

Audiovisual vowel monitoring and the word superiority effect in children

Mathilde Fort,¹ Elsa Spinelli,^{1,2} Christophe Savariaux,³ and Sonia Kandel¹

Abstract

The goal of this study was to explore whether viewing the speaker's articulatory gestures contributes to lexical access in children (ages 5–10) and in adults. We conducted a vowel monitoring task with words and pseudo-words in audio-only (AO) and audiovisual (AV) contexts with white noise masking the acoustic signal. The results indicated that children clearly benefited from visual speech from age 6–7 onwards. However, unlike adults, the word superiority effect was not greater in the AV than the AO condition in children, suggesting that visual speech mostly contributes to phonemic—rather than lexical—processing during childhood, at least until the age of 10.

Keywords

audiovisual speech perception, children, lexical access, noise, speech development, word recognition, word superiority effect

Several studies have shown that adults can rely on the information carried by the speaker's oro-facial gestures to identify speech in noisy situations (Benoit, Mohamadi, & Kandel, 1994; Erber, 1969; Sumby & Pollack, 1954; see also Green, 1998 for a review). For instance, it has been shown that, when the auditory information was adversely affected by white noise, consonant and vowel phonemes embedded in VCVVCVC¹ nonsense words were better identified in audiovisual than in auditory-only presentations (Benoit et al., 1994). Thus, decoding facial gestures enhances phoneme intelligibility when the auditory information has deteriorated. However, little is known about the contribution of visual information to lexical activation in adults (e.g., Brancazio, 2004; Fort, Spinelli, Savariaux, & Kandel, 2010) and to our knowledge, this question has never been addressed in children. The purpose of the present research was to investigate whether children use visual information in lexical access and, if so, at what age.

Audiovisual speech perception and lexical access in adulthood

Most studies have addressed the role of visual cues in audiovisual speech perception and lexical access process separately. Little is known about the interaction between these two sources of information. If visual speech benefits phoneme perception, it should also benefit word recognition at a lexical level in audiovisual face-to-face situations. Some of the studies investigating lexical access in adults in an audiovisual context used the McGurk paradigm (McGurk & MacDonald, 1976). The McGurk effect is a perceptual illusion in which an auditory /ba/ dubbed onto a visual /ga/ is perceived as /da/ or /θa/. This finding provided strong evidence that acoustic and visual signals integrate. Using this perceptual illusion, Brancazio (2004), Windmann (2004), and Barutchu, Crewther, Kiely, Murphy, and Crewther (2008) showed that visual information contributed to lexical access, whereas Sams, Manninen, Surakka, Helin, and Kättö (1998) were unable to find this pattern

of results. More recent data from a consonant monitoring experiment in French indicates that the presence of visual information not only facilitates phoneme detection, but also contributes to the process of lexical access in noise (Fort et al., 2010). Other studies using repetition priming paradigms showed that visual-only speech primes activate lexical representations (Buchwald, Winters, & Pisoni, 2009; Kim, Davis, & Krins, 2004).

Most of these studies suggest that the visual system codes information on facial movements during speech perception and that this information is exploited during the lexical access process. The purpose of the present research was to investigate this issue from a developmental perspective.

Visual speech influence in childhood

Numerous studies have shown that very young infants are sensitive to visual speech even at the first stages of development (Burnham & Dodd, 2004; Kuhl & Meltzoff, 1982; Patterson & Werker, 1999, 2003; Rosenblum, Schmuckler, & Johnson, 1997). Infants seem to be able to detect auditory-visual correspondences for vowels at 2 months old (Patterson & Werker, 2003) and they even possess the ability to integrate these two sources of information as early as 4.5 months old (Burnham & Dodd, 2004). Weikum et al. (2007) also reported that 4-month-olds are able to extract relevant information from a visual-only speech stream to discriminate two languages

¹ Laboratoire de Psychologie et de NeuroCognition, Grenoble, France

² University of California, Berkeley, USA

³ Gipsa-lab, Grenoble, France

Corresponding author:

Mathilde Fort, Laboratoire de Psychologie et de NeuroCognition—Psychologie, Bâtiment Sciences de l'Homme et Mathématiques BP47, Grenoble Cedex 9, 38040 France.

Email: mathilde.fort@upmf-grenoble.fr

(French vs. English). Thus the visual correlates of speech seem to influence its perception even in the early phases of human development. However, sensitivity to this information does not seem to be clearly understood in the later stages of childhood. In support of this claim, McGurk and MacDonald (1976) observed in their original study that children between the ages of 3–5 and 7–8 years exhibited a weaker McGurk effect than adults. Moreover, Massaro, Thompson, Barron, and Laren (1986) asked preschoolers and elementary school children (4–6 and 6–10 years old) to identify an auditory /ba/ dubbed onto a visual /da/. The results showed that children's responses were less dominated by the visual input (i.e., lower percentage of /da/ responses) than those of adults, suggesting that sensitivity to visual speech increases with age (Dupont, Aubin, & Menard, 2005; Hockley & Polka, 1994; Massaro, 1984).

A recent study (Jerger, Damian, Spence, Tye-Murray, & Abdi, 2009) pointed out that the diverging evidence of infants' and children's sensitivity to visual speech could be due to differences in experimental procedures and task demands. Infants' perception has been investigated with indirect measures via online responses (i.e., looking times), whereas most research on children uses direct procedures via offline responses (e.g., syllable identification) which—according to the authors—require a more conscious access and more detailed visual speech representations. To test their assumption, they examined visual speech influence on phonological processing by children using an indirect approach, the multimodal picture-word naming task. The idea underlying the original picture naming task (Jerger, Martin, & Damian, 2002) is that the simultaneous presentation of a congruent distractor sharing the same onset (e.g., "peach," /pi:tʃ/) would facilitate the word naming process of a picture item (e.g., "pizza," /pi:tza/) as compared to the simultaneous presentation of an unrelated distractor (e.g., eagle, /ɪgəl/). In this experiment, children of ages 4 to 14 years were asked to name a picture located on the speaker's chest. To test whether visual speech may influence this naming process, the participants could only hear (audio-only condition, AO) or both hear and see (audiovisual condition, AV) the speaker articulating either the congruent or the unrelated distractor. As expected, the authors found that the children named the picture faster when hearing simultaneously a congruent distractor than an unrelated distractor, even if they were told not to pay attention to it. Interestingly, this effect was greater for the AV than AO conditions but only for the younger (4-year-olds) and the older (10–14-year-olds) children, not for the intermediate age groups (5-, 6–7- and 8–9-year-olds).

