Visual Speech in the Head: The Effect of Cued-Speech on Rhyming, Remembering, and Spelling

Jacqueline Leybaert
Brigitte Charlier
Université Libre de Bruxelles

Deaf children rely mainly on lipreading to understand spoken language. The phonological representations they develop from the lipread signal are underspecified, leading to poor performances in all mental activities relying on such representations. To overcome these difficulties, systems have been designed that deliver entirely visually specified information about the phonological contrasts of spoken language. The paper explores the consequences of exposure to one of such systems, namely cued-speech (CS) on the development of phonological representations. Deaf children exposed early to CS at home show a reliance on inner speech for rhyming, remembering, and spelling similar to that displayed by hearing children but different from that of deaf children not exposed early to CS. We argue that the degree of specificity of phonological information delivered to the deaf children is more important than the modality though which they perceive speech for the development of phonological abilities.

Phonological representations play an important role in the mental life of hearing persons. In particular, a set of cognitive skills called the three R’s (rhyming, remembering, and reading) involve inner speech (Campbell, 1991). Hearing subjects typically get access to speech representations when judging whether two words rhyme, when maintaining verbal information in memory during a short time, and when recognizing printed words. To complete the picture, reliance on phonological representations is also evidenced when hearing individuals spell unfamiliar words. The term “inner speech” will be used in this article to refer to the phonological codes involved in these different activities, without assuming that it refers to one single phenomenon (see Besner, 1987). Experience with spoken language thus leads hearing individuals to develop phonological representations recruited at various levels of the cognitive architecture.

With regard to deaf persons, an important question is to what extent the development of inner speech depends in a necessary sense upon acoustic experience. Or, to formulate the problem differently: Can profoundly, prelingually deaf children acquire inner speech? To our knowledge, Dodd was the first researcher arguing theoretically against the view that acoustic information is a necessary condition for the development of inner speech. She pointed out that lipreading may constitute a primary input for deaf children to gain information about the phonological structure of spoken language (Dodd & Hermelin, 1977; see also Dodd, 1987).

Dodd’s view was elaborated in the context of a change of theoretical perspective about lipreading. Up to the mid-70s, lipreading was considered as improving speech perception in the case of a degraded acoustic signal, either because of an unfavorable signal-to-noise ratio or because of impairment of the sensory
acoustic channel (Binnie, Montgomery, & Jackson, 1974; Erber, 1969; Sumby & Pollack, 1954). Empirical studies performed since the mid-70s have shown that visual speech information is processed unavoidably by hearing listeners, even in the case of a nondegraded acoustic signal. Seeing a face saying “ga” while hearing the syllable “ba” leads to a compound percept, like “da” (McGurk & MacDonald, 1976). The influence of visual speech on the perception of acoustic speech information has also been demonstrated in serial recall experiments (Campbell & Dodd, 1980; Spoehr & Corin, 1978). These pioneer studies were later largely confirmed (e.g., the collection of articles edited by Dodd & Campbell, 1987). The main message of this bulk of studies is that the interaction between visual and acoustic information occurs at a low, precategorical level of the processing of speech input. The audiovisual interactions further suggest that auditory and visual speech information share some common stage of phonological processing (Summerfield, 1987). Developmental data are also worth considering in this discussion. The integration of auditory and visual speech information seems already established during the first weeks of life. Five- to six-month-old infants prefer to look to faces pronouncing the same syllable as that delivered acoustically than to faces pronouncing a different syllable (Kuhl & Meltzoff, 1982; MacKain, Studdert-Kennedy, Spieker, & Stern, 1983; Walton & Bower, 1993).

The research on audio-visual speech thus gives weight to Dodd's assumption that lipread information may be dealt with similarly to acoustic information by the brain. This conclusion leads to the supposition that visual information about speech contributes to the formation of phonological representations in hearing subjects. Mills's (1987) work about the phonological development of blind children illustrates this point nicely. For seeing babies, but not for blind ones, bilabials are easy to imitate, because the articulatory gesture is clearly visible. Mills found that blind children do not show the predominance of labials in their first oral productions, which characterize the early productions of seeing babies.

The consequence of bimodal (i.e., audio-visual) speech perception is that deafness per se does not preclude the development of phonological representations. While hearing subjects derive their representations from audio-visual input, deaf children depend more on the visual channel to sustain the development of their phonological system. The major obstacle for the development of phonological representations in deaf children does not lie in the modality through which they perceive speech, but in the nature of the perceived information. Lipreading, although providing information about some phonological contrasts, does not allow the perception of all of them. Lip movements give some clues about place of articulation, but no clues at all about features like nasality and voicing (Erber, 1979; Walden, Prosek, Montgomery, Scherr, & Jones, 1977). For example, visual information allows the discrimination of consonants articulated in the front from those articulated in the back of the mouth, but does not permit differentiation of consonants produced at the same articulation point (e.g., /b/, /p/, and /m/ or the vowels /a/ and /5/). Through lipreading, then, deaf children only have access to phonetically underspecified information. The representations they derive from this input are also necessarily underspecified.

The lack of well-formed phonological representations adversely affects all higher-order functions based on such representations. Deaf children generally have poor metap phonological abilities (Hanson & McGarr, 1987; Campbell & Wright, 1988), reduced memory span for linguistic stimuli (Conrad, 1979; Hanson, 1982), low reading and spelling levels (Conrad, 1979) and deficits in the retention and interpretation of free morphemes and articles (Taeschner, Devescovi, & Volterra, 1988; Volterra & Bates, 1989). Since these deficits are observed even in native sign language users, they cannot be ascribed to the lack of linguistic experience per se, but rather to the lack of complete, well-specified perception of the phonetic contrasts of spoken language.

