Impaired Speech Perception in Poor Readers: Evidence from Hearing and Speech Reading

Beatrice de Gelder* and Jean Vroomen

Tilburg University, The Netherlands and *Université Libre de Bruxelles, Belgium

The performance of 14 poor readers on an audiovisual speech perception task was compared with 14 normal subjects matched on chronological age (CA) and 14 subjects matched on reading age (RA). The task consisted of identifying synthetic speech varying in place of articulation on an acoustic 9-point continuum between /ba/ and /da/ (Massaro & Cohen, 1983). The acoustic speech events were factorially combined with the visual articulation of /ba/, /da/, or none. In addition, the visual-only articulation of /ba/ or /da/ was presented. The results showed that (1) poor readers were less categorical than CA and RA in the identification of the auditory speech events and (2) that they were worse in speech reading. This convergence between the deficits clearly suggests that the auditory speech processing difficulty of poor readers is speech specific and relates to the processing of phonological information.

There is at present a reasonable consensus that poor reading acquisition is related to phonological deficits. These phonological problems of developmental dyslexics are currently approached from a number of different vantage points and it seems too early to expect a unified picture to emerge. But it is clear that the notion of a phonological problem covers a great variety of processes involved in reading and in learning to read. To begin with, the understanding of reading disorders has greatly profited from new developments in the study of the adult reading process, and research on dyslexia is increasingly inspired by the state of the art in reading research (Bertelson & de Gelder, 1989, 1991) and the role of phonology in adult and in poor readers (Bosman, Leerdam, & de Gelder, 1997). Such insights on developmental dyslexia gained by putting developmental questions in the context of adult reading models must now be integrated with a long-standing tradition that focuses on the notion of phonological awareness, the crucial skill developed by beginning readers that is fully mastered by the time basic fluency is established. The poor readers manifest the same difficulties with phonological
awareness tasks that are by now well documented for prereaders (Bertelson, de Gelder & van Zon, 1997), illiterates (Bertelson, de Gelder, Morais & Tfouni, 1989) and nonalphabetic literates (Bertelson & de Gelder, 1991; de Gelder, Vroomen, & Bertelson, 1993). It is generally agreed that poor performance on tasks measuring segmentation skills and the associated reading difficulties stem from problems in the phonological domain. The assumption is that the ability to perform grapheme/phoneme conversions unites phonological awareness deficits and nonword reading problems. A different domain of troubles for poor readers is verbal memory. Here also evidence points to phonological aspects since memory deficits of poor readers appear to be specific for spoken material (see Brady, 1991, for an overview). Phonological confusability as well as word length have been pinned down as critical factors, but their importance for poorer memory is not at all that obvious when all relevant factors are controlled for (Irasquin & de Gelder, 1997). More critically, studies of phonological processes in memory have not made it clear how the phonological component of verbal memory might be related to phonological processes in reading or in nonword reading. The latter is clearly a skill that is poor in developmental dyslexics (Rack, Snowling, & Olson, 1992), but the importance of this observation is not fully understood. So while it is repeatedly affirmed that poor reading and deficient phonological skills are firmly linked, it remains for now unclear what phonological components present in these different skills are related or might have a common root or just what this phonological component might actually be.

While it remains central to the mainstream scientific literature on developmental reading deficits that in an alphabetic orthography these deficits are related to acquiring grapheme/phoneme conversion skills, the demands such a conversion imposes on the speech representations themselves are matters that have received surprisingly little attention. One intuition is that the development of grapheme/phoneme conversion skills might be at risk because, among other reasons, phonological representations of the developmental dyslexic do not have the robustness required or do not have the degree of fine-grainedness to play their assigned role in grapheme/phoneme conversion.