These results indicate that English children between 5 and 9 years old are less influenced by visual speech than adults on both indirect and direct tasks. This is in line with the studies described earlier. To assess whether this lack of visual influence was due to a temporary loss of speech-reading skills, the authors also administered the participants a visual-only speech-reading task. Surprisingly, they found that the speech-reading scores increased with age. The authors argue that the lack of visual speech influence within the age of 5 to 9 could be due to a period of transition (e.g., reflecting the re-organization of phonological representational knowledge) rather than a loss of visual speech processing per se (e.g., such as speech-reading skills; see Massaro et al., 1986, for such a claim).

Recent findings investigated whether English and Japanese children (aged 6, 8, and 11 years) benefited from the presence of visual facial information to perceive speech when the auditory signal is deteriorated by noise (Sekiyama & Burnham, 2008). Using a syllable identification task, their results showed that visual speech

benefits especially increased with age between 6 and 8 for English participants but remained stable for Japanese children. In the earlier stages of development (i.e., 6 years), however, the results showed that the size of visual influence was very weak and equivalent for Japanese and English children. These findings indicate that, from the age of 8, children are able to extract reliable information from the speaker's orofacial gestures to enhance the intelligibility of speech sounds. Because younger English children showed a weaker visual speech influence than their elders, these data provide evidence that speech-reading ability—to perceive speech in noise—becomes more accurate with age. Moreover, this study also showed that the size of visual speech influence across age differed between the English and Japanese participants. Together, these findings clearly indicate that language experience has differential developmental and cross-linguistic impacts on AV speech processing. Nonetheless, the mechanisms underlying the development of this capacity are still not well understood (e.g., Jerger et al., 2009; Sekiyama & Burnham, 2008) and little is known about the specific contribution of visual processing in children during lexical access.

Lexical access in adults

The realization of a word can vary throughout many different factors (e.g., such as speaker, speaking rate, contexts, presence of noise, etc.). Successful word recognition is thus a challenging and complex issue for novice speech perceivers. To be able to map these different realizations onto the same meaning, it is generally assumed that the mature speaker (and perceiver) has a mental lexicon (Treisman, 1960) which contains a representation of each word he/she knows (but see Goldinger, 1998; Johnson, 2006 for alternative hypotheses). Findings in adults such as the word superiority effect (Cutler, Mehler, Norris, & Segui, 1987), Ganong effect (Ganong, 1980), or the phonemic restoration effect (e.g., Samuel, 1981), suggest that lexical representations influence phoneme perception. Cutler et al. (1987) observed that a consonant (e.g., /b/) was detected faster in a French word (e.g., "belle," /bel/; i.e., beautiful) than in a pseudoword (e.g., "berre," /beɛʁ/). This word superiority effect suggests that lexical knowledge biases adults' phoneme perception/decision. However, little is known about the influence of word representations in children's speech perception.

Lexical access in children

The emergence and degree of phonological specification of word representations have been investigated in toddlers (Fennell & Werker, 2003; Hallé & de Boysson-Bardies, 1996; Stager & Werker, 1997; Swingley & Aslin, 2000, 2002; see also Best, Tyler, Gooding, Orlando, & Quann, 2009, for recent findings about this issue), but only a few studies (Ackroff, 1981; Walley, 1988) examined the influence of lexical knowledge on spoken word recognition and phonological processing in older children. Both of them used the phoneme restoration paradigm (Warren, 1970). When a portion of a word corresponding to a phoneme is replaced by white noise, adult listeners tend to hear the word as intact as when white noise is added to it; they "restore" the missing speech segment. The phonemic restoration effect is greater in words (e.g., "progress," /progrɛs/, where the bold letter indicates the missing phoneme) than in pseudo-words (e.g., "crogless," /krogles/), suggesting that a lexical bias is responsible for this effect in adults (Samuel, 1981). Walley (1988) used this paradigm on children and showed that 5-year-

olds elicited less phonemic restoration effect than adults (see also Ackroff, 1981, for similar results with 6- and 8-year-olds). As a consequence, we may hypothesize that lexical knowledge has less influence on phoneme perception in children than adults.

In sum, infants are able to process facial speech gestures from the first stages of their development (e.g., Kuhl & Meltzoff, 1982; Patterson & Werker, 2003), but the influence of visual speech still seems to increase with age during childhood (Dupont et al., 2005; Hockley & Polka, 1994; Jerger et al., 2009; Massaro, 1984; Massaro et al., 1986; McGurk & MacDonald, 1976; Sekiyama & Burnham, 2008), even up to the age of 11 (e.g., Sekiyama & Burnham, 2008). In addition, visual speech processing skills in children aged between 5 and 10 have rarely been documented in the literature as compared to the same ability in infants (Jerger et al., 2009; Sekiyama & Burnham, 2008). Moreover, there is no information available on how children use visual speech in lexical processing. The goal of the present study was to get insight into this issue from a developmental perspective.

Experiment I

Fort et al. (2010) conducted a consonant phoneme monitoring task in which adults had to detect French consonant targets in words and pseudo-words that were presented in noise. The results revealed that the targets were detected better and faster in words than in pseudo-words. This word superiority effect was more important in the audio-visual than audio-only modality. These results suggest that visual speech contributes in lexical access per se. The aim of the present research was to investigate whether children also exploit this information in the process of lexical access. To explore this question, we decided to use a phoneme monitoring task with vowel targets. We selected vowel targets instead of consonants for several reasons. First, vowels are entities that play a primary role in speech development (e.g., Locke, 1993). Second, vowels are more salient to listeners than consonants (Ladefoged, 2001) and seem to better resist in noise masking (Nooitboom & Doodeman, 1984, cited in Cutler, Sebastián-Gallés, Soler-Vilageliu, & Van Ooijen, 2000). This makes the task easier for young children and therefore more adapted for a developmental study. We thus conducted a vowel phoneme monitoring task in children involving words and pseudo-words presented in audio-only (AO) and audiovisual (AV) contexts. Stimuli were mixed with noise in the acoustic signal to avoid ceiling effects and to enhance the role of visual speech on phonological processing.

If the influence of lexical knowledge on phonological processing increases with age (e.g., Walley, 1988), we should observe a progressive increase of the word superiority effect. Similarly, if visual speech benefits increase as a function of age, there should be an AV advantage only for the older children (Sekiyama & Burnham, 2008). In other words, according to Sekiyama and Burnham's findings, we should only observe a weak AV advantage from ages 5 to 8. Finally, if visual speech contributes to lexical activation in childhood we should observe, as did Fort et al. (2010), a greater word superiority effect in the AV than AO presentations.

Method

Participants. Ninety-six native French-speaking children participated in the experiment, ranging in age from 5 years 2 months to 10 years 10 months. The children were distributed into five groups according to age and school year: 5–6 years, kindergarten (mean age: 5 years 8 months, $N = 19$); 6–7 years, first grade (mean age:

6 years 11 months, $N = 18$); 7–8 years, second grade (mean age: 7 years 11 months, $N = 20$); 8–9 years, third grade (mean age: 8 years 11 months, $N = 20$); and 9–10 years, fourth grade (mean age: 9 years 10 months, $N = 19$). They all had normal or corrected-to-normal vision and reported no auditory disorders.