**Cued-Speech: A System Providing Visually Well-Specified Speech Information**

To overcome these deficiencies, several systems have been created aimed at disambiguating lipreading by adding visual information carried by the hands. The most widespread among them is cued-speech (CS), devised by Cornett (1967). In CS, the speaker holds one
hand near the mouth while speaking to complement lipreading with a manual cue. A cue is made of two parameters: hand shape and place of execution around the mouth. The hand can adopt several shapes (eight in the French version of CS) at different positions around the mouth (five in French). Handshapes disambiguate the consonants and hand positions the vowels. Shapes and places are assigned to groups of two or three consonants (or vowels) in such a way that phonemes easy to discriminate by lipreading share the same handshape or place while items that are difficult to discriminate belong to different groups. For example, a particular handshape is shared by /p, d, ʒ/, another one by /b, n/ and so on. For the vowels, one place is shared by /i, ʁ, ø/, a second one by /a, o, ɔ/, and so on (see Figure 1). Each time the speaker produces a cue (a particular handshape at a specific place) while pronouncing a CV syllable, he is visually giving unambiguous information about this syllable and its phonemes. Syllabic structures like VC, CCV, CVC require additional cues to reveal the supplementary phonemes. CS has been designed in an oralist perspective. Hands alone do not give any useful information: No handshape nor hand place can be interpreted without taking the lips into account. Although the system may appear somewhat complex and artificial to uninitiated persons, it is easy to learn and, with some practice, normally hearing adults succeed at using it while speaking at a natural rate.

Initially, CS was aimed at improving the speech reception skills of deaf children. This improvement has been evaluated both for English- and French-speaking children. Nicholls and Ling (1982) studied a group of profoundly deaf children taught at school with CS since at least three years of age. They found that the speech reception scores of these children increased from about 30% for both syllables and words in the lipreading condition, to more than 80% in the lipreading + cues condition. The authors underlined that the children's average scores in the lipreading + cues condition are within the range of normal hearing listeners' reception of similar materials through audition. Pérétier, Charlier, Hage, and Alegria (1988; see also Alegria, Leybaert, Charlier, & Hage, 1992) investigated the effect of exposure to the French version of CS, called Langage Parlé Complété. (For the sake of clarity, here-after CS will be used instead of Langage Parlé Complété to describe the experiments run in French.) The subjects (mean age: 10;7 years; range: 5;11–16; 1 years) were divided in two groups: deaf children exposed to CS both at home since an early age and at school (the home group) and deaf children exposed to CS only at school (the school group). All children showed better understanding for sentences presented in the lipreading + cues condition than in the lipreading alone condition. However, the advantage provided by CS was larger for the home group than for the school group. These two studies thus strongly suggest that the addition of manual cues has the expected effect of disambiguating lipread information. This positive effect is

### Figure 1  French version of cued-speech. Reproduced with authorization of the Belgian LPC-Association.

<table>
<thead>
<tr>
<th>The consonants</th>
<th>The vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>[p] p (pas)</td>
<td>a (papa)</td>
</tr>
<tr>
<td>[d] d (dis)</td>
<td>au (peau)</td>
</tr>
<tr>
<td>[ʒ] ʒ (gel)</td>
<td>e (peux)</td>
</tr>
<tr>
<td>[k] k (cou)</td>
<td>m (pain)</td>
</tr>
<tr>
<td>[v] v (vu)</td>
<td>e (deux)</td>
</tr>
<tr>
<td>[t] t (tad)</td>
<td>[e]</td>
</tr>
<tr>
<td>[s] s (sur)</td>
<td>[ö]</td>
</tr>
<tr>
<td>[r] r (rod)</td>
<td>[i]</td>
</tr>
<tr>
<td>[b] b (bon)</td>
<td>[c]</td>
</tr>
<tr>
<td>[n] n (non)</td>
<td>ou (loup)</td>
</tr>
<tr>
<td>[t] t (tad)</td>
<td>[u]</td>
</tr>
<tr>
<td>[q] w (cuisine)</td>
<td>[o]</td>
</tr>
<tr>
<td>[m] m (maman)</td>
<td>[z]</td>
</tr>
<tr>
<td>[ɛ] n (son)</td>
<td>on (mon)</td>
</tr>
<tr>
<td>[l] l (loup)</td>
<td>[j]</td>
</tr>
<tr>
<td>[f] f (feu)</td>
<td>[j]</td>
</tr>
<tr>
<td>[j] j (joue)</td>
<td>[j]</td>
</tr>
<tr>
<td>[g] g (gu)</td>
<td>[j]</td>
</tr>
<tr>
<td>[ʃ] s (surf)</td>
<td>[j]</td>
</tr>
<tr>
<td>[ʒ] z (coup)</td>
<td>[j]</td>
</tr>
<tr>
<td>[ʒ] ʒ (gel)</td>
<td>[j]</td>
</tr>
</tbody>
</table>

* and any vowel not preceded by a consonant (arrêté)  
** and any isolated consonant (sec, prof) or followed by a silent e (lune)
particularly significant for children whose parents use CS.

Given the fact that CS allows deaf children to perceive visually the same amount of phonetic information as that perceived through audition by hearing children, a strong effect of exposure to CS on the acquisition of phonological representations might be expected. The general question we were interested in was whether deaf children educated with CS rely on phonological representations in a different way than do deaf children who are not exposed to CS. We first tested whether deaf children derive accurate phonological representations from CS, suitable for rhyme decision tasks. Next, we investigated whether these children rely on inner speech when involved in an immediate short-term memory task. Finally, we tested whether they use accurate phoneme-grapheme correspondences for word spelling. The results obtained in these three situations are summarized below.

Rhyming

One way to explore the nature and the quality of phonological representations consists of evaluating the capacity to judge whether two words rhyme or to generate rhymes in response to a word target. These tasks require access to lexical phonological representations and comparison of the words' final vowel (or vowel + coda).

If a hearing child is asked whether two words rhyme, what will determine his or her answer? First of all, the child's age. Under three years, children will probably respond at chance level; older, say around five, they will probably respond correctly. The emergence of the notion of rhyme is correlated with the quality of the child's oral productions (Webster & Plante, 1995). The degree of phonological similarity between the words is a second factor determining performance. Kindergartners are prone to judge assonant words like MASSA and LACA (in Portuguese) as rhyming as they do for SOLA and MOLA; first-graders, however, correctly categorize MASSA and LACA as nonrhyming and SOLA and MOLA as rhyming (Cardoso-Martins, 1994). Children of different reading-levels seem thus to rely on different cues to take their decision. The kindergartners probably use the global similarity between words. The first-graders likely analyze the word constituents and base their decisions upon the identity of phonological segments rather than on global similarity.

