European Journal of Cognitive Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pecp20

The Role of Lip-reading and Cued Speech in the Processing of Phonological Information in French-educated Deaf Children

Jesus Alegria , Brigitte L. Charlier & Sven Mattys
Published online: 21 Sep 2010.

To cite this article: Jesus Alegria , Brigitte L. Charlier & Sven Mattys (1999) The Role of Lip-reading and Cued Speech in the Processing of Phonological Information in French-educated Deaf Children, European Journal of Cognitive Psychology, 11:4, 451-472, DOI: 10.1080/095414499382255

To link to this article: http://dx.doi.org/10.1080/095414499382255

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 1999, *11* (4), 451–472

The Role of Lip-reading and Cued Speech in the Processing of Phonological Information in French-educated Deaf Children

Jesus Alegria
Laboratoire de Psychologie Expérimentale, Université Libre de Bruxelles, Belgium

Brigitte L. Charlier
Centre Comprendre et Parler, Brussels, Belgium

Sven Mattys
Departments of Psychology and Cognitive Science, Johns Hopkins University, Baltimore, USA

Deaf children exposed to Cued Speech (CS: system designed to reduce lip-reading ambiguity) either before age 2 ("early") or later at school ("late") were presented with words and pseudowords with or without CS. The first goal was to examine the effects of adding CS to lip-reading on phonological perception. Results showed that CS substantially improved performance suggesting that CS corrects for lip-reading ambiguities. CS effects were significantly larger in the "early" than the "late" group, particularly with pseudowords. The second goal was to establish the way in which lip-reading and CS combine to produce unitary percepts. To address this issue, two types of phonological misperception resulting from CS's structural characteristics were analysed; substitutions based on the similarity between CS units, and intrusions of a third syllable for bisyllabic pseudowords requiring three CS units. The results showed that the frequency of such misperceptions increased with CS. The integration of CS and lip-read information is

Requests for reprints should be addressed to Jesus Alegria, Université Libre de Bruxelles (CP: 191) – Laboratoire de Psychologie Expérimentale, 50 Avenue F.D. Roosevelt, 1050 Brussels, Belgium. Email: alegria@ulb.ac.be

This work was supported by the Belgian Fonds de la Recherche Scientifique Médicale (F.R.S.M. – convention No. 3.4571.94). Thanks are due to Mayra Manzano and Thierry Claes for their participation in data collection and for discussing the results with us. Thanks are also due to Olivier Périer, Jacqueline Leybaert, Catherine Hage, and Ruth Campbell, for helping to clarify some of the ideas presented here.

discussed as a function of CS's structural characteristics and the amount of exposure to CS.

Phonological codes support numerous fundamental cognitive activities in normal-hearing people. These activities include speech perception and production, as well as reading and writing. Research with deaf people has been concerned with the development of such phonological codes in the absence of auditory information. Several experimental paradigms have demonstrated that a significant proportion of deaf youngsters possess phonological codes and use them in a manner similar to their unimpaired counterparts (Alegria, Leybaert, Charlier, & Hage, 1992; Campbell & Wright, 1988; Conrad, 1979; Dodd & Hermelin, 1977; Leybaert & Alegria, 1993; Leybaert, Alegria, Hage, & Charlier, 1998; Oller & Eilers, 1988). Similarly, experiments focusing on the mechanisms of reading and spelling in the deaf have shown that these activities involve the use of phonological representations (Alegria & Leybaert, 1991; Hanson, Shankweiler, & Fischer, 1983; Leybaert, 1993; Leybaert & Alegria, 1995). Although these studies suggest that phonological representations can develop in the absence of auditory information, the origin of those representations remains to be established. Lip-reading is obviously an important determinant to consider.

LIP-READING

Lip-reading undoubtedly contributes to the development of phonological representations in people who are deaf. In fact, a vast literature shows that lip-reading improves speech recognition in both deaf and normal-hearing people. Until the mid-1970s, the function of lip-reading was believed to be limited almost exclusively to assisting speech perception in poor auditory conditions (Binnie, Montgomery, & Jackson, 1974; Erber, 1969, 1974; Sumby & Pollack, 1954). However, data collected during the past 20 years with normal-hearing populations have indicated that, when a speaker's face is seen, the visual information accompanying speech becomes part of an audio-visual compound that enhances encoding of the speech signal. McGurk and MacDonald (1976) showed that seeing a face pronouncing the syllable /ga/ while simultaneously listening to an auditory /ba/ produced a perceptually clear /da/. These results have since then been confirmed and expanded in several studies (e.g. Campbell, Dodd, & Burnham, 1998; Massaro, 1987, 1989). The McGurk effect radically modifies the old view that, in the hearing, speech processing is a purely auditory phenomenon.

One obvious reason why the speech processing apparatus incorporates lip-reading information is that human beings experience, from their earliest infancy, a highly correlated audio-visual speech signal. Results obtained with the paradigm of visual preference have shown that infants prefer to look at a face executing articulatory gestures that match simultaneous auditory information over articulatory gestures that do not match the auditory signal. This linkage indicates that the visual (lip-reading) information is processed in a linguistically relevant manner as early as five months after birth (Burnham, 1998; Kuhl & Meltzoff, 1982; MacKain, Studdert-Kennedy, Spieker, & Stern, 1983).

Thus, because lip-reading seems to be integrated into the speech perception process from the earliest days of life, some abstract representations common to both auditory and visual speech information may exist (Green, 1998; Liberman & Mattingly, 1985; Summerfield, 1987, 1991). In turn, phonological codes are likely to have a visual (lip-reading) dimension, at least in individuals hearing normally. This visual dimension of speech is important because it makes the phonological representations of the deaf appear comparable to those of the hearing. However, because lip-reading alone is an insufficient support for deaf people, methods for reducing lip-reading ambiguities have received considerable attention.

MANUAL AIDS TO LIP-READING: CUED SPEECH (CS)

Research has shown that the phonological representations in children with profound hearing loss are at least initially based on lip-reading (Dodd, 1976, 1987; Dodd & Hermelin, 1977; Leybaert & Alegria, 1995). However, for the vast majority of deaf children, lip-reading alone cannot adequately trigger language acquisition because the visual information available in speech is far more ambiguous than the auditory information (e.g. the voiced/voiceless distinction is not visible). Moreover, successful lexical development requires a rather systematic association between referents and their corresponding phonological strings. Lip-reading alone does not provide this association as reliably as the auditory signal does. Due to this shortage, speech-language pathologists and teachers for the deaf have attempted to reduce ambiguities in lip-reading with supplementary auditory, visual, and tactile information.

