming effect in the visual-only condition.

# EXPERIMENT 1

## Method

### *Participants*

Sixty-three native French speakers (15 men and 42 women, mean age = 22 years, ranged from 17 to 38 years) participated in the experiment. They reported no auditory or visual disorders.

### *Stimuli and recording*

The stimuli were 90 disyllabic French target words (e.g., /by.ʁo/ "bureau", desk, the dot indicates the syllable boundary, see Appendix A). They were selected from a French spoken word database (mean lexical frequency = 24.64 occurrences per million, opm, LEXIQUE 3.71, New, Pallier, Brysbaert, & Ferrand, 2004). Each word was paired with two monosyllabic primes that either shared its initial phonemes (matching /by/) or did not (unrelated /fo/). We selected the matching and unrelated experimental primes that were salient in the visual speech signal. In other words, the consonants always had a bilabial (/p,b,m/), a labiodental (/f,v/), or an alveolar (/s/) place of articulation. The vowel was always rounded, producing a lip protrusion (/o,u,y/). To avoid the fact that the phonemic and visemic composition of the primes (and their visual intelligibility) could be predictive of whether a matching or unrelated target was to follow, we used the same primes in the unrelated condition and for the unrelated fillers. We ensured that the relation (unrelated/matching) between the prime and target could not be anticipated on the basis of prime identity. We thus balanced the number of times that a prime was used in the matching or in the unrelated condition with the filler trials. We determined the offsets of the primes on the basis of the acoustic signal. As we only used Consonant-Vowel (CV) primes, we computed for each prime the moment corresponding to 90% of vowel nucleus energy. These 90% points was set as the offset for the auditory-only, audiovisual and visual-only conditions. Thus, for the audiovisual and the visual-only conditions, the prime ended with the mouth of the speaker still open. The mean duration for the experimental primes was $M$=582 ms ($SD$=50 ms). For the purpose of the lexical decision task, 90 disyllabic pseudo-word targets were each paired with two primes (matching, unrelated). In order to reduce the proportion of matching items to 25%, an additional 90 unrelated filler words and 90 unrelated filler pseudo-words were included. We used a small proportion of matching items and a very short Interval Inter-Stimuli (i.e., ISI = 50 ms) between the presentation of the prime and the target in order to minimise the involvement of conscious response strategies (e.g., Hamburger & Slowiaczek, 1996).

Primes and targets were recorded separately (recording format: PAL; size of the video: 720×576 pixels, recording frame rate: 25 frames/s) by a linguistically trained female native speaker of French. The primes were recorded as syllables, in isolation. As a consequence, the matching primes and the targets came from a different utterance. The head, neck and top part of the speaker's shoulders were visible. The speaker had to start producing each utterance with her mouth closed and was instructed to avoid blinking during the stimulus pronunciation. A tri-CCD SONY DXC-990P camera and an AKG C1000S microphone were used to make the recording. The recording was digitalised with the Dps Reality v 3.1.9 software to obtain *avi* video files (compression codec: Intel Indeo Video 4.5). The videos for the primes were then cut using the 90% offset that was determined on the basis of the acoustic signal.

### Procedure

Participants were tested individually in a quiet room. The video stimuli were shown at 25 frames/s and the auditory component was presented at a 44100 Hz sampling rate. Participants were seated at 40 cm from the CRT monitor (refresh rate = 100 Hz). The size of the image on the screen was 19 cm×22 cm. We used a phonological priming procedure with an auditory lexical decision task. The syllabic primes were displayed either audiovisually, in auditory-only or visually-only in three separate blocks. During the auditory-only presentation of the primes, the still face of the speaker was presented with her mouth closed. During the audiovisual condition of presentation of the primes, the moving face of the speaker was displayed with the auditory component of the prime. In the visual-only condition, the sound was simply muted while the moving face was visible. Stimuli were counterbalanced across six experimental lists so that each participant went through all the conditions (Auditory-Only, Audiovisual, Visual-Only×Matching, Unrelated), but heard each target only once. Each trial began with presentation of a prime, followed 50 ms later by the auditorily presented target item. The target item was always presented with the still face of the speaker. Participants were asked to decide whether or not the target was a word as accurately and quickly as possible by pressing one of two response buttons. There was no time-out, so the next trial could only start after the participant answered. The experiment was performed using E-Prime 2.0 software (Psychological Software Tools, Pittsburgh, PA, USA). The discrepancy between what E-prime 2 reported as target onset and actual target display onset was <3 ms. Participants' accuracy and response times from target onset were also collected with the same software.

## Results

Mean reaction times (RTs) on experimental words correct responses in the six conditions are presented in Figure 1. Errors (0.9%) and RTs longer than 1,500 ms (5%) were removed. Due to the low percentage of errors (<1%), no analyses were carried out on errors. For the response times we computed the mean response time of each participant for each condition separately. Then we discarded the data above/below two SD from their corresponding mean (2.3% of the RTs). We also discarded two target words ("bottine", ankle boot, and "musique", music) because their error rate was higher than 30%. A 3 (Modality: auditory-only vs. audiovisual vs. visual-only)×2 (Prime Type: matching vs. unrelated) repeated measures ANOVA was conducted by participants ($F_1$) and by items ($F_2$).