Are the profoundly and congenitally deaf children able to detect that two words rhyme as do hearing children? In a pioneer study, Dodd and Hermelin (1977) found that 12- to 14-year-old deaf children are able to judge above chance level whether pairs of lipread words rhyme. They argued that deaf children's notion of rhyme is based on visual lipread similarity between spoken words rather than on articulatory cues. Although of primary interest, this experiment did not determine whether deaf children distinguish between the notion of rhyme and the notion of lipread similarity. Indeed, all rhyming words automatically end with the same lipread image (e.g., in French LIT /li/-NID /ni/), but all pairs of words ending with identical lipread images do not necessarily rhyme (e.g., in French LIT /li/-NEZ /ne/). An important question then is to determine whether deaf people do discriminate between lipread similarity and rhyme. Unfortunately, this question has not been investigated empirically up to now. In all studies that have followed Dodd and Hermelin's, rhyming is confounded with lipread similarity.

Studies conducted after Dodd and Hermelin's (1977) have investigated the effect of another variable on rhyming judgment in deaf subjects, spelling similarity (Campbell & Wright, 1988; Hanson & Fowler, 1987; Hanson & MacGarr, 1988). Deaf youngsters, educated orally or with sign language, generally achieve reasonably good performance when the two items rhyme and are spelled similarly (e.g., BAT/HAT) but fall nearly to chance level when the items are spelled differently (e.g., HAIR/BEAR). Surprisingly enough, the deaf students were sensitive to word spelling not only when written words were used as targets, but also when pictures were used. This differentiates them from their hearing controls, who showed a slight effect of spelling congruency, but only for written targets. Campbell and Wright suggested that deaf people spontaneously rely more on spelling than hearing people, because the phonological representations derived from lipreading are not clearly enough specified to support rhyming judgment. However, reliance on spelling information varies among deaf individuals. In the three studies mentioned...
above, the deaf youngsters who had a better speech intelligibility and those who were better readers reached higher scores in rhyming judgments about word pairs spelled differently.

If the underspecified nature of lipreading is responsible for the poor quality of the phonological representations of deaf individuals, a system disambiguating lipreading, like CS, should have a positive effect both on the quality of speech representations and on rhyming skills. Children exposed to CS would be able to judge that the French words LIT and NID are rhyming, even if the rhyme is not spelled identically, and that LIT and NEZ do not rhyme although sharing the same lipread image.

This hypothesis was investigated in a series of experiments (Charlier & Leybaert, Hearing by eye: The rhyming skills of deaf children educated with cued-speech, submitted). The first one involved rhyming judgment about pairs of written words. In addition to the hearing subjects, three groups of deaf subjects, matched for grade level with the hearing children, were tested. The cued-speech+ group (CS+) included 15 children (mean age: 11;2 years) having the French version of CS at home, meaning that at least one of their parents used it in daily communication. CS had been introduced to this group at a mean age of three years. Most of these children were also provided with CS at school. The CS-group (CS-) included 20 children (mean age: 13;0 years) exposed to CS only in their school environment, since the mean age of six years. The oral group included 28 children (mean age: 13;5 years) educated exclusively with the oral/aural method both at home and at school. The results showed no difference between the groups of subjects: All children were able to make rhyme judgment with the same accuracy and all were influenced by spelling similarity. We concluded that this situation did not test access to lexical phonological representations for two reasons. First, the presentation of the written words may force all subjects to process the word spelling. Second, subjects may derive the word pronunciation from the word spelling itself instead of recovering the phonological information from the mental lexicon.

In a second experiment pairs of pictures were used. The reason for using pictures was that they do not provide any direct information about word pronunciation. This forces subjects to make comparisons between lexical phonological representations. New groups of hearing (n = 12; mean age: 8;7 years), CS+ (n = 16; mean age: 10;1 years), CS− (n = 18; mean age: 12;7 years) were recruited. In addition, new groups were included: children educated in spoken language only (the oral group; n = 29; mean age: 13;3 years) and children who used sign language, who were divided up into two groups. The children of the SL+ group (n = 12; mean age: 10;4 years) were all native signers: they had deaf parents and sign was used at home. At school they received an oral education mixed with sign language. Subjects of the SL- group (n = 20; mean age: 10;1 years) attended the same schools as the SL+ group, but had hearing parents and were educated in spoken language at home. In order to tap the processes used by the deaf children, two experimental variables were manipulated. The first one was spelling similarity. In the rhyming pairs, the pictures represented either words with similar spelling (RO+: e.g., TRAIN /trɛ /- MAIN /mɛ /) or words with different spelling (RO: e.g., LIT /li/ - RIZ ri/). The second variable was lipread similarity. Up to now, studies showing rhyme sensitivity in deaf subjects have confounded phonological and lipread similarity. To distinguish between these two factors, performance for nonrhyming pairs in which the names of the pictures have similar lipread images when spoken (NRLR: e.g., LIT /li/ - NEZ /ne/) was compared with performance for nonrhyming pairs with different lipread images (NRC: e.g., LUNE /lyn/ - FLEUR /flœr/). Subjects who believe that rhyme and similar lipread images are the same thing will have a tendency to judge erroneously that NRLR pairs rhyme. Subjects who have developed well-specified phonological representations would not be affected by lipread similarity.

Children were presented with pairs of drawings. For each pair, they had to say whether the pictures "were friends" or not. In a pretest, the notion of "friend" (i.e., rhyme) was assessed for each child, and only children who show a comprehension of this notion were included as subjects.