In this context, Cued Speech (CS), a system of hand gestures, was developed in an attempt to help deaf children decode speech by eliminating lip-reading ambiguities (Cornett, 1967; see Périer, 1987, for a comparison between CS and other similar systems). When using CS, the speaker forms various hand shapes at positions near and around the mouth so that the listener can see the speech information conveyed simultaneously by the lips and the hand. In the French version of CS, the hand can adopt eight hand shapes and be placed at five hand positions around the mouth. Hand shapes are intended to clarify consonants whereas hand positions disambiguate vowels.

As shown in Fig. 1, consonants and vowels are grouped into sets of two or three, with each set having a specific hand shape or hand position. Sets of consonants and vowels are represented by one hand shape or one hand position, respectively. Within a set, phonemes are easily discriminable through lip-reading. Conversely, items that are difficult to discriminate through lip-reading belong to different sets. For example, a particular hand shape is shared by the phonemes /p, d, ʒ/, which are easily distinguished on the basis of lip-read cues. In contrast, bilabials (e.g. /b, m, p/), which are difficult to discriminate from each other through lip-reading, are represented by different hand shapes (Fig. 1). The same principle holds with vowels: A particular hand position is shared by vowels with high visual discriminability (e.g. /i, ɔ̃, ã/ and /a, o, ə/), whereas vowels that share similar lip shapes (e.g. the rounded French vowels /y, u, o, ɸ/) are coded by different hand positions. It is worth mentioning that, compared to English, French has a small number of diphthongs. This aspect makes French CS relatively simpler than English CS given that coding diphthongs in CS requires several (and rapid) hand position changes.

As a result of this organisation, the simultaneous lip-read and CS information produced for, say, a CV syllable (i.e. one hand shape in one particular hand position) provides unambiguous information about the identity of this particular syllable. We will refer to the CV syllable structure as the "canonical" structure for the CS system because one such syllabic unit corresponds to one CS unit. For the other syllabic structures, CS has additional features. For example, a vocalic syllable (V) is represented by the hand position corresponding to the vowel combined with the "neutral" hand shape (the one found with consonants /f, t, m/, see Fig. 1). Likewise, syllables bearing several consonants, such as CCV and CVC, are coded using the "neutral" hand position (the one found with vowels /a, o, ə/). Figure 2, which displays the different syllabic structures used in the present experiment, illustrates various CS units. For instance, the vocalic syllable /ə/ is coded at the /ə/ hand position using the "neutral" hand shape. The syllable /bli/ is coded with two CS units: The first one is the /b/ hand shape in the neutral hand position while the second one is the canonical CV unit for the /li/ part of the syllable.

Critically, these structural characteristics of CS could potentially create phonological percepts not present in the input because the number of CS units does not always coincide with the number of spoken syllables. This inequity could produce perception of extra syllables. Further, CS could also cause confusions among consonants sharing the same hand shape

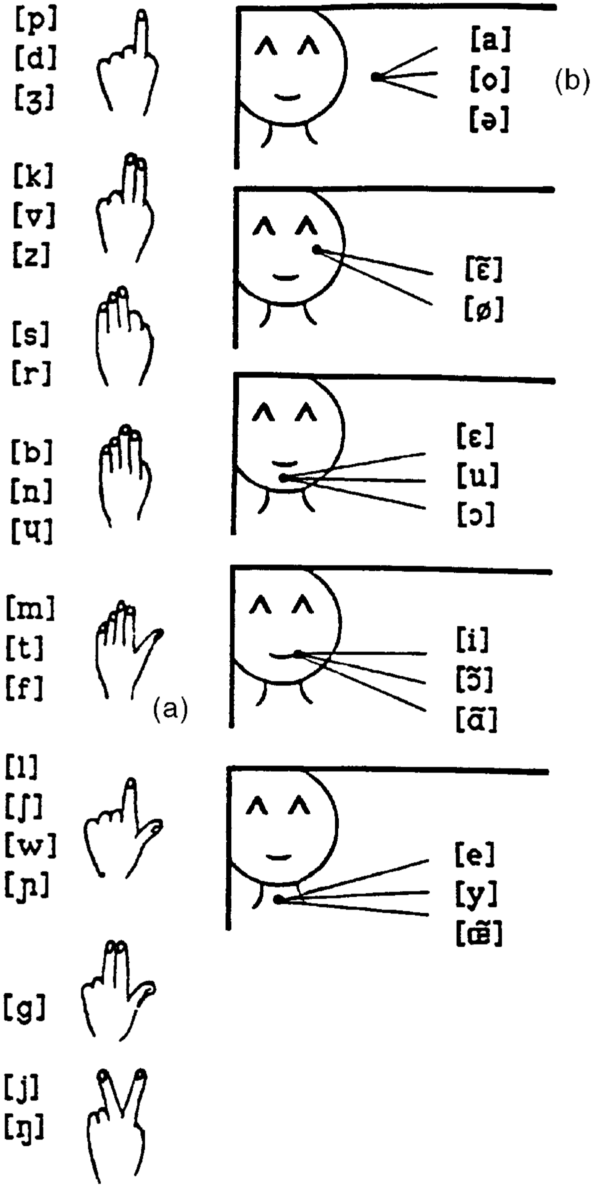

FIG. 1. French version of Cued Speech: (a) "neutral" hand shape used to code isolated vowels and (b) neutral hand position used to code consonants other than onsets of CV syllables.