The analysis revealed a main effect of Prime Type, $F_1(1, 62) = 64.96$, $p < .001$, $\eta_p^2 = .51$, $F_2(1, 87) = 71.08$, $p < .001$, $\eta_p^2 = .45$. The main effect of the modality was not significant, $F_1(1, 62) = 1.05$, $p > .05$, $\eta_p^2 = .015$, $F_2(1, 87) = 2.15$, $p > .05$, $\eta_p^2 = .024$. The interaction between Modality and Prime Type was significant, $F_1(2, 124) = 4.3$, $p < .05$, $\eta_p^2 = .07$, $F_2(2, 174) = 4.23$, $p < .05$, $\eta_p^2 = .05$. Planned comparisons revealed that the priming effect (matching vs. unrelated) in the visual-only condition was significant, $F_1(1, 62) = 3.93$, $p = .05$, $\eta_p^2 = .06$, $F_2(1, 87) = 4.27$, $p < .05$, $\eta_p^2 = .047$, but smaller than in the auditory and the audiovisual conditions, $F_1(1, 62) = 8.57$, $p < .005$, $\eta_p^2 = .12$, $F_2(1, 87) = 7$, $p < .01$, $\eta_p^2 = .075$. No difference was found between audiovisual and auditory-only conditions, $F_1(1, 62) < 1$.

## Discussion

This study investigated whether the visual information provided by the articulatory gestures that produced the first syllable of a word contributes to lexical activation during word recognition. The results indicate that the participants recognised words faster when they were preceded by their initial syllable than by an unrelated

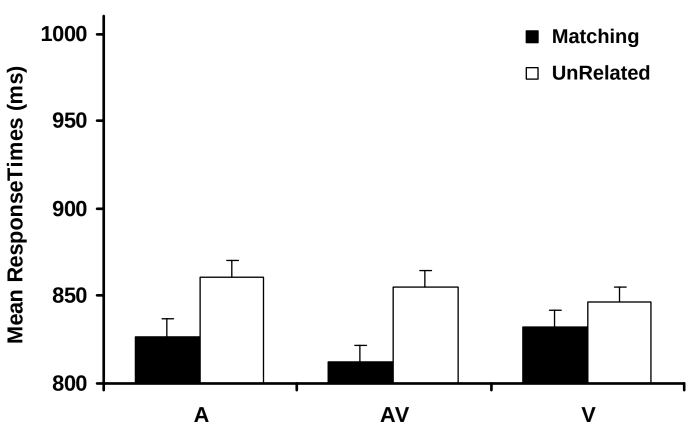


**Figure 1.** Mean response times (in ms) as a function of prime type (matching vs. unrelated) and modality [Auditory (A) vs. Audiovisual (AV) vs. Visual-only (V)]. Error bars represent mean standard error.

syllable prime. This facilitatory priming effect was significant for the auditory-only, audiovisual, and visual-only conditions. The facilitatory priming effect in the visual-only condition indicates that decoding the speaker's oro-facial gestures to produce the initial portion of the target words of a word will provide cues on the identity of that word. Seeing the speaker produce /by/ accelerates the recognition of /byʁo/ ("bureau"). This result is in line with previous priming studies (e.g., Buchwald et al., 2009; Kim et al., 2004) as well as the phoneme monitoring experiments conducted by Fort et al. (2010) in French and provides further evidence for the contribution of visual information in lexical activation.

It is noteworthy that the experimental design used in this experiment does not allow us to draw conclusions on the locus of the facilitation (i.e., lexical, prelexical, or else). If visual information can activate lexical units during word recognition process, the facilitatory priming effect observed in Experiment 1 should vary as a function of lexical variables. Indeed, most spoken-word recognition models assume that lexical frequency affects the activation level of lexical candidates during the word recognition process (e.g., Cohort II; Marslen-Wilson, 1987, 1990; TRACE; McClelland & Elman, 1986). For instance Cohort II and TRACE assume that increasing experience with a word results in a higher resting activation level for high frequency relative to low-frequency words. Thus, a high-frequency word needs less activation than a low-frequency word to be recognised. This explains why high-frequency words are processed faster than low-frequency words. Some other spoken-word recognition models such as the Neighbourhood Activation Model (NAM; Luce & Pisoni, 1998) posit that the locus of frequency effects takes place at a postlexical decision stage that follows initial lexical activation. For NAM, high-level lexical information such as frequency is assumed to operate by adjusting the activation levels represented within word-decision units but does not directly influence the resting-activation level of the lexical units per se. Thus, lexical frequency has in this model a postlexical rather than a lexical locus but nonetheless influence word recognition before lexical access is completed (e.g., Dahan, Magnuson, & Tanenhaus, 2001). Following the logic of these two types of models, lexical frequency should thus affect lexical or postlexical levels rather than the prelexical stages of word recognition process.

If the locus of the facilitation observed in Experiment 1 is lexical or postlexical, we may expect that word frequency modulates the priming effect (see e.g., Forster & Davis, 1984). To test whether the lexical frequency of the target had an influence on the size of the priming effect in the Experiment 1, we added lexical frequency as a continuous factor in the by-items analysis (mean lexical frequency = 25.18 opm, range: 0–433, LEXIQUE 3.71, New et al., 2004). The interactions between Prime Type and Lexical Frequency, $F_2(1, 87) = 1.41$, $p > .05$, $\eta_p^2 = .012$, or between Modality, Prime Type, and Lexical Frequency, $F_2(1, 87) < 1$, did not reach significance. This suggests that if lexical frequency has an impact on the priming effect in the visual-only condition, the frequency difference was not enough to modulate the size of the priming effect significantly. The aim of Experiment 2 was thus to manipulate lexical frequency as an a priori factor. We examined whether larger frequency differences could affect the visual-only priming effect observed in Experiment 1. If the observed facilitation is lexical (or postlexical) rather than prelexical, we should observe a different priming effect for high versus low-frequency target words. Alternatively, if the facilitation observed is not modulated by the target's frequency, the priming effect would rather be due to a prelexical than a lexical mechanism.