Stimuli. The stimulus set was composed of 40 stimuli of dissyllabic word/pseudo-word pairs. We selected the items that are known at age 5. Twenty pairs were target-present trials (i.e., the target vowel was in the carrier item; see Appendix 1) and 20 pairs were target-absent trials. Each pseudo-word was constructed by changing the first phoneme of the first syllable in the original word (e.g., the French word “bateau” = /batø/, boat, paired with the pseudo-word /lato/). We used this procedure to ensure that the pseudo-words were close to the original words but were nevertheless non-lexical items. We decided to specifically change the first phoneme in order to keep constant across the members of each pair the consonantal environment which preceded the vowel target phoneme (e.g., the target /ø/ in the word “bateau” = /batø/, or in the pseudo-word /lato/).

Target-present trials. For the 20 pairs of target-present trials (or carrier items), the critical vowel target was constant between each word/pseudo-word pair and was always located at the end of the second syllable (e.g., the target /ø/ in the word “bateau” = /batø/, or in the pseudo-word /lato/). The reason for choosing this position was that the word superiority effect should be stronger than if the target appeared in the initial syllable (cf. Frauenfelder, Segui, & Dijkstra, 1990). We used three vowel target phonemes: two rounded vowels (/ø/, /y/) and one stretched vowel (/ɛ/).

Target-absent trials. The 20 word/pseudo-word pairs of target-absent trials were constructed using the same phonemes as in the previous paragraph. However, these pairs were preceded (at the beginning of each block) by a non-matching vowel target phoneme (e.g., the target /ø/ in “merci” = /mɛ̃si/ or in the pseudo-word /lɛ̃si/). The non-matching vowel target was carefully chosen to be acoustically and visually (i.e., articulatory) different from the second vowel phoneme.

Stimuli recording. The stimuli and vowel targets were recorded in a soundproof room by a trained female native French speaker. We only presented the face of the talker (from her chin to her eyebrows) in front of a green background. The recording was done with a tri-CCD SONY DXC-990P camera and an AKG C1000 S microphone. The recording was digitalized with the Dps Reality v 3.1.9 software to obtain mpeg video files. In the AO condition we used the soundtrack extracted from the video so that the acoustic signal was identical in the AO and AV conditions. We used the Matlab 7.1 software to generate the noise and add it to each utterance. We used one noise level or signal to noise ratio (i.e., -9 dB). Signal to noise ratio, often written S/N or SNR, is a measure of signal strength relative to background noise. The ratio is usually measured in decibels (dB). We used the following formula: $SNR = 20 \log_{10}(Vs/Vn)$ in which Vs and Vn are respectively the original signal amplitude and the noise amplitude. As each utterance energy was dependent on its vowel and on its consonant type (e.g., plosive, fricative), we calculated the mean strength for each stimulus and then added white noise to keep the signal to noise ratio constant throughout the duration of the stimulus. The stimuli were distributed in two experimental lists corresponding to the two conditions of presentation of the stimuli (AO vs. AV). Each list contained 10 pairs of target-present trials and 10 pairs of target-absent trials presented in random order.

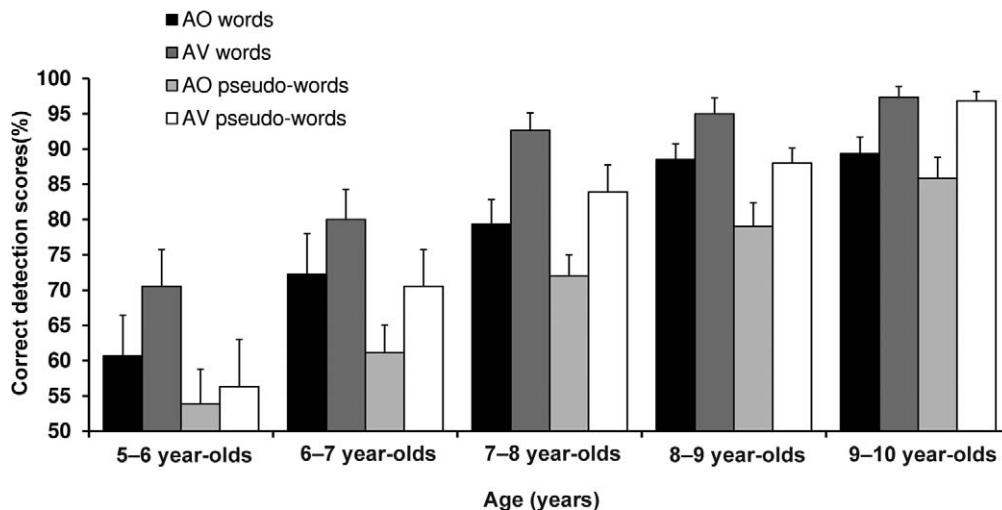


Figure 1. Percentages of correct detection scores as a function of age (age: 5–6-year-olds vs. 6–7-year-olds vs. 7–8-year-olds vs. 8–9-year-olds vs. 9–10-year-olds), modality (AO vs. AV), and lexical status (words vs. pseudo-words). Errors bars represent the standard error.

Procedure. Participants were tested individually in a room apart from their classroom inside the school. They sat at 50 centimetres from an LCD screen (Neovo 17 X-17A) in a darkened soundproof room. The video stimuli were presented at 25 frames/s with a resolution of 720×576 pixels. The auditory component of the stimuli was provided at a 44100 Hz sampling rate by two SONY SRS-88 speakers located on both sides of the screen. The experiment was conducted with E-Prime 2.0 software (Psychological Software Tools, Pittsburgh, PA). The experiment consisted of two different sessions separated by 1 or 2 weeks. In each session, participants had to detect the target vowel in the target item (a word or a pseudo-word). They were told that the vowel target could or could not be in the carrier utterance. A go/no go response task was employed whereby participants pressed the space bar of a keyboard as quickly as possible when they perceived the target in the carrier item. To limit the cognitive load, the vowel target type was displayed by a block so that the target was always displayed auditorily once before each block. Before each trial, the participants' response hand was always on the space bar. They were instructed to do nothing if they did not hear it and used the dominant hand to press the space bar if the target was present.