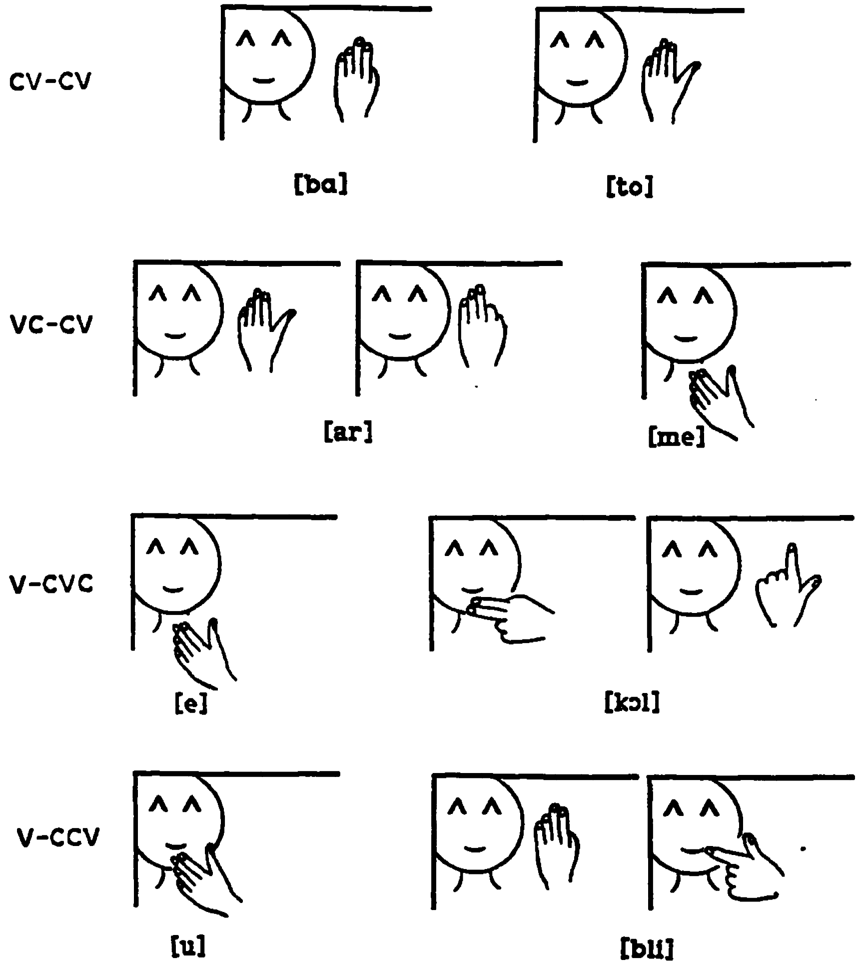

FIG. 2. Examples of the phonological structures used in the experiment and their corresponding Cues.

(e.g. confusing /d/ with /ʒ/ or /p/) or among vowels sharing the same hand position (e.g. confusing /o/ with /ə/ or /a/). We will refer to these two types of errors as "CS errors".

Errors due to the specific characteristics of CS might depend, in large part, on the way CS information is processed relative to lip-read information. For example, the only way to discriminate /da/ from /ʒa/, which share the same hand shape, is to decode the labial difference in their initial consonant. Similarly, the distinction between the syllables /lo/ and /lə/, which share the same hand position, depends on the lip-read information about the vowel. Therefore, the occurrence of CS-related errors would suggest that the CS information was decoded without the integration of lip-reading information. One of the goals of the present experiment is to explore this possibility. This analysis represents an important step towards understanding how CS and lip-reading information combine to produce a unique phonological compound.

THE ROLE OF CUED SPEECH IN SPEECH PROCESSING

Nicholls and Ling (1982) examined the role of CS in speech recognition. In their research, a group of deaf youngsters who had been exposed to CS in school for at leats four years were presented with lip-read sentences with and without CS. The results showed that the addition of CS produced a substantial improvement in comprehension. Similarly, a group of deaf children was examined in a sentence understanding task (Charlier, Hage, Alegria, & Périer, 1990; see Périer, Charlier, Hage, & Alegria, 1987 for an earlier report). In this study, some children had been exposed to CS since infancy (the "early" group). The other children of the sample had used CS exclusively in school from age 5 or 6 onwards (the "late" group). The two groups were presented with lip-read sentences with and without CS. For each sentence, the children were asked to choose, out of a set of four drawings, the one that best corresponded to the content of the sentence. The authors found that the presence of CS generated a significant improvement in both groups, but the improvement was greater in the "early" than in the "late" group. When "early" and "late" groups of children were paired off for the duration of CS exposure, the advantage for the "early" group was still present, despite the fact that the children in the "early" group were younger than those in the "late" group.

The present experiment has two goals. The first one is to further document the notion that CS improves lip-reading. To this end, words and pseudowords, presented with and without CS, were presented to children of different ages and with different durations of exposure to CS. The accuracy of their transcription was analysed for all conditions. The second goal is to explore the way in which CS and lip-reading information combine. To do so, we analyse the patterns of CS errors in canonical and non-canonical structures.

METHOD

Participants

The participants were 31 prelingually deaf French-communicating children and adolescents. All participants showed normal intelligence levels (IQ was routinely evaluated in school with non-verbal tests). They attended classes corresponding to their age. None exhibited learning or language disabilities. Their teachers were asked to provide an evaluation (on a six-point scale from 0, very poor, to 6, very good) of the children's reading and spelling abilities relative to normal expectations for deaf children attending specialised schools. The evaluations revealed that the children's spelling abilities were sufficient for the elementary phoneme-to-grapheme transcription task required in the present experiment. Participants were split into two groups as a function of the age at which they had been exposed to CS: "early" and "late".

The seven participants (six male and one female) belonging to the "early" group had been exposed to CS at home before age 2 and, since then, had used CS systematically to communicate with their parents. The average age in the "early" group was 10;9 (years; months), ranging from 8;6 to 12;0. Duration of CS exposure was 9;5 (years; months), ranging from 6;11 to 11;1. The average hearing level (calculated on 500, 1000, and 2000 Hz pure tones) was 102.9 dB HL, ranging from 100 to 115 dB HL. All children in this group wore two hearing aids. One of the participants had corrected vision.

The "late" group consisted of 24 participants (15 male and 9 female). The mean age in the "late" group was 15;9 (years; months), ranging from 11;8 to 19;10. All participants started using CS between age 5 and age 9 (i.e. at the end of pre-school or at the beginning of primary school). Except for two children who occasionally used CS to communicate with their parents, all participants in the "late" group used CS exclusively at school to communicate in the classroom. The average duration of CS exposure in this group was 6;5 (years; months), ranging from 3;0 to 10;11. The average hearing level as 100 dB HL, varying from 85 to 120 dB HL. All of the participants in the "late" group used wore two hearing aids except for one participant who wore a cochlear implant (implanted four years prior to the present experiment). Seven participants had corrected vision. The sample characteristics are reported in Table 1.