# EXPERIMENT 2

## Method

### Participants

Twenty native French speakers (5 men and 15 women, mean age = 25 years, ranged from 20 to 31 years) participated in the experiment. None of them participated in Experiment 1. They reported no auditory or visual disorders.

### Stimuli and recording

Sixty disyllabic French target words (e.g., /bo.ku/ "beaucoup" a lot, see Appendix B) were associated to two monosyllabic primes[2]: one for the matching (e.g., /bo/), and another for the unrelated condition (e.g., /ʁe/). Unlike the stimuli for Experiment 1, in Experiment 2 we made sure that there was no visemic overlap at all between the unrelated primes and the target, by using the classification established for French visemes by Gentil (1981). As in Experiment 1, we computed for the acoustic signal of each prime the moment time corresponding to 90% of vowel nucleus energy. These 90% point in time was then used for the visual-only condition of presentation of the prime. The mean duration for the experimental prime was $M$=637 ms ($SD$=100 ms). Half of the target words were high-frequency (mean lexical frequency = 124.61 opm; range: 26.8–626) and the other half were low-frequency words (mean lexical frequency = 0.78 opm; range: 0–3.65). Thus, each couple of matching and unrelated primes was matched to a high and a low-frequency word (e.g., matching prime /bo/ vs. unrelated prime /ʁe/ → high frequency /boku/, "beaucoup", low frequency /bolɛ/, "bolet", boletus, cf. Appendix B). There was a significant difference in lexical frequency between the low-frequency and the high-frequency words, $t(58) = 27.29$, $p<.001$. We controlled that high- and low-frequency target word groups were matched on duration [$M$=482 vs. $M$=477 ms, respectively, $t(58) < 1$]. We also ensured that the mean neighbourhood density for the low-frequency word group was not significantly different from the mean neighbourhood density for the high-frequency word group [$M$=11.6 vs. $M$=13.9, respectively, $t(58) = 1.37$, $p>.05$]. Indeed, this parameter, defined as the number of words that differ from a given target by one phoneme substitution, addition, or deletion (i.e., the number of similar-sounding word units that can compete with the target word unit) can influence word recognition performances (cf. Luce & Pisoni, 1998). Moreover, the target words were dissyllabic and had four phonemes. Finally, all of the 60 pairs of high/low-frequency target words displayed a uniqueness point (i.e., the position of the first phoneme from the left that distinguishes a word from all other words) at their fourth last phoneme (except for the item "vécu", /veky/, which has its uniqueness point at its third phoneme), as the position of this point may increase or decrease artificially the latencies estimation (Marslen-Wilson, 1990). The primes were different in the unrelated and in the matching condition for the experimental items (cf. Appendix B) but also different (from the unrelated and matching experimental items) in the fillers. However, as there were 300 trials per participant, each matching prime was only repeated four times among the 300 trials. It was therefore unlikely that the relation (unrelated/matching) could be anticipated on the basis of prime identity.

[2] In Experiments 1 and 2, the monosyllabic primes were CV syllables but could also be, in many cases, monosyllabic words. Indeed, most of the CV combinations in French actually correspond to monosyllabic words. Note however, that using a similar partial priming paradigm, Spinelli et al. (2001) found the same pattern of priming effects with words or pseudo-words as primes.

Regarding the visual intelligibility of the prime, 40% of the CV matching prime had a consonant that is visually salient (40%, i.e., that have a bilabial (/p,b,m/) a labiodental (/f,v/), or induce lip protrusion during constriction (/ʃ,ʒ/)), whereas 60% had a less salient consonant (i.e., alveolar (/t,d,n,l,s,z/), velar (/g,k/), or uvular (/ʁ/) place of articulation). Using the same criteria, the 43% CV unrelated primes had a consonant that is visually salient, whereas 57% had a less salient consonant. As the proportions of salient/less salient consonant were similar for the matching and the unrelated primes, it seems unlikely that prime visual saliency (i.e., visual intelligibility) could predict a relation (unrelated/matching) between the prime and the target.

These experimental items were matched to 60 disyllabic pseudo-words that contain the same number of phonemes and syllables than the target words. There were also 60 unrelated disyllabic filler words and 60 unrelated disyllabic filler pseudo-words to reduce the proportion of matching items to 25%. The stimulus recording and the speaker were the same than in Experiment 1.

### *Procedure*

We used a phonological priming procedure with a lexical decision task. The syllables primes were displayed in the visual-only condition. Stimuli were counter-balanced across two experimental lists so that each participant went through all the conditions (high-, low-frequency target words×matching, unrelated) but heard each target only once. The high-frequency and low-frequency target words were both displayed randomly. The procedure, size of the image on the screen, task, instructions software, and data-gathering were the same as in Experiment 1.

## Results

Mean RTs on experimental words for correct responses for the four conditions are shown in Figure 2. Incorrect responses (8.7%) and RTs longer than 1,500 ms (1.5%) were removed. We also removed two target pairs from the analysis ("baudet", mule/ "beauté", beauty; "titan", titan/"tirer", to pull) because both low-frequency members of each pair had more than 30% of errors. For each condition, we discarded data above and below two *SDs* from the mean (1.9% of the RTs). Percentages of errors for the four conditions were also computed.