The results (see Table 1) revealed interesting differences between the groups. A significant effect of experimental conditions was obtained in all deaf groups, but not for the hearing group. Post-hoc con-
First, it is possible that subjects get access to under-

Table 1  Mean accuracy of deaf CS+, deaf CS−, deaf oral, deaf SL+, deaf SL−, and preliterate deaf CS+ and hearing subjects for rhyme judgment about pictures

<table>
<thead>
<tr>
<th></th>
<th>RO</th>
<th>RNO</th>
<th>NRLR</th>
<th>NRC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Readers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaf CS+</td>
<td>97.4</td>
<td>94.4</td>
<td>93.8</td>
<td>100</td>
</tr>
<tr>
<td>Deaf CS−</td>
<td>86.4</td>
<td>73.9</td>
<td>68.3</td>
<td>95.6</td>
</tr>
<tr>
<td>Deaf oral</td>
<td>92.2</td>
<td>71.6</td>
<td>73.8</td>
<td>96.6</td>
</tr>
<tr>
<td>Deaf SL+</td>
<td>89.3</td>
<td>58.2</td>
<td>74.7</td>
<td>96.7</td>
</tr>
<tr>
<td>Deaf SL−</td>
<td>82.9</td>
<td>66.6</td>
<td>58.6</td>
<td>92.0</td>
</tr>
<tr>
<td>Hearing</td>
<td>95.8</td>
<td>97.0</td>
<td>97.7</td>
<td>99.2</td>
</tr>
<tr>
<td><strong>Prereaders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS+</td>
<td>100</td>
<td>94.4</td>
<td>89.0</td>
<td>100</td>
</tr>
<tr>
<td>Hearing</td>
<td>91.3</td>
<td>91.0</td>
<td>78.6</td>
<td>94.0</td>
</tr>
</tbody>
</table>

RO: Rhyming pairs orthographically similar; RNO: rhyming pairs orthographically different; NRLR: nonrhyming pairs with similar lip-read images; NRC: nonrhyming control pairs. Deaf CS+: Children educated with cued-speech both at home and at school; Deaf CS−: Children exposed to cued-speech at school only; Deaf oral: Children educated with spoken language only; Deaf SL+: Children educated with sign language at home and with cued-speech at school; Deaf SL−: Children educated with spoken language at home and with sign language at school (adapted from Charlier & Leybaert, submitted).

trasts revealed that all groups of deaf subjects except the CS+ one had a lower performance for RO-pairs than for RO+ pairs. The difference between RO+ and RO-conditions reached 12.5% for the CS- group, 20.6% for the oral group, 31.2% for the SL+ group, 16.3% for the SL- group, and only 3.0% for the CS+ group. This orthographic effect confirms that deaf children educated orally or with sign language use their knowledge of spelling to support rhyme judgment, probably because their phonological representations are underspecified. This seems also to be the case for the CS- children. By contrast, CS+ children and hearing children do not show any sensitivity to word spelling. This differentiates them from the other groups of deaf children tested in the experiment as well as from the deaf subjects evaluated in the literature.

The NRLR condition induced more errors than the NRC condition in all groups of deaf children, the CS+ one included. However, the effect of lipread similarity was much lower in the CS+ group (in which the difference between NRLR and NRC equals 6%) than in the other groups (CS−: 27%; oral: 28%; SL+: 22%; SL−: 33%). We thought that the NRLR condition may mislead subjects at two distinct stages of processing. First, it is possible that subjects get access to under-

specified representations. This is very likely the case for children educated orally or with sign language, but not for the CS+ children. A second possible locus of the lipread similarity effect is the articulatory loop. After getting access to the word representations, subjects hold them in working memory in order to compare them. They generally pronounced the items, aloud or silently, in order to compare them. Stimuli with similar lipread images also share some articulatory features. This articulatory similarity may induce a tendency to say that the two items rhyme. This tendency may be stronger in deaf than in hearing subjects, because the former do not have acoustic images that can confirm or deny the articulatory clue. Hearing subjects may be supposed to access, through articulation, an acoustic image of the word, in the same way as they can evoke the sound of a wave, of a music instrument, or of a bird singing. Profoundly deaf people probably do not have the same capacity to evoke, mentally, the sound of a word (or of the wave, etc.), although they may know a lot of things about the sounds related to different events (Marschark, 1995). The importance of acoustic images is supported by the study of deafened adults who show depressed performance for rhyming judgment about word pairs spelled differently (Lyxell, Arlinger, Andersson, Bredberg, Harder, & Ronnberg, 1996). It is thus possible that the slight effect of lipread similarity observed in the CS+ group is located at the working memory stage of comparison, while the larger effect displayed by the other groups is located both at the representational and at the comparison level.

A further advantage of using pictures as test material is that it allows the evaluation of very young children. An attempt to do so was made with deaf prereaders, which succeeded only in the case of CS+ children. The fact that only CS+ children were able to understand the notion of rhyme is already interesting. Indeed, although the development of metalinguistic abilities is supposed to follow the acquisition of reading and spelling in children educated orally or with sign language (Campbell, 1991), it is clearly not the case for CS+ children, in which sensitivity to rhyme develops before literacy.

A group of CS+ prereaders was thus compared to a control group of hearing children. Both groups showed the same pattern of results: no effect of spell-
ing, but a significant effect of lipread similarity. This effect is compatible with the two interpretations sketched above. Possibly the representations of prereaders are, to a certain extent, different from those of readers (e.g., Studdert-Kennedy, 1986). Or perhaps prereaders are mislead by articulatory similarity, because they detect rhyme on the basis of a global phonological similarity rather than on the basis of detailed segmental information (Treiman & Breaux, 1982; Cardoso-Martins, 1994).

The effect of lipread/articulatory similarity and of orthographic similarity cannot be interpreted as showing that subjects rely only on one of these clues. If this had been the case, subjects would have reached a 0% level of correct responses in the critical condition. This was never observed. These effects show only a tendency to prefer lipread/articulatory and/or spelling information for rhyme judgment in particular tasks.

To conclude up to this point: Exposure to visual speech information that specifies all phonological contrasts leads deaf children to internalize representations that are entirely specified. The notion of rhyme seems to emerge as a natural consequence of the acquisition of such representations, independently of the modality (visual versus acoustic) through which language has been perceived. Although the rhyming ability of deaf children educated with CS seems to be well established, the detailed processes they use to perform such tasks remain to be clarified. Hearing children have been found to use the articulatory loop (see next section) when performing rhyme judgment; their rhyming scores decrease significantly under articulatory suppression (Arthur, Hitch, & Halliday, 1994). The CS+ deaf children are also likely to use an articulatory loop. There is, however, one main difference between the two populations, namely the absence of acoustic images in deaf subjects. Most of them say, during the rhyming judgment experiments: “It's difficult for me, because I'm deaf.” What kind of knowledge do they use, then, to decide that two items rhyme? Is there a “visual image,” made up of lipread + cues information, that is equivalent to the acoustic image? Or do they access a more abstract knowledge? Further experimental work, involving articulatory suppression, is needed to address these questions.