In each session, all the items were either displayed in AO (with the still face of the speaker) or in AV (with the moving face of the speaker). Thus each item was presented twice to each participant, once in AV, once in AO. To avoid “learning” the items from one session to the other, the sessions were separated by 1 or 2 weeks. Within each session, each child was presented 20 word/pseudo-word pairs of target-present trials and 20 word/pseudo-word pairs of target-absent trials. Within each block, half of the items contained the target (target-present trials) and half did not (target-absent trials). For the AV condition, the participants were told to watch and listen to the stimuli carefully. This instruction intended to avoid that the participant focused on one modality more than the other (e.g., Alsius, Navarra, Campbell, & Soto-Faraco, 2005). The order of the modality of presentation was counterbalanced between participants. Thus if the first participant perceived the stimuli in AV in the first session and in AO during the second session, the second participant would perceive the stimuli in AO during the first session and then in AV in the second session. A training session of six stimuli preceded each condition and could be repeated as many times as necessary.

Results

Correct detection scores and mean response latencies (measured from target onset for the correct responses) were calculated for each participant and item pair. Due to a great variability of latencies, no analysis was carried out on this measure. To ensure that audiovisual enhancement was not masked over age by a floor effect for the 5–6 year-olds in AO ($M = 57\%$, chance level = 50%) and a ceiling effect for the 9–10 year-olds in AV ($M = 97\%$), we performed an arcsine transformation of the square root of the mean correct detection scores for each participant and for each item in each condition before the analysis (Winer, 1970). A 5 (age: 5–6/6–7/7–8/8–9 and 9–10 years) \times 2 (modality: AO vs. AV) \times 2 (lexical status: word vs. pseudo-word) mixed analysis of variance (ANOVA) was conducted, on one hand means computed for each participant (F_1) and on the other hand means computed for each item (F_2) using the STATISTICA 10 software (Statsoft, Inc, 1984–2011). Mean correct detection scores for all the conditions are shown in Figure 1.

The analyses revealed a main effect of age, $F_1(4, 91) = 21.28$, $p < .001$, $\eta^2_p = .48$; $F_2(4, 76) = 78.75$, $p < .001$, $\eta^2_p = .81$. Results yielded a main effect of modality, $F_1(1, 91) = 45.32$, $p < .001$, $\eta^2_p = .33$; $F_2(1, 19) = 48.48$, $p < .001$, $\eta^2_p = .80$, indicating that participants were better able to detect the vowel targets in AV than AO. The main effect of lexical status was also significant, $F_1(1, 91) = 51.24$, $p < .001$, $\eta^2_p = .36$; $F_2(1, 19) = 16.05$, $p < .001$, $\eta^2_p = .46$, suggesting that participants were better at detecting vowels embedded in words with respect to pseudo-words. The interaction between modality and age was significant by items, $F_1(4, 91) = 1.25$, $p = .30$, $\eta^2_p = .05$; $F_2(4, 76) = 4.9$, $p < .001$, $\eta^2_p = .20$. To further investigate the effect of modality, we performed planned comparisons indicating a significant audiovisual advantage for the 6–7-year-olds, $F_1(1, 91) = 6.91$, $p < .01$, $\eta^2_p = .07$; $F_2(1, 19) = 10.71$, $p < .005$, $\eta^2_p = .36$, 7–8-year-olds, $F_1(1, 91) = 17.7$, $p < .001$, $\eta^2_p = .16$; $F_2(1, 19) = 26.7$, $p < .001$, $\eta^2_p = .58$, 8–9-year-olds, $F_1(1, 91) = 8.44$, $p < .005$, $\eta^2_p = .09$; $F_2(1, 19) = 10.9$, $p < .005$, $\eta^2_p = .36$ and 9–10-year-olds, $F_1(1, 91) = 15.7$, $p < .001$, $\eta^2_p = .14$; $F_2(1, 19) = 40.4$, $p < .001$, $\eta^2_p = .68$, but not for the 5–6-year-olds, $F_1(1, 91) = 1.91$, $p = .16$, $\eta^2_p = .02$; $F_2(1, 19) = 1.35$, p

$= .25$, $\eta^2_p = .06$. No other two-way or three-way interaction was observed.

To make sure that the children did not develop any response strategies, we also computed a d' for each mean correct detection score of each child in each condition, using this formula: $d' = z(\text{CD}) - z(\text{FA})$, in which z represents the inverse of the normal cumulative distribution and CD and FA refer respectively to the mean probability of correct vowel detections and false alarms. A 5 (age: 5–6/6–7/7–8/8–9, and 9–10-year-olds) \times 2 (modality: AO vs. AV) \times 2 (lexical status: word vs. pseudo-word) mixed ANOVA was conducted by participants. The analyses on d' revealed a main effect of age, $F(4, 91) = 14.88$, $p < .001$, $\eta^2_p = .40$, and a strong AV advantage, $F(1, 91) = 160.22$, $p < .001$, $\eta^2_p = .64$. There was also a main effect of lexical status, $F(1, 91) = 19.63$, $p < .001$, $\eta^2_p = .18$. No interaction between these factors was observed.

Discussion

The aim of experiment 1 was to investigate whether children exploit visual information in the process of lexical access. We conducted a vowel monitoring task with words and pseudo-words, in AO and AV, with white noise masking the acoustic signal. The results showed the influence of age on detection scores. They also provided evidence that lexical knowledge affected the children's performance: children had better scores to detect a vowel embedded in a word than in a pseudo-word. Moreover, children had greater scores in AV than AO modalities from age 6–7. Thus this study indicates that even young children (from 6–7 years old) are able to disentangle an auditory signal from a noisy background by processing the articulatory gestures of a speaker when the auditory information is degraded.

Contrary to our expectations, no significant interaction was obtained between lexical status and modality. In other words, the word superiority effect was not significantly higher in AV than in AO for participating children. This is at odds with Fort et al. (2010), who showed that this interaction was significant in adults. To insure that the absence of such a pattern of results was not due to the fact that we used vowel targets (experiment 1) instead of consonants (Fort et al., 2010), we conducted a vowel phoneme monitoring task in adults. As in experiment 1, adult participants had to detect vowel targets in words and pseudo-words presented in AO and AV contexts. As in Fort et al. (2010), the stimuli were displayed with white noise in the acoustic signal. If visual speech activates lexical representations in adults, we should observe a greater word superiority effect in AV rather than AO, both on correct vowel detection scores and latencies.

Experiment 2

Method

Participants. Sixty native adult French speakers (17 men, 43 women) participated in the experiment. They all had normal or corrected-to-normal vision and reported no auditory disorders.

Stimuli. The stimulus set was composed of 180 stimuli of dissyllabic word/pseudo-word pairs. Ninety pairs were target-present trials (i.e., the target was in the carrier item; see Appendix 2) and 90 pairs were target-absent trials. As in experiment 1, each pseudo-word was constructed by changing the first phoneme of the first syllable in the original word (e.g., “jumeaux” = /ʒymo/, twins à /lymo/).