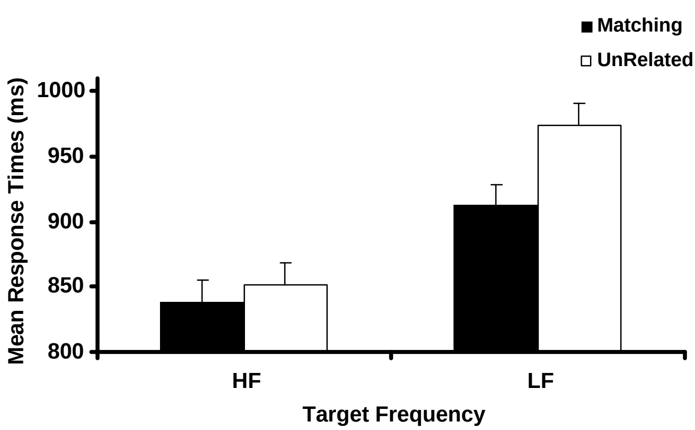


**Figure 2.** Mean response times (in ms) as a function of prime type (matching vs. unrelated) and target lexical frequency [low-frequency (LF) vs. high-frequency (HF)] for the visual-only primes. Error bars represent mean standard error.

### Reaction times

A 2 (Prime Type: matching vs. unrelated)×2 (Target Lexical Frequency: high-frequency vs. low-frequency words) repeated measures ANOVA was conducted by participants ($F_1$) and by items ($F_2$). Analysis of RTs revealed a main effect of Prime Type, $F_1(1, 19)=14.33$, $p<.005$, $\eta_p^2=.43$, $F_2(1, 27)=17.591$, $p<.005$, $\eta_p^2=.39$, and Target Lexical Frequency, $F_1(1, 19)=75.15$, $p<.001$, $\eta_p^2=.80$, $F_2(1, 27)=39.97$, $p<.001$, $\eta_p^2=.60$. The interaction between Prime Type and Target Lexical Frequency was significant, $F_1(1, 19)=7.01$, $p<.05$, $\eta_p^2=.27$, $F_2(1, 27)=12.82$, $p<.005$, $\eta_p^2=.32$. Planned comparisons revealed that the priming effect (matching vs. unrelated) was significant for the low-frequency target words, $F_1(1, 19)=13.2$, $p<.005$, $\eta_p^2=.59$, $F_2(1, 27)=27.1$, $p<.001$, $\eta_p^2=.50$ but not for the high-frequency target words, (both $F_s<1$).

### Errors

A 2 (Prime Type: matching vs. unrelated)×2 (Target Lexical Frequency: high-frequency vs. low-frequency words) repeated measures ANOVA was conducted by participants ($F_1$) and by items ($F_2$) on the percentage of errors. The analyses showed a main effect only of Target Lexical Frequency [*M* high frequency = 1.1%, *M* low frequency = 15.5%, $F_1(1, 19)=39.21$, $p<.001$, $\eta_p^2=.67$, $F_2(1, 27)=17.32$, $p<.001$, $\eta_p^2=.39$]. There was no significant effect of Prime Type, [*M* unrelated = 8.7%, *M* matching = 7.8%, $F_1(1, 19)<1$] nor interaction (both $F_s<1$).

## Discussion

Experiment 2 was designed to determine the locus of the facilitatory effect observed in the visual-only prime condition. As in Experiment 1, visual-only primes showing the articulation of the first two phonemes of a word facilitated its processing when presented auditorily as target. The results of Experiment 2 revealed that this facilitation effect was modulated by the word's frequency. The matching visual-only primes provided facilitation for the low-frequency words but not for high-frequency words (see Forster & Davis, 1984, for a similar trend in visual word recognition). We presume that the lack of facilitatory effect of visual speech primes on high-frequency words could be due to the fact that their recognition is already advantaged relative to the other lexical candidates (see Goldinger, Luce, Pisoni, & Marcario, 1992 for a similar claim).[3] Because lexical frequency had an impact on visual speech priming, it is likely that the locus of the facilitation observed in Experiments 1 and 2 was not prelexical. Thus, this result revealed that the visual information provided by the articulatory gestures of a syllable corresponding to the two first phonemes of a word activates the lexical units of that word. In other words, these findings indicate that the articulatory gestures that produce the initial portion of a word are enough information to reach the lexical level. This suggests that visual speech can activate lexical units as soon as the first articulatory gestures are available in the visual signal (see also, Dahan et al., 2001; Warren & Marslen-Wilson, 1987, 1988, for a similar claim regarding the auditory component of speech).

[3] Note that we had 63 participants and 90 items counterbalanced across six lists in Experiment 1, and 20 participants and 30 pairs of high/low-frequency words counterbalanced across two lists in Experiment 2. As a consequence, in both Experiments 1 and 2 we had about 10 data points per item in each condition (i.e., $63/6 \approx 20/2 \approx 10$) and 15 data points per participant in each condition (i.e., $90/6=30/2=15$). This rules out the possibility that the reason why no priming effect was found for the high-frequency condition was due to less statistical power in Experiment 2 compared to Experiment 1.

## GENERAL DISCUSSION

The objective of this study was to examine the role of visual information in the process of word recognition and determine whether this information mediated prelexical rather than lexical (or postlexical) processing. In Experiment 1 we used audiovisual, auditory-only, and visual-only primes to investigate the early phases of word recognition. The primes were syllables that could or could not match the onset of a word. The results indicated that priming the first syllable of a word facilitates its auditory recognition. The data for the visual-only condition indicate that viewing the articulation of a syllable that matches the onset of a word facilitates its recognition. Experiment 2 shed some light onto the level of processing that can be affected by the visual information on the articulatory gestures with this priming procedure. In this experiment, visual-only syllables primed target words of contrasting lexical frequency. There was a significant priming effect for low-frequency words but not for high-frequency words.