Remembering

Let us suppose somebody showed you a phone number written on a paper; you cannot write it down. What will you do in order to retain it? As a hearing person, you will very probably repeat it to yourself, again and again. You may repeat it aloud, or mentally, without producing any sound. This particular kind of memory will be discussed in this section, that is, the ability to retain during a brief interval a series of verbal items (consonants, words, pictures) in order to repeat them in the same order in which they were presented.

Conrad (1962, 1964) demonstrated that when hearing people memorize series of consonants presented visually, they show confusions that have an acoustic rather than a visual nature. For example, hearing subjects will have a tendency to confuse B, D, G, but not B and F. These confusions indicate that verbal information is held in short-term memory in a speech-related form, even if it is presented visually and if the recall is to be written.

Baddeley and his collaborators (Baddeley & Hitch, 1974; Baddeley, 1986) have elaborated a model of “verbal working memory.” There is no room here to go into all the details of this model but only to point to some of its characteristics. Working memory involves an auditory-verbal component, the “articulatory loop.” This system consists of two elements: a phonological input store within which memory traces fade if not refreshed, and a rehearsal process that serves to refresh the memory traces by means of subvocal rehearsal. The items presented auditorily enter directly into the phonological store. The items presented visually (like written words or pictures) are fed into the phonological store provided that they are encoded phonologically and subvocalized through the rehearsal process.

Two main effects have been considered as the signature of the articulatory loop: phonological similarity and length effects. The phonological similarity effect refers to the fact that rhyming items are harder to memorize than nonrhyming ones. The source of this effect is located at the phonological store, where the representations of phonologically similar words are activated at the same time. The candidate representation for a particular serial recall position is confused with
other, phonologically similar, representations. Maintaining a sequence of such representations in correct order is therefore more difficult and error-prone than maintaining a sequence of dissimilar representations in correct order.

The word length effect refers to the fact that longer words are harder to retain than shorter words. Following the standard view, this effect is a direct consequence of the articulatory nature of rehearsal. Longer words take longer to articulate, and longer to rehearse, than shorter words. Consequently, their phonological store traces can be refreshed less frequently than is possible for short words (but see Brown & Hulme, 1995, for a model simulating the length effect without any rehearsal process).

Empirical studies have demonstrated that the articulatory loop is involved both in short-term memory and in rhyme judgment, in adults (Besner, Davies, & Daniels, 1981) as well as in children (Arthur, Hitch, & Halliday, 1994). There is, however, one main difference between the two tasks. Whereas in rhyme judgment the articulatory loop is necessarily requested to perform a phonologically-based decision, it is spontaneously used in a short-term memory task.

Do deaf people use their articulatory loop spontaneously when memorizing a series of written words or pictures? Conrad (1970) was the first researcher to investigate this issue. He showed that the errors of some deaf children exposed only to spoken language were phonologically based (e.g., the B/D confusion), while for others children they were visually motivated (e.g., the N/V/X confusion). The quality of the deaf subjects’ speech was correlated with the type of errors made. Children with good oral skills made mainly phonological confusions whereas children with poor oral skills made essentially visual confusions. These data were replicated in a large-scale study involving nearly all deaf school-leavers (age: 15;6 to 16;6) of England and Wales. In this study, the ability to retain series of written items extracted from a rhyming set (DO, FEW, BLUE, WHO, ZOO, TRUE, SCREW, THROUGH) was compared to the ability to retain a series extracted from a visually similar, nonrhyming set (BARE, BEAN, DOOR, FURS, HAVE, HOME, FARM, LANE). (When presented in lower case, the words of the non-rhyming set share the same visual envelope.) Some of these deaf teenagers made more errors in the rhyming than in the nonrhyming set, evidencing for the use of a phonological code; the magnitude of the rhyming effect was, however, lower in deaf than in hearing subjects. Other deaf youngsters made more errors on the nonrhyming visually similar set than on the rhyming visually dissimilar set, showing evidence for the use of a visual code. The use of a phonological code was strongly related to hearing loss; severely deaf youngsters develop internal speech more often than profoundly deaf ones. Other studies (Hanson, 1982; Hanson, Liberman, & Shankweiler, 1984; Waters & Doehring, 1990; Wandel, 1990) have confirmed that some of the deaf youngsters exposed only to spoken language or exposed to sign language rely on inner speech to memorize series of written words.

The use of phonological coding for memorizing set of pictures has been investigated by Campbell and Wright (1990). In one experiment, children had to memorize pairs of pictures that either rhyme (e.g., CAT-HAT or TIE-FLY), were semantically related (e.g., KNIFE-FORK), or were randomly associated (e.g., BELL-LEAF). The semantic relationship was used very efficiently both by hearing and deaf subjects. The rhyme cue, however, was used only by the hearing subjects. At first sight, this finding may be surprising in regard to the studies showing that at least some deaf people are sensitive to rhyme when memorizing a series of written words. A critical difference, underlined by the authors, is the nature of the material: For the deaf, written words, but not pictures, may have a favored connection to their phonological form (see also Leybaert, Alegria, & Fonck, 1983; Leybaert & Alegria, 1993, for Stroop-like evidence about the strong relationship between written and spoken words in deaf people).