Target-present trials. For the 90 pairs of target-present trials (or carrier items), the vowel target was located, as in experiment 1, at the end of the second syllable (e.g., the target /o/ in “jumeaux” = /ʒymo/ or in /lymo/, meaningless). We used five vowel target phonemes: three rounded vowels (/o/, /u/, /y/) and two stretched vowels (/i/, /e/). According to the LEXIQUE 2 adult's spoken word database (New, Pallier, Brysbaert, & Ferrand, 2004), the average frequency of the carrier words was 62 occurrences per million.

Target-absent trials. The 90 word/pseudo-word pairs of target-absent trials were constructed using the same phonemes as in the previous paragraph. However, these pairs were always preceded by a non-matching vowel target (e.g., the target /e/ in “bougie” = /buʒi/, candle, or in the pseudo-word /toʒti/).

Stimuli recording. The stimuli's recording and the speaker were the same as in experiment 1. As 9–10-year-olds' correct detection scores were at ceiling at –9 dB in experiment 1, we decided to select a weaker signal to noise ratio (–18 dB) that was already used by Fort et al. (2010). This was done to avoid a ceiling effect on adult's performances and to be able to directly compare our data with Fort et al.'s findings. The SNR of the acoustic signal of each item was computed and modified as described in the experiment 1 Method section. The stimuli were distributed in two experimental lists corresponding to the two presentation conditions (AO –18 dB; AV –18 dB). Each list contained 90 word/pseudo-word pairs (i.e., 45 pairs of target-present trials and 45 pairs of target-absent trials).

Procedure. The participants were tested individually. The material and software were the same as in experiment 1. The procedure was also similar to the one used in experiment 1. However, in experiment 2, the target vowel phonemes were not presented by block but in a random order. As a consequence, the vowel target was displayed auditorily before each item. Then the word or the pseudo-word (carrier) was presented. As in experiment 1, participants had to detect the target vowel they just perceived in the carrier word or pseudo-word. A go/no go response task was employed whereby participants pressed the space bar of a keyboard as quickly as possible when they heard the target in the carrier item.

The experiment consisted of two blocks. Within one block, half of the carrier items were presented in AV, with the moving face of the speaker. Within the other block, the other half was displayed in AO, with the still face of the speaker. Between each block, a black screen informed the participants of a change in presentation modality. Within each block, half of the items contained the vowel target (target-present trials) and half did not (target-absent trials). The presentation condition for each item was counterbalanced between participants so that each word/pseudo-word pair was displayed in all the conditions, but only once by each participant. Block order and noise conditions were counterbalanced across participants. Within each block, the order of the stimuli was randomized. A training session of eight stimuli preceded each condition.

Results

Correct detection scores and mean response latencies (measured from target onset for the correct responses) were calculated for each participant and for each item pair. For the latencies, we computed the mean of each participant for each condition separately. Then we discarded the data above/below two standard deviations from their corresponding mean (3.2% of our data). A 2 (modality: AO vs. AV) \times 2 (lexical status: word vs. pseudo-word) within-participants ANOVA

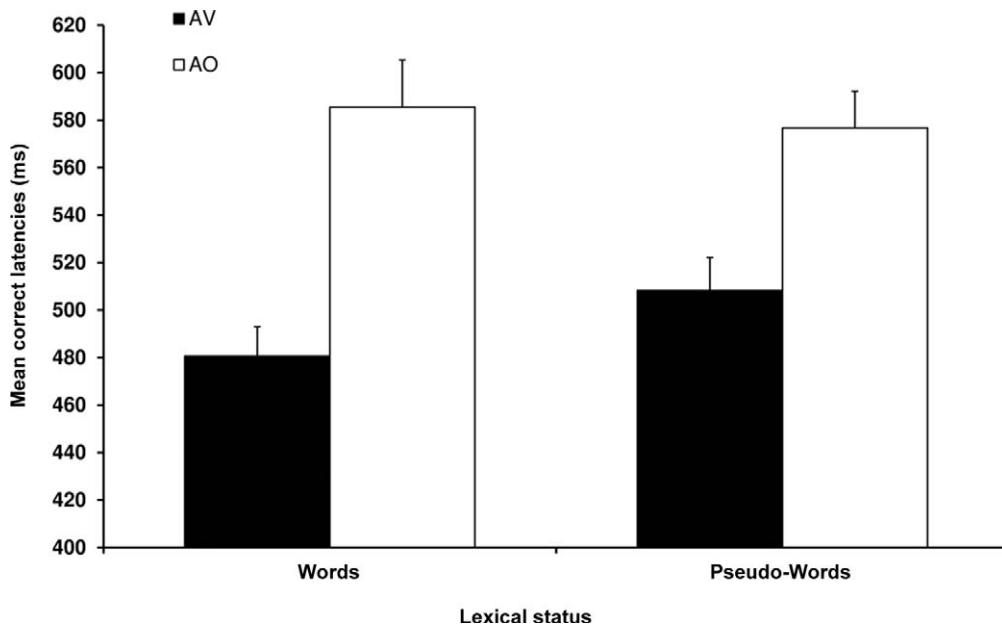


Figure 2. Mean correct latencies as a function modality (AO vs. AV) and lexical status (words vs. pseudo-words). Errors bars represent the standard error.

was conducted both by participants (F_1) and by items (F_2) on correct detection scores and latencies.

Response latencies. Mean latencies for the four conditions are shown in Figure 2.

First, the analyses revealed that participants detected the vowel target phonemes faster in AV than in AO, $F_1(1, 59) = 38.62, p < .001, \eta^2_p = .40$; $F_2(1, 89) = 75.68, p < .001, \eta^2_p = .50$. The main effect of lexical status was not significant, $F_1(1, 59) = 1.68, p = .19, \eta^2_p = .03$; $F_2(1, 89) = 2.2, p < .014, \eta^2_p = .03$. The interaction between lexical status and modality was significant by participants, $F_1(1, 59) = 6.17, p < .05, \eta^2_p = .10$, and marginally significant by items, $F_2(1, 89) = 3.46, p = .07, \eta^2_p = .04$, indicating that the word superiority effect was significant in AV, $F_1(1, 59) = 17.21, p < .001, \eta^2_p = .23$; $F_2(1, 89) = 14.95, p < .001, \eta^2_p = .14$, but not in AO, F_1 and $F_2 < 1$.

Correct detection scores. Mean correct detection scores for the four conditions are shown in Figure 3.

The analyses revealed a significant main modality effect in favour of the AV condition, $F_1(1, 59) = 385.61, p < .001, \eta^2_p = .87$; $F_2(1, 89) = 273.13, p < .001, \eta^2_p = .87$. The main effect of the lexical status was not significant, F_1 and $F_2 < 1$. However, the interaction between lexical status and modality was significant by participants, $F_1(1, 59) = 4.96, p < .05, \eta^2_p = .08$, $F_2(1, 89) = 2.51, p = .11, \eta^2_p = .03$, indicating that the lexical effect was significant in AV, $F_1(1, 59) = 5.27, p < .05, \eta^2_p = .08$, $F_2(1, 89) = 3.95, p = .051, \eta^2_p = .04$, but not in AO, $F_1(1, 59) = 1.68, p > .05, \eta^2_p = .03$, $F_2(1, 89) < 1$.