Experiment 1 also revealed that there was less facilitation for the visual-only condition than the audiovisual and auditory-only conditions. The facilitation for the audiovisual condition was not significantly greater than the auditory-only condition. This visual-only $<$ audiovisual $\approx$ auditory-only pattern could be due to the fact that a visible speech gesture can match more than one acoustic phoneme. For example, the facial gesture in the prime /by/ is similar to the gesture for articulating /py/. This means that in the visual-only condition, the prime /by/ activated the lexical units of the words compatible with the articulatory gestures of /by/—like "bureau", /byʁo/ or "rébus", /ʁeby/, rebus—but also by /py/ (e.g., "purée", /pyʁe/, puree, "trapu", /tʁapy/, stocky). The visual-only primes may have activated more lexical candidates than the audiovisual and auditory-only primes, increasing the lexical competition and thus decreasing the size of the facilitation effect. For the same reason, the benefit of the visual information in a clear and audible speech signal (i.e., in the audiovisual condition) may have been too small to be significantly larger than the facilitation observed for the auditory information alone (i.e., in the auditory-only condition). Another explanation for the visual-only $<$ audiovisual $\approx$ auditory-only trend could be that in Experiment 1 the relationship between prime and target was established on a phoneme—not a viseme—basis. Thus, 20% of the unrelated primes were visually similar with the onset of the targets (Appendix A). The visual speech gestures for the prime and onset of the target shared the same visemes (e.g., Gentil, 1981). Although this overlap did not prevent us from observing a significant priming effect in the visual-only condition, it may have contributed to decrease its size when compared to the one observed in the audiovisual and auditory-only conditions.

The lack of benefit in the audiovisual condition as compared to the auditory-only condition could be that the priming effect was already at ceiling from the clear auditory information alone and hence seeing the speaker did not provide any more information above and beyond what was already obtainable from listening. This interpretation is in line with our idea that visual speech may play a significant role in lexical access especially when this process is difficult. It is in line with previous studies (Fort et al., 2010) that revealed that the benefit of visual information in lexical access is especially notable in adverse conditions. Second, the visual-only $<$ audiovisual $\approx$ auditory-only trend may be explained by the fact that the target words used in Experiment 1 were relatively frequent (mean lexical frequency = 24.64 opm). Indeed, in Experiment 2, we only found a significant facilitation for the low-frequency words (mean lexical frequency = 0.78 opm) but not for the high-frequency words (mean lexical frequency = 124.61 opm). This interpretation remains however speculative, since there was no correlation between the frequency of the target and the size of the priming effect in Experiment 1. Further studies are required to determine whether this trend can be accounted for by these or other variables.

The priming effect for low-frequency words observed in Experiment 2 suggests that visual speech may play a determinant role especially when the lexical access process is more time consuming or constitutes a cognitive load. Decoding visual information would especially enhance the activation level of a lexical unit when the conversational situation is somewhat adverse. As a consequence, visual information would contribute to the word recognition process essentially when a lexical unit requires a large amount of activation to be recognised. However, other studies found that decoding visual speech information (i.e., speechreading) seems to be easier for "easy words" (i.e., high-frequency words and words with sparse neighbourhood density) rather than for "hard words" (i.e., low-frequency words and words with high neighbourhood density) (e.g., Auer, 2002; Auer & Bernstein, 1997; Kaiser, Kirk, Lachs, & Pisoni, 2003, Mattys, Bernstein, & Auer, 2002). This discrepancy might be due to the fact that these authors used offline identification tasks whereas the design of the present study allowed us to measure the specific contribution of visual speech in lexical access with an online procedure. Moreover, these studies did not investigate the role of visual information in lexical access but how high-level lexical parameters such as frequency or neighbourhood density influence offline speechreading in a rather "top-down" perspective. Finally, it is noteworthy that in these experiments speechreading performance could be biased by guessing strategies (Lyxell & Rönnberg, 1987). In contrast, our study provides insight on how visual information enhances the retrieval of the correct lexical unit during speech perception, in a more "bottom-up" fashion (see van Linden & Vroomen, 2007, for a distinction between top-down lexical influences vs. bottom-up visual speech benefits).

In sum, these findings suggest that visual information taps into lexical or postlexical rather than prelexical levels of processing during word recognition. This is in line with studies that presented whole words as primes (e.g., Buchwald et al., 2009; Kim et al., 2004). The results of our study allow us to go a step further. The data provide evidence that the mere presentation of the first two phonemes—that is, the articulatory gestures of the initial syllable—is enough visual information to activate a lexical unit and initiate the lexical access process. As stated above, this study suggests that visual speech is processed continuously to activate lexical units, as soon as the first articulatory gestures are available in the visual signal. This claim is consistent with the temporal precedence of visual information in conversational speech. In French there is a natural asynchrony in the availability of visual and auditory information (e.g., Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009). The precedence of the onset of mouth movements (i.e., the visual information) with respect to the onset of the voicing (i.e., the auditory information) may vary from 100 to 300 ms. In other words, the visual information is available at least 100 ms before the auditory information and can therefore be processed before any acoustic processing is initiated. Studies conducted in French by Cathiard, Lallouache, Mohamadi, and Abry (1995) indicate that the visual system decodes the visual information on lip movements such that we can identify a vocalic phoneme /y/ before its acoustic onset (see also Smeele, 1994, for similar results with bilabial stop consonants in Dutch and Jesse & Massaro, 2010, for a large-scale study). This means that the visual detection of an articulatory gesture (e.g., labial closure) may activate a pool of lexical candidates compatible with this visual information before the auditory information becomes available. For instance, in the case of the prime /bo/, all the words containing a bilabial stop and a rounded vowel could be activated at least 100 ms before the auditory information could be decoded. The up-coming auditory information would refine the initial pool by adding some acoustic features that are not salient in the visual signal (e.g., voice, nasality, etc.). It could thus be argued that the role of the auditory information would be to decrease the number of lexical candidates for recognition. This is in line with the idea that visual speech may play a priming role (Munhall & Tohkura, 1998) and we extend it to the domain of lexical access. Further research is of course necessary to investigate this issue.