Stimuli and Conditions

All stimuli, words, and pseudowords, were four-phoneme-long bisyllables. Four different structures were examined: CV-CV, VC-CV, V-CVC, and

TABLE 1
Sample Characteristics

	Early	*Late*
n	7	24
Age		
mean	10;9	15;9
range	(8;6–12;0)	11;8–19;10)
Exposure Duration		
mean	9;5	6;5
range	(6;11–11;1)	(3;0–10;11)
Hearing Loss		
mean	102.9	100.0
range	(100–115)	(85–120)
Reading-Spelling		
mean	4.6	3.7
range	(3–6)	(2–6)

Number of participants (*n*), mean age (years; months) and range, Duration of Exposure to CS, Hearing Loss (dB HL), and Reading-Spelling marks (evaluated by the teachers in a scale from zero: very poor, to six: very good), per group.

V-CCV, with eight words and eight pseudowords per structure (the entire list can be seen in the Appendix). Pseudowords were generated by combining syllables of structurally similar words. For example, in condition CV-CV, the pseudoword /by-vɸ/ was created by pasting the first syllable of the word /by-ʀo/ (desk) and the second syllable of the word /ʃə-vɸ/ (hair). Words and pseudowords were presented in separate blocks of 64 items each. The two possible orders of presentation were counterbalanced across participants. Each item appeared twice: once with lip-reading alone, and once with lip-reading plus CS (the latter condition is henceforth labelled CS). Half the items appeared first with lip-reading alone and then with CS. The other half appeared in the opposite order. The number of items inserted between the two versions of the same stimulus varied between 29 and 36 (mean = 32). All items were videotaped by a female pathologist specialising in speech and language, who was fluent in CS (she had over 10 years of daily CS practice) and familiar to the children. The same tape was presented to all of the participants. To signal the onset of an item, the speaker looked down at the floor. Upon pronunciation, she raised her head in the upright position and produced the item. Then, her face tilted down again and the tape was stopped. All items were pronounced at a normal rate (less than one second per item). They were exclusively visual (the videotape was played without sound).

Procedure

The participants were tested either individually or in groups of two or three in a quiet and dimly light room. They were seated in front of a TV screen 65 cm in diagonal. The distance between the children and the TV screen was ~ 120 cm. First, participants were shown a series of practice items with the same phonological structures as the test items. The training session ended as soon as the participants showed that they understood the task. Before each experimental block of trials, they were informed about the block's lexical status: words or pseudowords. After each trial, the participants gave a written transcription of the item, without any time constraints. The next item was presented as soon as all the children had written the response and were looking at the screen again. For each item, the participants were asked to provide a response as complete as possible. They were encouraged to guess if they were uncertain.

At the beginning of the session, the experimenter told the children that the standard spelling of the words did not matter. Participants were given examples showing acceptable spellings of various phonemes, words, and pseudowords. French spelling, like English spelling, involves numerous irregularities at the phoneme-to-grapheme transcription level. For example, the phoneme /o/ is spelled "eau" in the word "bureau" /by-ʀo/ (desk) but spellings like "o", "au", and "aud", found in other words, are also pronounced /o/. In this experiment, all such spellings were considered correct. Interpreting the target phoneme from deviant spellings, was not problematic since French, unlike English does not present irregularities at the grapheme-to-phoneme transcription level. The phonology of any spelling can be deduced without ambiguity. For example, if a participant gave the response "buraud" (instead of the standard spelling "bureau"), she or he could be unambiguously credited with having correctly perceived the final phoneme /o/. This property allowed for evaluation of the participant's written response at a phonological level regardless of the spelling.

RESULTS

Effects of CS on Phonological Processing

The first goal of the present experiment was to determine the role played by CS in the processing of words and pseudowords and its relationship to the age at which children were exposed to CS for the first time as well as the amount of CS exposure they had had in their life. In this section, we present the results in the Lexicality (word vs pseudoword) and Cueing (lip-reading vs CS) conditions, broken down into the "early" and "late" groups. The mean percentages of correct responses appear in Fig. 3.

An analysis of variance was performed on the accuracy data, factoring Group ("early" vs "late") as a between-subject factor, and Lexicality (words vs pseudowords), Syllabic Structure (CV-CV, VC-CV, V-CVC, and V-CCV), and Cueing (lip-reading vs CS) as within-subject factors. Although the main effect of Group ("early" vs "late") did not reach significance, $F < 1$, CS was found to significantly improve the overall identification performance, $F(1, 29) = 134.38$, $p < .001$. This effect was greater in the "early" than in the "late" group, as shown by the interaction between Group and Cueing, $F(1, 29) = 6.39$, $p < .02$. None the less, the improvement in the "late" group was still significant, $F(1, 21) = 69.43$, $p < .001$.

The results also show that words were identified better than pseudowords, $F(1, 29) = 203.62$, $p < .001$. A significant interaction between Lexicality and Group, $F(1, 29) = 7.43$, $p < .05$, indicated that pseudowords were better identified by the "early" than the "late" participants, $F(1, 29) = 4.12$, $p = .052$. This difference did not appear with words, $F < 1$. The lexicality effect applied to both the CS and lip-reading conditions, as shown by the absence of a three-way interaction between Lexicality, Cueing, and Group, $F < 1$.

Finally, Syllabic Structure also generated reliable differences, $F(3, 87) = 20.32$, $p < .001$. The only factor that interacted significantly with

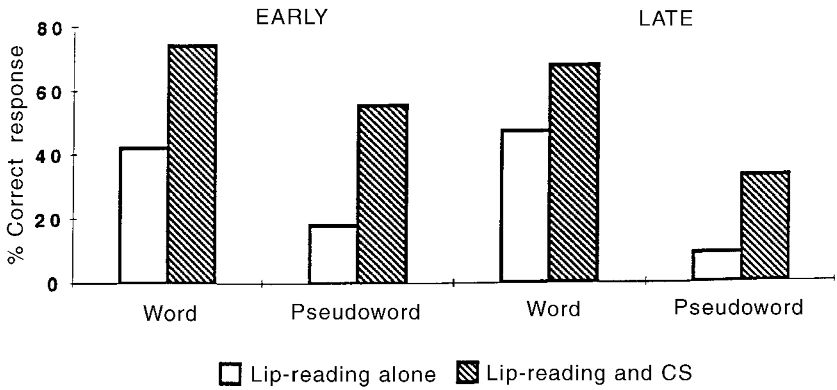

FIG. 3. Mean percentage of correct responses for words and pseudowords per condition (lip-reading alone vs lip-reading plus CS). The figure shows the results obtained in the "early"

Structure was Lexicality, $F(3, 87) = 16.81$, $p < .001$. For clarity's sake, Structure is not represented in Fig. 3. The percentage of correct responses varied between 48.8% (V-CVC) and 33.5% (V-CCV). The only theoretically motivated prediction regarding differences between structures was the superiority of the canonical condition in the CS condition. However, this effect did not emerge. Given that the influence of Syllabic Structure is specifically related to the hypothesis of CS-generated errors, this variable will be considered in more detail in the next section.