To make sure that the adult participants did not develop any response strategies, we also computed a d' for each stimulus pair, using the same method as described in experiment 1. A 2 (modality: AO vs. AV) \times 2 (lexical status: word vs. pseudo-word) within participants ANOVA was conducted by participants on these data. The results yielded a main effect of modality in favor of the AV condition, $F(1, 59) = 646.43, p < .001, \eta^2_p = .92$, but also a main word

superiority effect, $F(1, 59) = 9.54, p < .001, \eta^2_p = .14$. The interaction between lexical status and modality was not significant, $F(1, 59) = 2.1, p = .15$, but planned comparisons revealed that the word superiority effect was significant in AV, $F(1, 59) = 10.17, p < .005, \eta^2_p = .15$, but not in AO, $F < 1$.

Discussion

The goal of experiment 2 was to replicate Fort et al.'s findings using vowel targets instead of consonants. Hence, as in Fort et al. (2010), we conducted a vowel phoneme monitoring task with words and pseudo-words displayed in AV and in AO with white noise in the acoustic signal. Notably, the results yielded a significant effect of modality on correct detection scores, d' and latencies, suggesting that the presence of visual information increases vowel intelligibility (e.g., Benoit et al., 1994) and also accelerates its detection process. As we expected, latency data and correct detection scores showed a greater word superiority effect in the AV rather than in the AO condition. This confirms the findings presented by Fort et al. (2010) and supports the idea that visual speech contributes to the activation of lexical representations in adults. These results go a step further than Fort et al. (2010), as they only observed a word superiority effect for correct detection scores. The fact that our results are also significant for latencies suggests that visual speech not only contributes to the activation of lexical representations in adults, but it also *accelerates* the process of word recognition.

In sum, experiment 2 provided vowel phoneme monitoring data indicating that seeing the articulatory gestures of a speaker contributes to and even accelerates the lexical representation activation process in adults.

General discussion

The goal of this study was to examine the influence of visual speech on a lexical access process from a developmental perspective. We thus conducted a vowel phoneme monitoring task

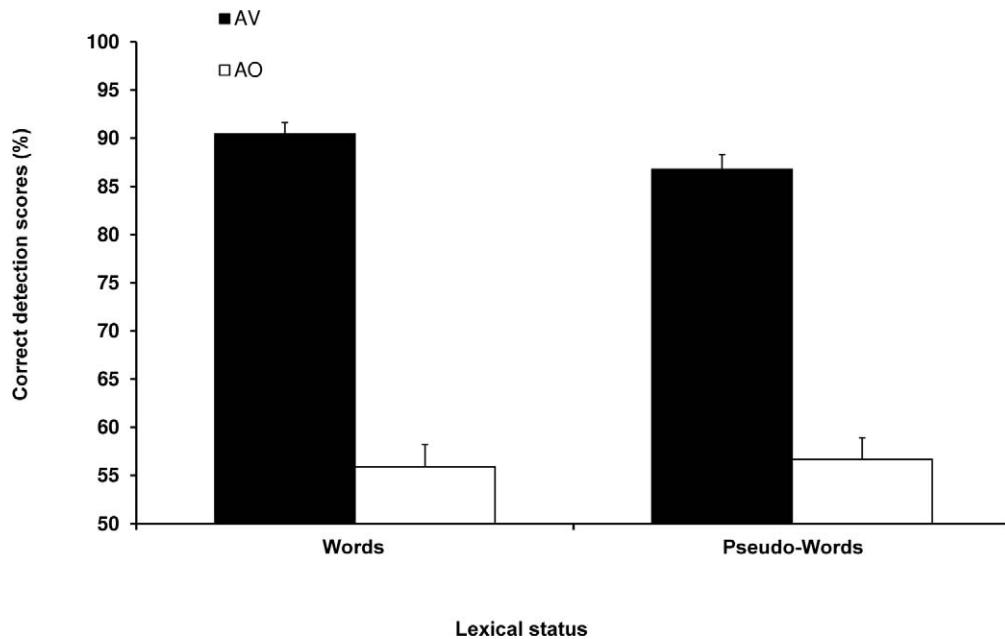


Figure 3. Percentages of correct detection scores as a function of modality (AO vs. AV) and lexical status (words vs. pseudo-words). Errors bars represent the standard error.

in words and pseudo-words in AO and AV modality with noise in the acoustic signal in children aged from 5–6 to 10 years-old (experiment 1) and in adults (experiment 2).

First, our study provided evidence that lexical knowledge affects the children's performance. Children had better scores to detect a vowel phoneme embedded in a word than a pseudo-word. This word superiority effect suggests that lexical knowledge biases vowel detection processes in children at least from the age of 5. In other words, it seems that children can rely on lexical context to enhance vowel intelligibility. To our knowledge, our study is the first showing a strong influence of lexical information on the child phoneme detection process.

Second, in line with previous findings (e.g., Sekiyama & Burnham, 2008) we observed a significant increase of the AV advantage over the AO condition with age. Planned comparisons showed that this interaction was due to the fact the 5–6-year-olds did not significantly benefit from the presence of visual information. However, the audiovisual advantage was clearly significant for the other age groups. More specifically, children had greater correct detection scores in AV than AO modalities from age 6–7 onwards.

To our knowledge, this research is the first study reporting a size-similar benefit across age of coherent visual speech on performance in children before the age of 8. Indeed, these data suggest that, from the age of 6–7 years old, children are able to process visual speech to compensate for the lack of information in the auditory signal. This study seems to be the first set of data showing that young children (from 6–7 years old) are able to disentangle an auditory signal from a noisy background by processing the articulatory gestures of a speaker when the auditory information is degraded. The audiovisual benefit could be explained by the fact that under deteriorated acoustic conditions, visual and acoustic signals complement each other. The auditory information that has been masked by the noise is available in the visual signal and can be recovered by seeing the lips, teeth, tongue and jaw movements (see Benoît et al., 1994, for results with vowels, and Summerfield, 1987, for data with consonants). As a conclusion, this study

indicates that children of ages 6–7 to 10 use not only acoustic cues (e.g., Allen, Wightman, Kistler, & Dolan, 1989) but also visual and lexical information to disentangle speech from the masker (i.e., the background noise) to perform the task, at least in situations where the visual information is coherent with the auditory information presented in the speech signal.