In a second experiment, the authors examined the influence of word length on memory for series of pictures. Two recall modes were tested: reconstruction of picture pairs and naming. Deaf teenagers were compared to two groups of hearing controls: younger children, and chronologically age-matched children. Deaf teenagers differed from the hearing group matched for chronological age, but had the same pattern of perfor-
Table 2  Mean accuracy of deaf CS+, deaf CS—, deaf oral, and hearing children for memorizing series of pictures

<table>
<thead>
<tr>
<th></th>
<th>Monosyllabic</th>
<th>Rhyming</th>
<th>Polysyllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaf CS+</td>
<td>63.8</td>
<td>48.9</td>
<td>49.7</td>
</tr>
<tr>
<td>Deaf CS—</td>
<td>54.5</td>
<td>57.1</td>
<td>57.6</td>
</tr>
<tr>
<td>Deaf oral</td>
<td>45.9</td>
<td>45.3</td>
<td>51.5</td>
</tr>
<tr>
<td>Hearing</td>
<td>63.6</td>
<td>47.4</td>
<td>57.6</td>
</tr>
</tbody>
</table>

Monosyllabic: monosyllabic nonrhyming words; Rhyming: monosyllabic rhyming words; Polysyllabic: polysyllabic nonrhyming words. Deaf CS+, Deaf CS—: see legend of Table 1. Adapted from Charlier (1994).

formance as the younger hearing group. Whereas the older hearing subjects show a length effect, both in the naming and in the picture reconstruction modes, the deaf and the younger hearing children displayed a length effect only for naming. There was thus no clear indication that the deaf and the younger hearing children spontaneously relied on inner speech when memorizing series of pictures. Indeed, the length effect in the naming condition can be ascribed to the rehearsal, but also to the naming process itself.

It is thus worthwhile to ask whether children educated from an early age with CS spontaneously rely on the articulatory loop in a serial recall task involving pictures. If so, they would exhibit the phonological similarity effect and the length effect. To this end, Charlier (1994) compared the performance of CS+ children, CS- children, and hearing children. The experiment included three experimental conditions: monosyllabic nonrhyming words (e.g., CHAT /a/, BANC /bā/), FEU /fe/), monosyllabic rhyming words (e.g., DE /de/, FEE /fe/, NEZ /ne/), and polysyllabic nonrhyming words (e.g., HELICOPTERE /elikopter/, CROCODILE /krɔkodil/, TELEPHONE /telefɒn/). Pictures of these words were drawn on cards that served to present the stimuli to the subjects. The number of items in a series was determined individually on the basis of a pretest so that the subjects achieved around 60% of correct responses with nonrhyming monosyllabic words. The experimenter showed the determined number of cards to the subject, at the rate of one per second. Then he turned over the cards. After the last card was turned over, the children had to reconstruct the series of pictures in the order in which they were presented.

Groups of CS+ (n = 12; mean age: 8.8 years), CS- (n = 14; mean age: 10.9 years), oral (n = 11; mean age: 9.11 years), and hearing children (n = 30; mean age: 9.4 years) were tested. The performance for the monosyllabic and nonrhyming condition did not differ significantly between groups, indicating that all children were performing at an equal level of difficulty (see Table 2). The experimental variables, however, affected the groups of children very differently. First, both the CS+ and the hearing groups displayed a large rhyming effect, while the CS- and the oral groups display no sensitivity to phonological similarity. The difference of accuracy between the control and the rhyming condition reached 6.3% for the hearing and 5.0% for the CS+ group, but −2.6% for the CS- and .6% for the oral groups. Second, both the CS+ and the hearing groups exhibited a clear effect of spoken length, whereas the CS- and the oral groups did not display any sensitivity to word length. The difference between accuracy in the control and the polysyllabic conditions reached 6.1% for the hearing and 14.1% for the CS+, but −3.0% for the CS- and −5.6% for the oral group.

The results of the CS- children suggested that their recall of series of pictures was not supported by phonological representations. This confirms Campbell and Wright's (1990) results. The data of the CS+ contrast with those of these two groups. The phonological similarity effect and the length effect indicated that the CS+ children spontaneously used the articulatory loop, based on phonological representations. Access to phonological representations seems to be as unavoidable in this situation for CS+ subjects as it is for hearing subjects. Indeed, this access limits their performance in both rhyming and polysyllabic conditions. It is also interesting to note that the use of inner speech seems to enhance their memory span. Indeed, the number of items used to run the experiment was higher for the CS+ and the hearing groups (5.17 in average) than for the CS- (4.86) and 4.67 for the oral group.

Reading and Spelling

The development of well-specified phonological representations in deaf children may also have important consequences for their acquisition of reading and spelling. It is generally assumed that hearing subjects recognize printed words in two ways. Familiar words may be identified very rapidly, on the basis of the activation of an orthographic representation stored in long-term
memory. Unfamiliar words, for which no orthographic representations are available, may be identified through the application of grapheme-phoneme conversion rules, leading to the assembly of the word's phonological form. The sound of the word then gives access to its meaning. Similar procedures are available for spelling. Hearing children may use their lexical knowledge about how a particular word is spelled, or they may construct a spelling output by applying phoneme-to-grapheme rules to the word's phonological form. Theorists have underlined that the use of sublexical correspondences between graphemes and speech segments plays an important role at the beginning of reading and spelling acquisition (Frith, 1985; Seymour & McGregor, 1984).

One may wonder to what extent deaf children use their phonological representations for reading and spelling. In other words, are the deaf limited to the exclusive use of lexical word-specific knowledge, available only for familiar words, or are they also able to use sublexical relationships between phoneme and grapheme? In the present article, we will limit ourselves to the discussion of the procedures used for spelling.

The typical method for determining whether deaf children use phoneme-to-grapheme knowledge consists in comparing their performance for regular and irregular words. Regular words are those for which the correct spelling can be straightforwardly derived from their pronunciation, by applying phoneme-to-grapheme rules. Irregular words are those for which the correct spelling violates the phoneme-to-grapheme rules and must thus be retained by rote. The rationale of the comparison is the following: An advantage for regular over irregular words indicates that children rely on phoneme-grapheme rules for spelling. On the other hand, if children rely exclusively on word-specific knowledge, no difference is expected in the spelling of the two types of words.

In several studies, deaf youngsters have been found to show a regularity effect, that is, better performance for regular than for irregular words. This is true not only for deaf students who have achieved an exceptional reading level (Hanson, Shankweiler, & Fischer, 1983), but also for children with a reading age of approximately nine years (Burden & Campbell, 1994; Leybaert & Alegria, 1995; but see Waters & Doehring, 1990, for negative results). In our study (Leybaert & Alegria, 1995), the regularity effect was less marked in deaf than in hearing subjects, not because the deaf were better with irregular words, but because they were poorer with regular words. The deaf subjects thus took less advantage than the hearing subjects of the existence of regularities between phoneme and grapheme.