Our results suggest that visual speech enhances the activation of lexical units when lexical access is difficult (i.e., for the low-frequency words). We also posit that visual speech may facilitate lexical access by *activating* lexical units and anticipating auditory processing. In Kim et al.'s (2004) Experiment 4, the presentation of a visual-only prime that matched the two initial phonemes of the target but not the coda (e.g., articulatory gestures for "back" → written target "BAND") resulted in inhibition (compared to a control condition such as "leaf" → "BAND"). In our study however, we observed a facilitatory priming effect when presenting a visual-only prime that matched the two first phonemes of the target (e.g., /by/ → /byʁo/, "bureau", desk) compared to a control condition (e.g., /fo/ → /byʁo/, "bureau"). These findings suggest that the mere presentation of the facial movements corresponding to the two initial phonemes is enough to spread activation towards lexical units. Furthermore, the visual presentation for one mismatching consonant seems to be sufficient to provide *inhibition*. This has several implications. First, the data suggest that visual-speech is fine-tuned. Second, both results are in line with the observations in the auditory word recognition literature (facilitation: Spinelli et al., 2001 vs. inhibition: e.g., Dufour & Peereman, 2003a, 2003b; Slowiaczek & Hamburger, 1992). Third, these data indicate that processing occurs at the lexical (or postlexical) level. If the locus of the observed facilitation is lexical, we posit that one role of visual speech in lexical access would be to increase and/or decrease the size of the pool of lexical candidates by activating and/or inhibiting them during lexical competition. If this locus of this effect is rather postlexical, this would mean that visual speech would influence lexical access rather at decisional stage than at lexical level per se. The latter seems nonetheless unlikely given the natural precedence of visual speech over its auditory counterpart (Chandrasekaran et al., 2009). Indeed, it has been shown that seeing the first articulatory gestures of a speaker could contribute to making predictions about the identity of the forth-coming auditory output (Van Wassenhove, Grant, & Poeppel, 2005). In line with this idea, the visual information could "prime" the lexical units before the auditory information becomes available. This hypothesis remains however speculative and further studies should be done to test it.

Current models of spoken-word recognition such as Cohort II (Marsen-Wilson, 1987, 1990), TRACE (McClelland & Elman, 1986), and Shortlist (Norris, 1994) do not consider the contribution of visual information in the process of lexical access. Our study provides evidence that such models should incorporate the orofacial gestures as a source of information in their architecture (see Brancazio, 2004; Fort et al., 2010, for a supplementary discussion about this issue). The originality of our study was to show that seeing the articulatory gestures corresponding to the initial portion of a word provides enough information to activate the lexical level.

The present finding provides further constrains on how visual speech could be integrated in psycholinguistic models of spoken word recognition. It suggests that visual speech is processed continuously to activate lexical units, as soon as it is present in the speech stream.

Manuscript received 30 May 2011
Revised manuscript received 4 June 2012
First published online 25 September 2012



## REFERENCES

Auer, E. T. (2002). The influence of the lexicon on speech read word recognition: Contrasting segmental and lexical distinctiveness. *Psychonomic Bulletin & Review*, *9*, 341–347.

Auer, E. T., & Bernstein, L. E. (1997). Speechreading and the structure of the lexicon: Computationally modeling the effects of reduced phonetic distinctiveness on lexical uniqueness. *Journal of the Acoustical Society of America*, *102*, 3704–3710.

Barutchu, A., Crewther, S., Kiely, P., & Murphy, M. (2008). When /b/ill with /g/ill becomes /d/ill: Evidence for a lexical effect in audiovisual speech perception. *European Journal of Cognitive Psychology*, *20*(1), 1–11.

Benoît, C., Mohamadi, T., & Kandel, S. (1994). Effects of phonetic context on audio-visual intelligibility of French. *Journal of Speech and Hearing Research*, *37*, 1195–1203.

Brancazio, L. (2004). Lexical influences in audiovisual speech perception. *Journal of Experimental Psychology. Human Perception and Performance*, *30*, 445–463.

Buchwald, A. B., Winters, S. J., & Pisoni, D. B. (2009). Visual speech primes open-set recognition of spoken words. *Language and Cognitive Processes*, *24*, 580–610.

Cathiard, M.-A., Lallouache, M. T., Mohamadi, T., & Abry, C. (1995). Configurational vs. temporal coherence in audio-visual speech perception. In K. Elenius & P. Branderud (Eds.), *Proceedings of the 13th International Congress of Phonetic Sciences* (Vol. 3, pp. 218–221). Stockholm: ICPhS.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*, e1000436.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*, 317–367.

Dodd, B., Oerlemens, M., & Robinson, R. (1989). Cross-modal effects in repetition priming: A comparison of lip-read graphic and heard stimuli. *Visible Language*, *22*, 59–77.

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

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning Memory and Cognition*, *10*, 680–698.

Fort, M., Spinelli, E., Savariaux, C., & Kandel, S. (2010). The word superiority effect in audiovisual speech perception. *Speech Communication*, *52*, 525–532.

Ganong, W. F., III. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology. Human Perception and Performance*, *6*, 110–125.

Gentil, M. (1981). *Etude de la perception de la parole: Lecture labiale et sosies labiaux* [Study on speech perception: Lipreading and visemes]. Rapport technique, France: IBM.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 1211–1238.

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

Jesse, A., & Massaro, D. (2010). The temporal distribution of information in audiovisual spoken-word identification. *Attention, Perception, & Psychophysics*, *72*, 209–225.

Kaiser, A. R., Kirk, K. I., Lachs, L., & Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *46*(2), 390–404.