Finally, no significant interaction was obtained between lexical status and modality in children. Unlike adults (experiment 2; Fort et al., 2010), this study also indicates that even if 6–7-year-old children can rely both on lexical context and visual speech separately, they do not seem to combine these two sources of information to enhance vowel phoneme intelligibility in noise. The non-significant interaction between modality and lexicality suggests that the lexical bias (i.e., word superiority effect) is not greater when the visual information is available in the speech signal. Consequently it seems that visual speech only contributes to phonemic (or pre-lexical) processing until the later stages of childhood (i.e., after ages 9–10). It is likely that, during this period, visual speech may only spread activation towards pre-lexical units but not lexical representations. However, further research should be done to collect online measures (i.e., latencies) and to be able to directly compare performance in adults and children. Nonetheless, because our results also provided evidence that lexical knowledge biased vowel phonological processing in children, we may posit that they can process and rely separately on these two signals to perceive speech but they do not exploit them together to optimize lexical access process. Indeed, if visual speech only enhances pre-lexical processing during childhood but also contributes to lexical access in adulthood, we would expect to observe this shift during adolescence. Further research is in progress to determine the time period at which visual speech starts to influence lexical processing per se.

Funding

This research was supported by a governmental PhD fellowship.

Note

1. V = vowel, C = consonant.

References

- Ackroff, J. M. (1981). The interrelationship of verbal transformations, phonemic restorations and age. *Dissertation Abstracts International*, 42, 2106.
- Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research*, 32, 317–322.
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audio-visual integration of speech falters under high attention demands. *Current Biology*, 15, 839–843.
- Barutchu, A., Crewther, S., Kiely, P., Murphy, M., & Crewther, D. (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, 1–11.
- Benoit, 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.
- Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, 20, 539–542.
- 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.
- Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, 44, 209–220.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, 19, 141–177.
- Cutler, A., Sebastián-Gallés, N., Soler-Vilagelius, O., & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, 28, 746–755.
- Dupont, S., Aubin, J., & Menard, L. (2005). A study of the McGurk effect in 4- and 5-year-old French Canadian children. *ZAS Papers in Linguistics*, 40, 1–17.
- Erber, N. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research*, 12, 423–425.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46, 245–264.
- Fort, M., Spinelli, E., Savariaux, C., & Kandel, S. (2010). The word superiority effect in audiovisual speech perception. *Speech Communication*, 52, 525–532.
- Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 77–91.
- Ganong III, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 110–125.
- Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279.
- Green, K. P. (1998). The use of auditory and visual information during phonetic processing: Implications for theories of speech perception. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eyes II: Advances in the psychology of speechreading and auditory-visual speech* (pp. 3–26). Hove, UK: Psychology Press.
- Hallé, P. A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19, 463–481.
- Hockley, N., & Polka, L. (1994). A developmental study of audiovisual speech perception using the McGurk paradigm. *Journal of the Acoustical Society of America*, 96, 3309.
- Jerger, S., Damian, M. F., Spence, M. J., Tye-Murray, N., & Abdi, H. (2009). Developmental shifts in children's sensitivity to visual speech: A new multimodal picture word task. *Journal of Experimental Child Psychology*, 102, 40–59.
- Jerger, S., Martin, R., & Damian, M. (2002). Semantic and phonological influences on picture naming by children and teenagers. *Journal of Memory and Language*, 47, 229–249.
- Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, 34, 485–499.
- Kim, J., Davis, C., & Krins, P. (2004). Amodal processing of visual speech as revealed by priming. *Cognition*, 93, B39–B47.
- Kuhl, P., & Meltzoff, A. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138–1141.
- Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of language*. Oxford, UK: Blackwell.
- Locke, J. L. (1993). *The child's path to spoken language*. Cambridge, MA: Harvard University Press.
- Massaro, D. (1984). Children's perception of visual and auditory speech. *Child Development*, 55, 1777–1788.
- Massaro, D., Thompson, L., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology*, 41, 93–113.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36, 516–524.
- Nooteboom, S. G., & Doedeman, G. J. N. (1984). Speech quality and the gating paradigm. In M. P. R. van den Broecke & A. Cohen (Eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences: Utrecht* (pp. 481–485). Dordrecht, the Netherlands: Foris.
- Patterson, M., & Werker, J. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, 22, 237–247.
- Patterson, M., & Werker, J. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science*, 6, 191–196.
- Rosenblum, L., Schmuckler, M., & Johnson, J. (1997). The McGurk effect in infants. *Perception and Psychophysics*, 59, 347–357.
- Sams, M., Manninen, P., Surakka, V., Helin, P., & Kättö, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences: Effects of word meaning and sentence context. *Speech Communication*, 26, 75–87.
- Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, 110, 474–494.
- Sekiyama, K., & Burnham, D. (2008). Impact of language on development of auditory-visual speech perception. *Developmental Science*, 11, 306–320.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388, 381–382.

- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212–215.
- Summerfield, Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 3–51). London, UK: Erlbaum.
- Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76, 147–166.
- Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13, 480–484.
- Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242–248.
- Walley, A. C. (1988). Spoken word recognition by young children and adults. *Cognitive Development*, 3, 137–165.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392–393.
- Weikum, W., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, 316, 1159.
- Windmann, S. (2004). Effects of sentence context and expectation on the McGurk illusion. *Journal of Memory and Language*, 50(2), 212–230.
- Winer, B. J. (1970). *Statistical principles in experimental design*. London, UK: McGraw-Hill.

Appendix 1

Material used as experimental target words and target pseudo-words in experiment 1

Target words	Translation	Target pseudo-words
Bateau [bate]	Boat	Lateau [lato]
Cadeau [kado]	Gift	Madeau [mado]
Casser [kase]	To break	Dasser [dase]
Chanter [ʃäte]	To sing	Panter [päte]
Chapeau [ʃapo]	Hat	Tapeau [tapo]
Début [deby]	Beginning	Nébut [neby]
Dessus [dësy]	Over	Peussus [pësy]
Donner [done]	To give	Lonner [done]
Fermer [ferme]	To close	Termer [terme]
Garder [gazde]	To keep	Narder [naðde]
Gâteau [gato]	Cake	Nateau [nato]
Goûter [gute]	Snack	Nouter [nute]
Jeter [ʒäte]	To throw	Deter [dete]
Laver [lave]	To wash	Daver [dave]
Manteau [mäto]	Coat	Ganteau [gäto]
Monter [môte]	To take up	Lonter [löte]
Parler [parle]	To speak	Darler [daðle]
Perdu [peðdy]	Lost	Serdu [seðdy]
Tortue [toxti]	Turtle	Gortu [goxti]
Venue [vəny]	Coming	Leunu [ləny]

Note. Letters in bold represent target phonemes

Appendix 2

Material used as experimental target words and target pseudo-words in experiment 2