Although the absence of a global significant difference between "early" and "late" groups seems surprising at first, several sample characteristics may have contributed to this result. For example, the "early" group involved participants considerably younger than the "late" group, and all the "early" participants were in primary school, whereas 18 of the 24 participants in the "late" group had reached the secondary level. Moreover, the difference between the "early" and "late" groups was not exclusively due to the age of first CS exposure. This factor was partly confounded with the duration of the exposure to CS. Children in the "early" group had generally been exposed to CS not only at an earlier age but also during a longer period of time than those in the "late" group. Finally, hearing losses in the "late" group were both lower and more variable relative to those of the "early" group. Part of the differences between groups may arise from the fact that some children in the "late" group might be better at exploiting their residual hearing resources and less motivated to use CS.

In order to assess the impact of such differences between the "early" and "late" groups on the results, three factors were introduced into the previous analyses as covariates: Age, Duration of Exposure to CS, and Hearing Loss (mean loss for 500, 1000, and 2000 Hz tones in the best ear). The results showed that only Age emerged as a significant contributor to the performance, $F(1, 28) = 9.10$, $p < .01$; Duration of Exposure to CS, $F(1, 28) = 2.31$, $p > .10$, ns; Hearing Loss, $F < 1$. An interesting consequence of introducing Age as a covariate was that Group, which was not a significant factor in the ANOVA, clearly emerged, $F(1, 28) = 7.46$, $p < .05$, with "early" children performing better than "late" children. The interaction between Group and Cueing remained significant, $F(1, 27) = 8.03$, $p < .01$.

Furthermore, because Age and Duration of Exposure to CS were strongly correlated to each other (especially in the "late" group), we ran another set of control analyses in which these two factors were introduced as covariates simultaneously. The results of these analyses showed that Age, but not Duration of Exposure to CS, was significant, $F(1, 27) = 11.05$, $p < .01$, and $F(1, 27) = 4.08$, $p = .13$, respectively). Finally, it is worth mentioning that the Group effect and the Group by Cueing interaction remained significant, $F(1, 27) = 5.89$, $p < .01$, and $F(1, 27) = 6.72$, $p < .05$, respectively.

Cued Speech Processing: Effects of Some Specific CS Characteristics

To address the second issue of the experiment—the possibility that CS generates specific errors, we focused on the incorrect responses related to CS characteristics. The relevant analyses were done exclusively on pseudowords so as to factor out possible lexical influences. Two types of evidence were explored. The first one entailed tracking down the presence of substitutions of segments induced by CS characteristics (e.g. giving a /t/ response to an /f/ input because the consonants share the same hand shape). The second one concerned the influence of the number of CS units on the number of syllables reported.

Substitution Errors Based on CS Structure. In the following analyses, substitutions of a segment for one that shares a CS parameter (i.e. hand shape or hand position) are referred to as CS substitutions. We examined the frequency of CS substitutions in the CS condition and in the "lip-reading alone" condition.

We analysed the stimulus-response matrix of the pseudoword data and determined the frequency of CS substitutions relative to the total number of errors (i.e. any erroneous phoneme other than those belonging to the specific CS hand shape/position) for each phoneme. Consonants and vowels were analysed separately. For higher accuracy in coding the errors, we considered only the responses with a syllabic structure identical to that of the stimulus (V for V, VC for VC, etc.). The results are presented in Fig. 4. An analysis of variance was run with Group ("early" vs "late") as a between-subject factor, and Cueing (lip-reading vs CS) and Type of Phoneme (vowels vs consonants) as within-subject factors.

The Cueing factor failed to reach the conventional significance level, $F(1, 29) = 3.22$, $p = .063$. Neither Group nor Type of Phoneme was significant (Fs < 1). However, there was a tendency for CS substitutions to appear with consonants, $F(1, 29) = 3.56$, $p = .059$, but not with vowels, $F(1, 29) = 1.59$, $p > .10$.

The Importance of CS Units in Syllabic Segmentation. If the number of CS units affects the number of perceived syllables, then non-canonical syllabic structures (VC-CV, V-CVC, and V-CCV), which have three CS units, should produce a greater number of three-syllable responses than the canonical structure (CV-CV), which has only two. Figure 5 shows the

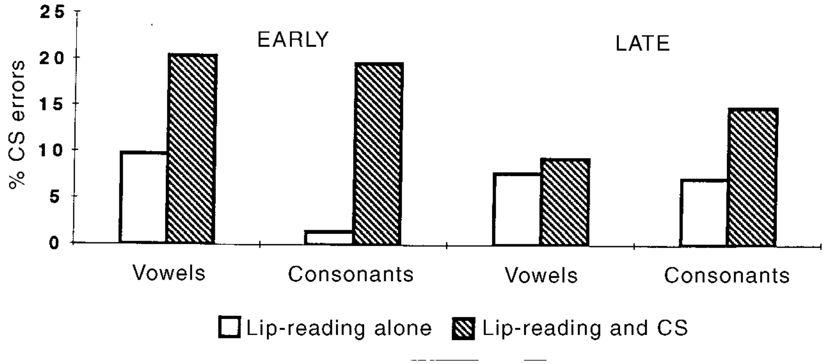

FIG. 4. Mean percentage of CS errors per group and per condition (lip-reading alone vs lip-reading plus CS). CS errors include confusions between segments that share the same hand position (vowels) or the same hand shape (consonants).

mean number of three-syllable responses as a function of Cueing in the canonical and non-canonical conditions, per group of subjects.