Kim, J., Davis, C., & Krins, P. (2004). Amodal processing of visual speech as revealed by priming. *Cognition*, *93*, B39–B47.

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

Lyxell, B., & Rönnberg, J. (1987). Guessing and speechreading. *British Journal of Audiology*, *21*, 13–20.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*, 71–102.

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

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

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

Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 576–585.

Mattys, S. L., Bernstein, L. E., & Auer, E. T. (2002). Stimulus-based lexical distinctiveness as a general word-recognition mechanism. *Perception & Psychophysics*, *64*, 667–679.

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

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Munhall, K. G., & Tohkura, Y. (1998). Audiovisual gating and the time course of speech perception. *Journal of the Acoustical Society of America*, *104*, 530–539.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524.

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

Sams, M., Manninen, P., Surakka, V., Helin, P., & Kättö, R. (1998). McGurk effect in Finnish syllabes, isolated words, and words in sentences: Effects of word meaning and sentence context. *Speech Communication*, *26*, 75–87.

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

Smeele, P. (1994). *Perceiving speech: Integrating auditory and visual speech* (Published doctoral dissertation). Delft University of Technology, The Netherlands.

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

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*, 212–215.

van Linden, S., & Vroomen, J. (2007). Recalibration of phonetic categories by lipread speech versus lexical information. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1483–1494.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 1181–1186.

Warren, P., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*, *41*, 262–275.

Warren, P., & Marslen-Wilson, W. D. (1988). Cues to lexical choice: discriminating place and voice. *Perception & Psychophysics*, *43*, 21–30.

Windmann, S. (2004). Effects of sentence context and expectation on the McGurk illusion. *Journal of Memory and Language*, *50*, 212–230.

# APPENDIX A: EXPERIMENT 1

Material used as Experimental target words, unrelated (UnR), and matching (M) primes in Experiment 1. Primes with an asterix (*) represents the unrelated primes that shared the same viseme with the two initial phonemes of the target

| *UnR primes* | *M primes* | *Target words* | *UnR primes* | *M primes* | *Target words* |
|---|---|---|---|---|---|
| /mu/* | /bo/ | Beauté /bote/ | /my/* | /po/ | Poney /ponɛ/ |
| /vo/ | | Bottine /botin/ | /my/* | | Potier /potje/ |
| /vu/ | | Bonnet /bonɛ/ | /by/* | | Polaire /polɛʁ/ |
| /vy/ | /bu] | Bouchon /buʃõ/ | /sy/ | | Pochette /poʃɛt/ |
| /fo/ | | Bouquet /bukẽ/ | /sy/ | | Pommier /pomje/ |
| /my/* | | Boulon /bulõ/ | /vy/ | | Poème /poɛm/ |
| /py/* | | Bouteille /butɛj/ | /vy/ | | Potion /posjõ/ |
| /so/ | | Bougie /buʒi/ | /fy/ | /pu/ | Poumon /pumõ/ |
| /so/ | | Boulet /bulɛ/ | /fy/ | | Poupée /pupe/ |
| /vo/ | | Bouton /butõ/ | /fy/ | | Pouvoir /puvwaʁ/ |
| /fo/ | /by/ | Bureau /byʁo/ | /py/* | | Poulet /pulɛ/ |
| /fu/ | | Bûcheron /byʃəʁõ/ | /sy/ | | Poussière /pusjɛʁ/ |
| /mu/* | | Buffet /byfɛ/ | /sy/ | | Poussin /pusẽ/ |
| /bu/ | fo/ | Folie /foli/ | /vo/ | | Poubelle /pubɛl/ |
| /pu/ | | Fossile /fosil/ | /vo/ | | Poulain /pulẽ/ |
| /py/ | | Faucon /fokõ/ | /vo/ | | Poussette /pusɛt/ |
| /su/ | | Forêt /foʁɛ/ | /fu/ | /py/ | Public /pyblik/ |
| /sy/ | | Fauteuil /fotœj/ | /so/ | | Punaise /pynɛz/ |
| /by/ | /fu/ | Fourrure /fuʁyʁ/ | /vu/ | | Purée /pyre/ |
| /sy/ | | Foulard /fulaʁ/ | /bu/ | /so/ | Sauna /sona/ |
| /mo/ | /fy/ | Fusil /fyzi/ | /by/ | | Sonnette /sonɛt/ |
| /po/ | | Futur /fytyʁ/ | /fu/ | | Solide /solid/ |
| /pu/ | | Fumer /fyme/ | /fy/ | | Sommeil /somɛj/ |
| /so/ | | Fusée /fyze/ | /my/ | | Saucisse /sosis/ |
| /by/* | /mo/ | Module /modyl/ | /py/ | | Sauterelle /sotəʁɛl/ |
| /fy/ | | Modèle /modɛl/ | /pu/ | | Soja /soʒa/ |
| /fy/ | | Mollet /molɛ/ | /vy/ | | Saumon /somõ/ |
| /py/* | | Motif /motif/ | /bo/ | /su/ | Soutien /sutjẽ/ |
| /sy/ | | Moteur /motœʁ/ | /by/ | | Souper /supe/ |
| /vy/ | | Moment /momã/ | /by/ | | Sourire /suʁiʁ/ |
| /vy/ | | Moral /moʁal/ | /fu/ | | Souris /suʁi/ |
| /by/* | /mu/ | Moutarde /mutaʁd/ | /mo/ | | Souci /susi/ |
| /fo/ | | Mouchoir /muʃwaʁ/ | /bo/ | | Soucoupe /sukup/ |
| /po/* | | Moulin /mulẽ/ | /mo/ | | Soudure /sudyʁ/ |
| /po/* | | Mouton /mutõ/ | /py/ | | Souffrance /sufʁãs/ |
| /py/* | | Moucheron /muʃəʁõ/ | /vo/ | | Souplesse /suplɛs/ |
| /bu/* | /my/ | Musée /myze/ | /fu/ | /sy/ | Sucette /sysɛt/ |
| /po/* | | Muguet /mygɛ/ | /fu/ | | Support /sypoʁ/ |
| /su/ | | Musique /muzik/ | /mo/ | | Sujet /syʒɛ/ |
| /vo/ | | Museau /myzo/ | /mo/ | | Supplice /suplis/ |
| /su/ | /po/ | Pommade /pomad/ | /bu/ | /vo/ | Volaille /volaj/ |
| /su/ | | Potage /potaʒ/ | /bu/ | | Voleur /volœʁ/ |
| /my/* | | Paupière /popjɛʁ/ | /su/ | | Vautour /votuʁ/ |
| /my/* | | Poteau /poto/ | /su/ | | Volume /volym/ |
| /my/* | | Police [polis/ | /sy/ | | Volant [volã/ |