Target words	Frequency	Translation	Target pseudo-words
Bambou /bäbu/	1.32	Bamboo	Kambou /käbu/
Bateau* /bato/	106.55	Boat	Lateau* /lato/
Beaucoup /boku/	626	A lot	Docoup /doku/
Bisou /bizu/	13.99	Kiss	Risou /rizu/
Bureau /byro/	156.68	Desk	Gureau /gyro/
Cadeau* /kado/	98.09	Gift	Madeau* /mado/
Caillou /kaju/	4.11	Stone	Naillou /naju/
Casser* /kase/	9.05	To break	Dasser* /dase/
Cerveau /servo/	57.67	Brain	Perveau /pervo/
Chanter* /ʃäte/	48.12	To sing	Panter* /päte/
Chapeau* /ʃapo/	48.61	Hat	Tapeau* /tapo/

(continued)

(continued)

Target words	Frequency	Translation	Target pseudo-words
Choisi /ʃwazi/	58.19	Chosen	Roisi /ʁwazi/
Compter /kōte/	31.49	To count	Sonter /sōte/
Copie /kopī/	16.88	Copy	Dopie /dopi/
Crédit /kredi/	25.82	Credit	Brédit /bredi/
Debout /dəbu/	91.81	Standing	Kebout /kəbu/
Début* /deby/	109.88	Beginning	Nébut* /neby/
Défi /defi/	10.23	Challenge	Léfi /lefɪ/
Dessous /dəsū/	18.14	Under	Meussous /məsū/
Dessus* /dəsy/	111.19	Above	Peussus* /pasy/
Dîner /dine/	51.78	Dinner	Riner /rīne/
Donner* /done/	233.3	To give	Lonner* /done/
Drapeau /drapō/	14.66	Flag	Grapeau /grapō/
Fermer* /ferme/	13.72	Closed	Termer* /terme/
Folie /foli/	49.01	Craziness	Bolie /boli/
Fourmi /fʊrmī/	2.78	Ant	Dourmi /duʁmī/
Fusée /fyze/	6	Rocket	Dusée /dyze/
Fusil /fyzi/	36.52	Shotgun	Rusi /ʁyzi/
Garder* /gaʁde/	21.71	To keep	Narder* /naʁde/
Gâteau* /gato/	42.33	Cake	Nateau* /nato/
Geler /ʒale/	3.61	To freeze	Neler /nale/
Génie /ʒenī/	34.65	Genius	Ménie /meni/
Genou /ʒənu/	11.43	Knee	Veunou /vənu/
Goûter* /gute/	1.84	Snack	Nouter* /nute/
Jaloux /ʒalu/	29.87	Jealous	Paloux /palu/
Jeter* /ʒete/	192.18	To throw	Deter* /dete/
Jeudi /ʒœdi/	24.58	Thursday	Feudi /fœdi/
Joli /ʒoli/	94.55	Pretty	Roli /ʁoli/
Jouer /ʒue/	1.9	To play	Pouer /pue/
Journée /ʒuʁne/	165.35	Day	Bournée /buʁne/
Jumeau /ʒymo/	1.27	Twin boy	Lumeau /lymo/
Laver* /lave/	9.73	To wash	Daver* /dave/
Lundi /lədi/	36.01	Monday	Vundi /vədi/
Lycée /lisē/	41.96	High school	Rissée /risē/
Manteau* /māto/	36.16	Coat	Ganteau* /gāto/
Mardi /maʁdi/	22.38	Tuesday	Sardi /saʁdi/
Menu /mœny/	9.87	Menu	Renu /ʁœny/
Merci /meʁsi/	378.44	Thank you	Lerci /leʁsi/
Monter* /mōte/	6.11	To take up	Lonter* /lōte/
Niveau /nivo/	45.46	Level	Tiveau /tivo/
Nouveau /nuvo/	170.28	New	Gouveau /guvo/
Panneau /pano/	9.87	Sign	Tanneau /tano/
Parler* /paʁle/	15.82	To speak	Darler* /daʁle/
Partout /paʁtu/	141.94	Everywhere	Nartout /naʁtu/
Perdu* /peʁdy/	217.23	Lost	Cerdu* /meʁdy/
Pérou /peʁu/	0.01	Peru	Guerou /geʁu/
Petit /pəti/	573.72	Small	Metit /məti/
Plateau /plato/	15.73	Tray	Clateau /klato/
Prévu /pʁevy/	55.54	Predicted	Crévu /kʁevy/
Radis /ʁadi/	1.81	Radish	Fadis /fadi/
Reçu /ʁəsy/	76.46	Receipt	Beussu /bəsy/
Rendu /ʁədy/	48.31	Return	Lendu /lədy/
Revue /ʁəvy/	7.79	Magazine	Pevue /pəvy/
Rideau /ʁido/	10.81	Curtain	Sideau /sido/
Saler /sale/	9.2	To salt	Naler /nale/
Samedi /samədi/	44.51	Saturday	Gamedi /gamədi/
Secoue /səku/	4.77	To shake	Beucoue /bəku/
Série /səri/	33.34	Series	Térie /təri/
Soirée /sware/	94.36	Evening	Doiree /dware/
Sonnerie /sɔnri/	6.03	Ringing	Monnerie /monri/
Sortie /sɔrti/	42.58	Exit	Dortie /doʁti/
Souris /sʊri/	21.94	Mouse	Gouris /gʊri/

(continued)

(continued)

Target words	Frequency	Translation	Target pseudo-words
Surtout /syʁtu/	148.66	Above all	Durtout /dyʁtu/
Tabou /tabu/	0.54	Taboo	Nabou /nabu/
Tapis /tapi/	20.13	Carpet	Rapis /ʁapi/
Tenu /təny/	27.1	Clothes	Penue /pəny/
Tester /teste/	93.68	To test	Mester /meste/
Têtu /tety/	6.17	Stubborn	Retu /ʁəty/
Tirer /tire/	10.35	To pull	Guirer /giʁe/
Tissu /tisy/	9.21	Material	Bissu /bisy/
Tomber /tõbe/	180.25	To fall	Romber /ʁõbe/
Tortue* /tɔʁty/	4.0	Turtle	Gortue* /gɔʁty/
Toupie /tupi/	1.5	Top	Foupie /fupi/
Tribu /triby/	0.01	Tribe	Fribu /friby/
Vaisseau /veso/	67.11	Vessel	Naisseau /neso/
Vaudou /vodu/	2.92	Vaudoo	Paudou /podu/
Vendu /vãdy/	37.62	Sold	Nendu /nãdy/
Venue* /vəny/	93.14	Coming	Lenue* /ləny/
Verrou /veʁu/	3.54	Bolt	Terrou /teʁu/
Voler /vole/	3.88	To fly	Noler /nole/

Notes. Frequency (in occurrences per million) indicates word frequency. Letters in bold represent target phonemes. Stimuli with * were used as experimental targets in experiment I.