An analysis of variance was performed on the number of three-syllable responses with Group ("early" vs "late") as a between-subject factor and Cueing (lip-reading vs CS) and Structure (canonical: CV-CV vs non-canonical: a pool of the three others) as within-subject factors. The results suggest that the introduction of CS globally reduces the frequency of three-syllable responses, as shown by the Cueing effect, $F(1, 29) = 19.30$, $p < .001$. Structure was not significant, $F(1, 29) = 1.04$, ns, but its interaction with Cueing reached significance, $F(1, 29) = 5.66$, $p < .05$. This interaction indicates that the reduction of three syllable responses

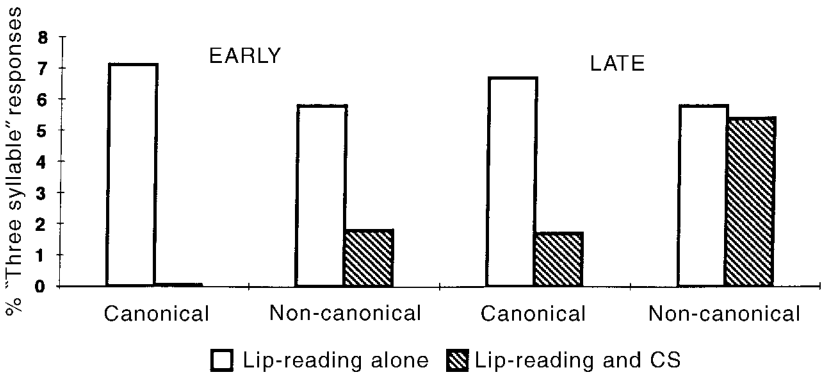

FIG. 5. Mean percentage of three-syllable responses per structure (canonical (CV-CV) vs non-canonical (VC-CV, V-CVC, and V-CCV)), per condition (lip-reading alone vs lip-reading plus CS), and per group.

was greater in the canonical than in the non-canonical condition; Cueing reduced the tendency to produce three-syllable responses in the canonical condition but not in the non-canonical condition, $F(1, 29) = 19.20$, $p < .001$ and $F(1, 29) = 1.78$, ns, respectively. Finally, the Group factor was non-significant, $F(1, 29) < 1$, as were its interactions with the other factors.

DISCUSSION

The motivation of the present experiment was twofold. The first goal was to examine the effects of adding CS information (Cues) to lip-reading information in the decoding of words and pseudowords by deaf children who had different experiences with CS. The second goal was to further understand how lip-reading and CS combine to produce a unitary percept. To address this issue, we examined cases in which the CS information was apparently not properly combined with the lip-read information, resulting in perceptual errors dubbed CS errors. Both the positive and the negative effects of CS are discussed later.

Lips and Cues Combination

The results obtained demonstrate that the addition of Cues to lip-reading substantially improves speech processing performances among the deaf children who use CS. Moreover, the improvement was more pronounced in children exposed to CS at home from an early age than in children exposed to CS later in their lives and solely at school. Such a discrepancy remained unchanged when the duration of exposure to CS was controlled. Thus, the initial age of exposure to CS turns out to be a critical factor. This finding strongly concurs with previous results (Charlier et al., 1990; Nicholls & Ling, 1982).

The present data also support the idea that successful labial/CS integration depends on two important cognitive capacities. First, deaf people must possess a mechanism responsible for extracting information from the lip-reading input and, when present, from CS information. This mechanism would enable them to elaborate an appropriate phonological input code. Second, adequate labial/CS processing necessitates a set of stored lexical entries accessible via the phonological input code so elaborated. The distinction between the extraction mechanism and the lexical entries is important because some children with hearing loss could possess lexical entries that are not accessible through the phonological code. Such a dissociation could happen in children whose lexical entries emerged through sign language, which typically does not transmit conventional phonological information. Based on the lexical effect observed across all groups, we presume that the participants in our experiment did possess and use phonologically accessible lexical entries that could be reached even with partial phonological input information. Access to the lexicon was obviously impossible in the case of pseudoword identification.

The performance disparities between groups might reflect differences in the quality of the phonological code attached to the lexical entries presented, or alternatively, differences in the ability to extract phonological input information. For example, deaf children could fail to identify the word /bu-RO/ (desk) because their lexical entry for this word does not contain all the necessary phonemic specifications. In contrast, the word could not be identified because processing of the speech signal has failed, even though the input word is correctly specified at the lexical level.

The pseudoword task was intended to explore more directly the pure bottom-up phonological processing mechanisms. The superiority of the "early" over the "late" group in decoding pseudowords suggests that children exposed to CS early in their life develop a more efficient phonetic processor than children learning CS later. Although the effects of age and/or schooling were substantial in the "late" group's pseudoword processing performance, the gap between "early" and "late" children remained considerable when this variable was partialled out.

When Lips and Cues Fail to Combine

The present research examined the possibility that the structural characteristics of CS may, under some circumstances, create incorrect perception of the speech input. Our data provided some evidence in support of this claim, which indicates that CS is sometimes interpreted independently of lip-reading. Two potential types of CS errors were examined: CS substitutions and errors in the number of perceived syllables.

Cued Speech Substitution Errors. Although statistical analyses between the number of CS substitutions with and without CS failed to show significant differences, the data revealed a tendency to make more CS substitutions when CS information was added to the labial input. This tendency was particularly marked with consonants and in the "early" group. However, the numbers of CS substitutions and non-CS substitutions found in the present sample are too small to draw definitive conclusions.

Despite the limited impact of CS errors on the general improvement provided by CS, the existence of CS errors is important on theoretical grounds because such errors could constitute an important source of information about the way in which CS affects phonological processing and how CS and lip-reading combine.

Number of Cued Speech Units and Syllable Perception. The second type of CS-induced errors we examined relates to the lack of a one-to-one match between the number of actual syllables of the input and the number of corresponding CS units. Only CV syllables translate into single CV units. VC, CVC, and CCV syllables are coded using two CS units. Our experiment sought to discover whether participants would be misled by this structural discrepancy, and hence, in the case of non-canonical syllables, perceive more syllables than those really present. Data consistent with this hypothesis would unambiguously indicate that the number of CS units is (or at least can be) interpreted autonomously from the labial information. The number of three-syllable responses was examined in two-syllable items made of canonical versus non-canonical structures. Contrary to the "negative-CS-effect" hypothesis, our results revealed that the number of three-syllable responses was smaller in the CS condition than in the "lip-reading alone" condition. However, this difference was reliable only in the canonical structure. No significant difference was found for the non-canonical structures. This finding suggests that CS helps determine the exact number of syllables in canonical CS structures whereas, in non-canonical structures, where the number of Cues is greater than the number of syllables, the role of CS is partly hindered by the tendency to interpret extra Cues as extra syllables.