# APPENDIX B: EXPERIMENT 2

Material used as Experimental HF (high-frequency) and LF (low-frequency) target, unrelated (UnR), and matching (M) primes in Experiment 2. Frequency (in occurrences per million) indicates word frequency. Neigh. Density corresponds to the number of neighbours for each target

| *UnR primes* | *M primes* | *HF Target-words* | *Frequency* | *Neigh. density* | *LF Target-words* | *Frequency* | *Neigh. density* |
|---|---|---|---|---|---|---|---|
| /le/ | /ba/ | Bateau /bato/ | 106.55 | 19 | Baquet /bakɛ/ | 0.46 | 25 |
| /ʁe/ | /bo/ | Beaucoup /boku/ | 626 | 1 | Bolet /bolɛ/ | 0 | 17 |
| /ʁe/ | /bo/ | Beauté /bote/ | 68.57 | 23 | Baudet /bodɛ/ | 0.11 | 11 |
| /tã/ | /by/ | Bureau /byʁo/ | 156.68 | 14 | Burin /byʁẽ/ | 0.57 | 10 |
| /vo/ | /ka/ | Cadeau /kado/ | 98.09 | 22 | Caban /kabã/ | 0 | 15 |
| /mẽ/ | /ko/ | Côté /kote/ | 250.51 | 30 | Copeau /kopo/ | 0.1 | 12 |
| /pa/ | /ku/ | Couteau /kuto/ | 51.08 | 11 | Coupon /kupõ/ | 0.51 | 17 |
| /ku/ | /de/ | Début /deby/ | 109.88 | 7 | Débit /debi/ | 1.1 | 14 |
| /my/ | /di/ | Dîner /dine/ | 84.73 | 21 | Divan /divã/ | 1.03 | 9 |
| /ʒə/ | /do/ | Donner /done/ | 233 | 16 | Doper /dope/ | 0.35 | 11 |
| /ʁi/ | /fa/ | Façon /fasõ/ | 212.6 | 15 | Fagot /fago/ | 0.03 | 9 |
| /gi/ | /fo/ | Folie /foli/ | 122.47 | 8 | Fauter /fote/ | 0 | 21 |
| /gẽ/ | /fu/ | Fumer /fume/ | 35.91 | 14 | Futon /fytõ/ | 0.28 | 11 |
| /by/ | /ga/ | Gâteau /gato/ | 42.33 | 14 | Gadoue /gadu/ | 0.37 | 7 |
| /ʃy/ | /ma/ | Matin /matẽ/ | 265.03 | 22 | Magot /mago/ | 2.24 | 11 |
| /ku/ | /mi/ | Midi /midi/ | 35.15 | 6 | Mica /mika/ | 0.34 | 12 |
| /kẽ/ | /mo/ | Moment /momã/ | 403 | 11 | Moka /moka/ | 0.54 | 12 |
| /kẽ/ | /mo/ | Monnaie /monɛ/ | 26.82 | 14 | Momie /momi/ | 2.45 | 11 |
| /ʒu/ | /ni/ | Niveau /nivo/ | 45.46 | 5 | Nicher /niʃe/ | 0.35 | 12 |
| /li/ | /nu/ | Nouveau /nuvo/ | 170.28 | 1 | Nougat /nuga/ | 0.89 | 5 |
| /so/ | /pa/ | Paquet /pakɛ/ | 36.9 | 25 | Patin /patẽ/ | 1.12 | 24 |
| /ʃi/ | /po/ | Poser /poze/ | 73.73 | 21 | Potée /pote/ | 0.04 | 24 |
| /mi/ | /sa/ | Salon /salõ/ | 37.06 | 25 | Sabot /sabo/ | 1.79 | 13 |
| /fe/ | /so/ | Sauter /sote/ | 57.89 | 26 | Saumon /somõ/ | 3.65 | 11 |
| /ʁo/ | /se/ | Série /seʁi/ | 33.34 | 16 | Sénat /sena/ | 1.38 | 9 |
| /pə/ | /sy/ | Sujet /syʒɛ/ | 107.92 | 9 | Sumo /symo/ | 0.88 | 4 |
| /ʁy/ | /ti/ | Tirer /tiʁe/ | 113.71 | 31 | Titan /titã/ | 1.06 | 13 |
| /me/ | /tu/ | Toucher /tuʃe/ | 49.46 | 16 | Toupie /tupi/ | 1.5 | 5 |
| /ʁy/ | /ve/ | Vécu /veky/ | 51.14 | 1 | Vérin /veʁẽ/ | 0.05 | 5 |
| /ʁy/ | /ve/ | Vélo /velo/ | 32.95 | 7 | Verrue /veʁy/ | 0.66 | 7 |