Another indicator of the use of phoneme-to-grapheme regularities is the percentage of phonologically accurate errors. When a hearing child ignores how to spell a particular word, he produces, most of the time, a spelling compatible with the word pronunciation (e.g., BRANE for BRAIN). This kind of misspelling reveals that the child starts from an accurate phonological representation to which he applies phoneme-to-grapheme rules. Deaf people also make such misspellings, but less than their hearing peers. For example, in our study, the percentage of phonologically accurate errors was 21.8% for young deaf children against 90.6% for young hearing children.

These observations may be interpreted in two ways. First, deaf children may rely less than the hearing on phoneme-to-grapheme rules. Second, they may rely on phoneme-to-grapheme rules, but start from inaccurate phonological representations. Evidence supporting both hypotheses was found in Leybaert and Alegria's study. The fact that deaf children start from inaccurate, underspecified phonological representations is exemplified by the systematic part of their phonologically inaccurate errors. For example, they often reduced consonantic clusters by deleting the liquid, misspelling ARMOIRE as AMOIRE, CARTABLE as CATABLE. They substituted voice, voiceless, and nasal consonants, misspelling OUVERT as OUFERT, AVION as AFEO. These errors indicate that the deaf apply conversion rules to inaccurate representations. For example, they often reduced consonantic clusters by deleting the liquid, misspelling ARMOIRE as AMOIRE, CARTABLE as CATABLE. They substituted voice, voiceless, and nasal consonants, misspelling OUVERT as OUFERT, AVION as AFEO. These errors indicate that the deaf apply conversion rules to inaccurate representations (see also Burden & Campbell, 1994, for similar observations). On the other hand, lower reliance on phoneme-grapheme rules is indicated by the occurrence of misspellings consisting of keeping the letters of a word, but in a wrong order (e.g., SPTE instead of SEPT). These misspellings occurred only in the deaf group. They probably result from a tendency to retain the word spelling by rote, without reference to its pronunciation.

To summarize up to this point, although deaf youngsters do use phoneme-to-grapheme mappings,
Table 3 Characteristics of the subjects of the spelling experiment

<table>
<thead>
<tr>
<th></th>
<th>Hearing</th>
<th>Deaf CS+</th>
<th>Deaf CS-</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>28</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Mean age</td>
<td>9;2</td>
<td>9;8</td>
<td>12;2</td>
</tr>
<tr>
<td>% Correct responses spelling test</td>
<td>85.1</td>
<td>86.0</td>
<td>75.0</td>
</tr>
<tr>
<td>Score at the reading test</td>
<td>23.6</td>
<td>23.3</td>
<td>10.9</td>
</tr>
<tr>
<td>Mean school years (in years; months)</td>
<td>3;1</td>
<td>3;4</td>
<td>5;5</td>
</tr>
</tbody>
</table>

they are unable to take a full advantage of the transparency of alphabetic orthographies by the underspecified nature of their phonological representations. Exposure to visual systems that entirely specify the phonological information, like CS, may thus have a strong impact at this level. Deaf children who have internalized accurate phonological representations may benefit more from the regularities between phonemes and graphemes than do other deaf children not using CS.

The effect of exposure to CS on the development of spelling has been investigated in an experiment still in progress (Leybaert, The effect of cued-speech on the development of deaf children’s spelling, in prep.). Again, three groups of children were compared: deaf children educated with CS at home (the CS+ group), deaf children exposed to CS only at school (the CS- group), and hearing children (see Table 3). These children were assessed on a spelling test, which consisted of writing the names of pictures representing simple concrete words. They also passed a reading sentence comprehension test. The preliminary results show that the CS+ group achieved the same spelling level and the same reading level as the hearing controls at approximately the same chronological age. This contrasts with the CS- children who, although older than their hearing peers, were delayed in their spelling and even more in their reading achievement. These descriptive data already indicate a positive effect of exposure to CS at home on the development of literacy.

What interests us primarily, however, is the effect of exposure to CS on the procedures used for spelling and, more precisely, on the use of phoneme-to-grapheme correspondences. The type of errors constitutes one indicator of this. In the hearing group, 92.9% of the errors were phonologically accurate (e.g., SIGARETTE for CIGARETTE; TRIN for TRAIN). As expected, this was also the case for most of the misspellings of the CS+ children (74.3%). This contrasts with the CS- children who made, besides phonologically accurate misspellings (30.5%), a large amount (32.1%) of phonologically inaccurate spellings (e.g., COPAT for COPAIN, JOU or ZUS for JUS). The error analysis thus confirms that CS+ children start from accurate phonological representations for spelling.

Given that CS+ children seem to access accurate phonological representations, it is possible that they develop spelling along lines similar to those of their hearing peers. The development of two aspects of their spelling is summarized below: spelling of words containing dominant and nondominant phoneme-grapheme transcriptions and spelling of words containing consonant clusters.

**Phoneme-to-grapheme dominance.** Although French orthography is quite consistent for reading, this is not the case for spelling; most of the phonemes can be spelled in different ways. They have a statistically dominant transcription (e.g., /s/ at the beginning of words, and before e and i is usually written as the letter S) and one or several nondominant transcriptions (i.e., the phoneme /s/ is also transcribed, but less frequently, by the letter C). The interest in using such variables experimentally is the following. The spelling of the words containing a dominant transcription constitutes a good evaluation of the degree of mastery of the frequent phoneme-grapheme correspondences. On the other hand, the percentage of correct spelling for words containing nondominant transcriptions is an indicator of the development of the orthographic lexicon (Alegría & Mousty, 1994).

The children were thus asked to spell words containing either dominant (e.g., SEL, SECRET) or nondominant transcriptions (e.g., CIEL, CIGARETTE). The words were varied in frequency. The score was the percentage of correct spelling for the grapheme under
Table 4  Percentage of correct responses in spelling words containing dominant and nondominant phoneme-grapheme transcriptions

<table>
<thead>
<tr>
<th></th>
<th>Frequent</th>
<th>Rare</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dominant</td>
<td>Nondominant</td>
</tr>
<tr>
<td>Hearing</td>
<td>87.2</td>
<td>51.5</td>
</tr>
<tr>
<td>Deaf CS+</td>
<td>74.1</td>
<td>50.9</td>
</tr>
<tr>
<td>Deaf CS-</td>
<td>96.8</td>
<td>82.4</td>
</tr>
</tbody>
</table>

Deaf CS+, Deaf CS–: see legend of Table 1. From Leybaert (in prep).