Perhaps, the perception of extra syllables in non-canonical structures results from the participants' tendency to interpret neutral Cues as real phonemes. For example, in the syllable /aʀ/ (Fig. 2), the vowel /a/ is coded with the neutral hand shape (shared with the consonants /m, t, f/) and the consonant /ʀ/ is coded with the neutral hand position (shared with the vowels /a, o, ə/). In consequence, /aʀ/ could be erroneously interpreted as /ta-ʀo/ or /fa-ʀa/, which both have two syllables. Thus, neutral Cues that convey no segmental meaning can be correctly inferred only if they are considered in combination with the lip-reading information. In the present experiment, we occasionally observed cases of neutral Cue intrusion (e.g. perceiving /bə-li/ instead of /bli/). Research has shown that several segmental contexts can favour the occurrence of vowel intrusion. For instance, Montgomery, Walden, and Prosek (1987) showed that highly visible consonants, like bilabials and labio-dentals, could reduce the intelligibility of lip-read vowels. Similarly, vowels with highly visible features, such as the French vowel /u/, tend to alter the perception of consonants (Owens & Blazek, 1985). Along these lines, one can hypothesise that segment visibility favours the interpretation of neutral Cues as real phonemes (e.g. a /bə-li/ response to /bli/ would be more frequent than a /gə-li/ response to /gli/). The number of intrusions observed in our data, however, was insufficient to permit a systematic analysis.

Questions for Further Research

The results obtained in the present study are compatible with two contrasting speech processing models. The truth probably lies somewhere in between. In the first type of model, both sources of speech information, lip-reading and CS, are hierarchically processed, with the former providing the core phonological information and the latter intervening optionally, and presumably later, in the decision process. Specifically, CS would mainly be used to solve the ambiguities remaining after a first pass based exclusively on lip-reading. Thus, in the first conceptualisation, the contribution of CS can be seen as rather superficial.

In the second type of model, CS is viewed as one of the several, and equally weighted, inputs of an automatic speech processing device. This device would deal with lip-read and CS information in the same way that vision and hearing are dealt with in normal-hearing people (Massaro, 1987; Summerfield, 1987; Summerfield & McGrath, 1984). The Lip-reading/Cues compound would produce a unique amodal phonemic percept conceptually similar to Summerfield's "common metric" which integrates auditory and lip-reading information to generate a vocal tract filter function. This class of model implies that the speech processing system accepts a combination of "exotic" signals (Cues) and natural ones (lip-reading and residual hearing). Some results by Fowler and Dekle (1991) suggest that this type of interaction is possible, at least under certain conditions. They demonstrated that auditory and tactile lip information (i.e. information gathered from touching the lips of the speaker while hearing speech) combines to produce McGurk-like effects. According to Fowler and Dekle, this combination arises because both inputs provide information about the movements of the vocal tract. In contrast, the authors did not find any significant evidence that auditory and orthographic information combine.

From the theoretical frame provided by Fowler and Dekle, the central question for CS could become whether CS supplies information about speech articulation or not. Although answering this question would necessitate further empirical evidence, research suggests that this type of combination is not exceptional. It has been shown repeatedly that non-speech auditory signals combine with lip-read information to produce genuine phonetic percepts (e.g. Breeuwer & Plomp, 1986; Green & Miller, 1985; Rosen, Fourcin, & Moore, 1981).

Finally, data and speculations about the integration of tactile speech information are also a significant source of inspiration about the constraints that determine the integration of different sources (see Oerlemans & Blamey, 1998, for a recent discussion). Oerlemans and Blamey invoke a notion similar to the one proposed here in which the integration of tactile information might differ between experienced and inexperienced users. In their account, only the latter would process the tactile input in combination with the auditory and lip-reading information during the first steps of speech processing.

In conclusion, we hope that our results will lay the grounds for future research concerned with the cognitive mechanisms involved in Cued Speech processing. For example, it is important to define the impact of CS on the acquisition of phonological representations in the deaf. In addition, one should find out whether CS is used as an optional or mandatory process in spoken word recognition. These issues should be considered in light of our "early" versus "late" results given that the degree of combination between labial information and Cues might depend on the age of first exposure to CS.

Manuscript received 7 April 1997
Revised manuscript received 11 September 1998

REFERENCES

Alegria, J., & Leybaert, J. (1991). Mécanismes d'identification des mots chez le sourd. In R. Kolinsky, J. Morais, & J. Segui (Eds.), *La reconnaissance des mots dans les différentes modalités sensorielles* (pp. 277–304). Paris: PUF.

Alegria, J., Leybaert, J., Charlier, B., & Hage, C. (1992). On the origin of phonological representations in the deaf: Hearing lips and hands. In J. Alegria, D. Holender, J. Junça de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 107–132). North-Holland: Elsevier Science Publishers.

Binnie, C.A., Montgomery, A.A., & Jackson, P.L. (1974). Auditory and visual contributions to the perception of consonants. *Journal of Speech and Hearing Research*, *17*, 619–630.

Breeuwer, M., & Plomp, R. (1986). Lip-reading supplemented with auditorily presented speech parameters. *Journal of the Acoustical Society of America*, *79*, 481–499.

Burnham, D. (1998). Language specificity in the development of auditory visual speech perception. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech* (pp. 27–60). Hove, UK: Psychology Press.

Campbell, R., Dodd, B., & Burnham, D. (1998). *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech*. Hove, UK: Psychology Press.

Campbell, R., & Wright, H. (1988). Deafness, spelling and rhyme: How spelling supports written word and picture rhyming skills in deaf subjects. *Quarterly Journal of Experimental Psychology*, *40A*, 771–788.

Charlier, B.L., Hage, C., Alegria, J., & Périer, O. (1990). Evaluation d'une pratique prolongée du PLC sur la compréhension de la parole par l'enfant atteint de déficience auditive. *Glossa*, *22*, 28–39.

Conrad, R. (1979). *The deaf school child*. London: Harper & Row.

Cornett, O. (1967). Cued speech. *American Annals of the Deaf*, *112*, 3–13.

Dodd, B. (1976). The phonological system of deaf children. *Journal of Speech and Hearing Disorders*, *41*, 185–198.

Dodd, B. (1987). Lip-reading, phonological coding and deafness. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 177–189). Hove, UK: Lawrence Erlbaum Associates Ltd.

Dodd, B., & Hermelin, B. (1977). Phonological coding by the prelinguistically deaf. *Perception and Psychophysics*, *21*, 413–417.

Erber, N.P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research*, *12*, 423–424.

Erber, N.P. (1974). Visual perception of speech by deaf children: Recent developments and continuing needs. *Journal of Speech and Hearing Research*, *39*, 178–185.

Fowler, C.A., & Dekle, D.J. (1991). Listening with eye and hand: Cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 816–828.

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 eye II: Advances in the psychology of speechreading and auditory-visual speech* (pp. 3–26). Hove, UK: Psychology Press.

Green, K.P., & Miller, J.L. (1985). On the role of visual rate information in phonetic perception. *Perception and Psychophysics*, *38*, 269–276.

Hanson, V.L., Shankweiler, D., & Fischer, F.W. (1983). Determinants of spelling ability in deaf and hearing adults: Access to linguistic structure. *Cognition*, *14*, 323–344.

Kuhl, P.K., & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science*, *218*, 1138–1141.

Leybaert, J. (1993). Reading in the deaf: The roles of phonological codes. In M. Marschark & D. Clark (Eds.), *Psychological perspectives on deafness* (pp. 269–309). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

Leybaert, J., & Alegria, J. (1993). Is word processing involuntary in deaf children? *British Journal of Developmental Psychology*, *11*, 1–29.

Leybaert, J., & Alegria, J. (1995). Spelling development in deaf and hearing children: Evidence for the use of morpho-phonological regularities in French. *Reading and Writing: An Interdisciplinary Journal*, *7*, 89–109.

Leybaert, J., Alegria, J., Hage, C., & Charlier, B. (1998). The effect of exposure to phonetically augmented lipspeech in the prelingual deaf. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech* (pp. 283–301). Hove, UK: Psychology Press.

Liberman, A.M., & Mattingly, I.G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

MacKain, K.S., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left hemisphere function. *Science*, *219*, 1347–1349.

Massaro, D.W. (1987). Speech perception by ear and eye. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 53–83). Hove, UK: Lawrence Erlbaum Associates Ltd.

Massaro, D.W. (1989). Multiple book reviews of “Speech perception by ear and eye”. A paradigm for psychological inquiry. *Behavioral and Brain Sciences*, *12*, 741–794.

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

Montgomery, A.A., Walden, B.E., & Prosek, R.A. (1987). Effects of consonantal context on vowel lipreading. *Journal of Speech and Hearing Research*, *30*, 50–59.

Nicholls, G.H., & Ling, D. (1982). Cued speech and the reception of spoken language. *Journal of Speech and Hearing Research*, *25*, 262–269.

Oerlemans, M., & Blamey, P. (1998). Touch and auditory-visual speech perception. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech* (pp. 267–282). Hove, UK: Psychology Press.

Oller, D.K., & Eilers, R.E. (1988). The role of audition in infant babbling. *Child Development*, *59*, 441–449.

Owens, E., & Blazek, B. (1985). Visemes observed by hearing-impaired and normal-hearing adult viewers. *Journal of Speech and Hearing Research*, *28*, 381–393.

Périer, O. (1987). L'enfant à audition déficiente. Aspects médicaux, éducatifs, sociologiques et psychologiques. *Acta Oto-Rhino-Laryngologica Belgica*, *41*, 129–420.

Périer, O., Charlier, B., Hage, C., & Alegria, J. (1987). Evaluation of the effects of prolonged Cued Speech practice upon the perception of spoken language. In I.G. Taylor (Ed.), *The Education of the deaf: Current perspectives. Vol. 1: International Congress on Education for the Deaf* (pp. 616–625). Beckenham, UK: Croom Helm.

Rosen, S.M., Fourcin, A.J., & Moore, B.C.J. (1981). Voice pitch as an aid to lipreading. *Nature*, *291*, 150–152.

Sumby, W., & Pollack, I. (1954). Visual contributions to speech visibility in noise. *Journal of the Acoustic 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. 3–52). Hove, UK: Lawrence Erlbaum Associates Ltd.

Summerfield, Q. (1991). Visual perception of phonetic gestures. In I.G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 117–137). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Summerfield, Q., & McGrath, M. (1984). Detection and resolution of audio-visual incompatibility in the perception of vowels. *Quarterly Journal of Experimental Psychology*, *36A* 51–74.

APPENDIX

List of Items

Words		*Pseudowords*
CV-CV		
/by-ʀo/	desk	/by-vɸ/
/ʃə-vɸ/	hair	/ʃə-ʀo/
/fi-ni/	finished	/fi-te/
/ko-te/	side	/ko-ni/
/ku-ʃe/	lying	/ku-te/
/de-by/	starting	/de-mɛ̃/
/ʃə-mɛ̃/	road	/ʃə-by/
/bo-te/	beauty	/bo-ʃe/
VC-CV		
/aʀ-ʒɑ̃/	money	/aʀ-sɑ̃/
/as-pɛ/	aspect	/as-mi/

Words		*Pseudowords*
/ab-sã/	absent	/ab-ʒɛ/
/ɔb-ty/	obtuse	/ɔb-me/
/aʀ-me/	army	/aʀ-pɛ/
/al-mã/	German	/al-ʒã/
/ɛn-mi/	enemy	/ɛn-ty/
/ɔb-ʒɛ/	object	/ɔb-mã/
V-CVC		
/o-təʀ/	author	/o-muʀ/
/a-fɛʀ/	business	/a-pol/
/e-kɔl/	school	/e-maʒ/
/e-tyd/	study	/e-təʀ/
/a-muʀ/	love	/a-tyd/
/i-maʒ/	picture	/i-tɛl/
/e-pol/	shoulder	/e-fɛʀ/
/o-tɛl/	hotel	/o-kɔl/
V-CCV		
/ã-glɛ/	English	/ã-bli/
/a-pʀi/	learn	/a-gʀɛ/
/e-kʀu/	nut	/e-tʀe/
/ã-tʀe/	entry	/ã-pʀi/
/ã-gʀɛ/	fertiliser	/ã-kʀu/
/u-bli/	forget	/u-gʀɛ/
/ã-klo/	closure	/ã-bʀi/
/a-bʀi/	cover	/a-klo/