Table 5  Percentage of correct responses for spelling words containing consonant clusters and control words

<table>
<thead>
<tr>
<th></th>
<th>Frequent</th>
<th>Rare</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clusters</td>
<td>Controls</td>
</tr>
<tr>
<td>Hearing</td>
<td>94.3</td>
<td>100</td>
</tr>
<tr>
<td>Deaf CS+</td>
<td>89.0</td>
<td>100</td>
</tr>
<tr>
<td>Deaf CS–</td>
<td>80.5</td>
<td>96.7</td>
</tr>
</tbody>
</table>

Deaf CS+, Deaf CS–: see legend of Table 1. From Leybaert (in prep.).

investigation. For example, if CIEL was spelled with a C, a correct response was attributed, independently of the way the other phonemes were spelled.

The CS+ children behaved similarly to the hearing controls (see Table 4). (Because this experiment is unfinished, the results described here are only preliminary, and we limit ourselves to describing the patterns, without giving the statistics.) At the beginning of learning to spell, they spell correctly words containing the dominant rule, irrespective of word frequency. They also made many errors for words including a nondominant rule, which were usually spelled with the dominant transcription (e.g., SIEL instead of CIEL). Their pattern of performance, at this first stage of learning to spell, is clearly more influenced by the dominance of the transcription than by word frequency, which indicates a main reliance on phoneme-to-grapheme mapping rules.

The results of the CS– group contrasted with those of the other two groups (see Table 4). At the first level of spelling achievement, the CS– children had higher scores than the other two groups for frequent words, containing dominant as well as nondominant transcriptions. This probably resulted from the fact that they were older and had had a longer exposure to orthographic material. Their scores for rare words, although higher for the nondominant transcriptions, were clearly depressed for the dominant transcriptions. Their whole pattern was more determined by word frequency than by the use of phoneme-to-grapheme dominance, which indicates a main reliance on lexical knowledge.

Consonant clusters. Bruck and Treiman (1990) have shown that consonant clusters like BR, TR, GR are difficult to spell for young hearing children who frequently omit the liquid in their written productions. Our previous study of the spelling of deaf children educated orally (Leybaert & Alegria, 1995) showed that they have persistent difficulties in spelling consonant clusters, probably because the two consonants do not have distinct images on the lips. The delivery of precise phonological information by CS may play a positive role on spelling such syllabic structures, because each consonant is indicated by a manual cue. In the same study (Leybaert, in prep.), the spelling of R and L either in a consonant cluster (e.g., TRAIN, FLEUR, TROMPETTE) or at the beginning of a syllable (e.g., LAPIN, ROUGE, REVOLVER) was thus compared in the three groups of subjects. Again, target words were varied in frequency. A correct response was scored if the R or the L was present in the subject's response.

The data of the hearing beginners (see Table 5) show that words containing consonant clusters entail slightly more errors than control words, thus confirming Bruck and Treiman's data. Of course, no effect of
word frequency appears: the graphemes R or L, either in consonantic clusters or in control words, could be spelled on the basis of consistent phoneme-to-grapheme rules. The CS+ group, although achieving slightly lower performance for the clusters, show the same pattern as the hearing group. This contrasts with the data of the CS- children, who were more impaired by consonant clusters, especially for rare words, and who showed a stronger effect of word frequency.

To sum up, the CS+ children exhibit spelling development qualitatively and quantitatively similar to that of hearing children. This indicates that deaf children can establish correspondences between visual representations of speech and the alphabet, provided that these representations are entirely specified.

Conclusion

The fact that the CS+ children rely on inner speech processes to the same extent as their hearing peers for rhyming, remembering, and spelling has important implications for several theoretical and educational issues.

First, the observations made for the CS+ children are relevant to the understanding of the potentialities available to deaf people in general. Because of their sensory impairment, deaf children are often thought to develop necessarily different cognitive functioning. At first sight, this opinion may seem supported by the available data. Up to now, the studies have shown that some deaf people may develop phonological and metaphonological abilities, but not to the same degree as hearing people (see e.g., Conrad, 1979). The fact that the CS+ children, despite profound and prelingual hearing loss, behave similarly to the hearing in a number of situations indicates clearly that it is possible for deaf children to develop phonological representations and inner speech processes, provided that they are exposed to an adequate input.

A second question concerns the critical conditions for the development of such inner speech processes. The case of CS+ children indicates that acoustic experience of speech is in no way a necessary condition for the development of such processes. Information delivered exclusively through the visual modality, as in CS, may trigger off the development of phonological abilities. What seems to be the critical factor is the delivery of accurate, well-specified information about the phonological contrasts of spoken language. Indeed, the performance of the children exposed to lip reading clearly shows only that exposure to an underspecified input is not sufficient to start the development of inner speech processes.

In addition, the differences observed between CS+ children and CS- children suggest that the critical factor for acquisition of accurate phonological representations is the exposure to CS at home. Indeed, children exposed to CS only at school do not use phonological representations to the same extent and with the same accuracy as the CS+ children. The two groups may differ in a number of aspects. It is possible that the children exposed to CS at school suffer only from a lack of linguistic experience in terms of quantity of language. Another possibility is that the quality of the language addressed to the child is different at school and at home. A third explanation could be that the preciousness of exposure to CS plays a determinative role in the emergence of phonological abilities. Our data do not allow us to choose between these explanations because the children educated with CS at home were also exposed to it at an early age. Further experimental research is needed to determine whether these factors can be disentangled in terms of their effect on deaf subjects' phonological skills.

Our data also confirm the close relationship between the use of inner speech in the activities that have been called the three R's (rhyming, remembering, reading, and spelling). The phonological representations derived from CS are recruited in cognitive functioning sometimes automatically, as for memorizing, sometimes perhaps more deliberately, as for rhyming or spelling. Further work is necessary to determine if phonological information also becomes automatically available when these children recognize printed words.

References