Such a possibility has only emerged in the literature sporadically. Poor readers count as normal in on-line perception of spoken language, but a small number of studies have reported inferior perception of auditory verbal stimuli by poor readers (Brady, Poggie, & Merlo, 1986; Brady, Shankweiler, & Mann, 1983; de Weerd, 1988; Godfrey, Syrdal-Laskey, Millay, & Knox, 1981; Lieberman, Meskill, Chatillon, & Schupack, 1985; Reed, 1989; Tallal, 1980; Werker & Tees, 1987). In a preliminary report of the data presented below we also found impairments in speech processing (de Gelder & Vroomen, 1988). The latter study, like the ones by Godfrey et al. (1981), Werker and Tees (1987), and De Weerd (1988), used a categorical speech perception paradigm. So did Irasquin and de Gelder (1997) in a study that found impaired perception for a place of articulation continuum.
Brady, Shankweiler, and Mann (1983) found evidence that the deficit is speech specific. They compared normal and slow readers on their ability to discriminate speech and nonspeech sounds under ideal and noisy listening conditions. No reading-group differences were found in the recall of environmental sounds, but performance was significantly different with words presented in noise. In contrast with this specific focus on speech perception, Tallal (1980) argued that reading-backwardness involved difficulties in general auditory processing. She found that reading-disabled children were worse than normal children in making temporal order judgments for very brief nonspeech tones presented at short interstimulus intervals. In a similar vein, Reed (1989) reported that reading-disabled children were impaired in making judgements with very brief tones and with stop consonants, while there were no difficulties with vowels, vowels in background noise, or briefly shown visual figures. The critical distinction in the latter experiment was that brief tones and stop consonants, but not vowels, have short-lived auditory cues. However, the subtle impairments in phonological processing cannot be reduced to differences in auditory processing skills between normal and poor readers, as has been claimed recently again by Tallal and collaborators, but are speech specific in nature rather than consequences of dysfunctions in general auditory processing. A clear argument in favor of such a sharp distinction was recently made by Studdert-Kennedy and Mody (1995). The present study pursues this line and focuses on speech-specific impairments, not by contrasting general auditory and speech-specific processes, but by taking advantage of the fact that speech is conveyed to the ear as well as to the eye. Obviously, for the latter input channel the problem of sorting out what is general sound and what is speech does not arise.

Auditory input is not the only source of speech information. As a matter of fact, for normal subjects whether in normal or in poor auditory input conditions, as well as for subjects with sensorial deficits, speech reading contributes significantly to speech understanding (Summerfield, 1987; Massaro, 1987; Campbell, 1989). Evidence for the notion of development in visual speech processing was made by data from Massaro and his colleagues (Massaro, Thompson, Barron, & Laren, 1986) showing clearly that visual speech perception develops over time. Summerfield (1991) has shown that large individual variations notwithstanding, speech reading skill is not related to general intellectual abilities but seems to be modular, together with linguistic competence in general. Understanding of the visual speech processor is progressing, just as are insights into the auditory speech processor and in the mechanisms whereby visual and auditory input combine. One possibility much favored by advocates of the “speech is special” perspective (Liberman, 1996) is that both input modalities share computational resources such as those that might be provided by common abstract speech representations of a speech module. Against this background we conjecture that when a memory deficit is observed with auditorily presented verbal material, a simi-
lar deficit also shows up in visual speech processing. Multimodality of speech input raises the issue whether populations suffering from perceptual or representational problems would gain an advantage from concurrent presentation in two modalities. For example, one might imagine that the extra visual information in audiovisual speech would boost performance if compared with auditory or visual-alone presentation. De Gelder and Vroomen (1995) pursued this notion in a study that looks at young poor readers as well as adult dyslexics. The study showed very clearly that the two groups of poor readers were impaired to the same extent in auditory speech processing and in visual speech processing. The overall performance of poor readers on the auditory as well as on the visual only memory lists lagged behind that of the two control groups for a pattern of performance that was structurally the same in both populations.

The present study extends the speech-specific view of language processing impairments of poor readers by taking advantage of the fact that speech is conveyed to the ear as well as to the eye. Poor readers and normal children were compared not only on the identification of auditory speech, but also on visual speech. It is known that hearing and speech reading are in part subserved by the same underlying speech processing mechanism (e.g., Summerfield, 1987). For instance, when incongruent auditory and visual speech are dubbed, the percept is sometimes a phoneme which is an intermediate solution between the heard and the visual information (McGurk & MacDonald, 1976). This so-called McGurk effect is rather clear evidence that heard and lipread speech are integrated somewhere in speech processing (e.g., Massaro, 1987). Given that hearing and speech reading are part of speech processing, it becomes interesting to investigate speech reading in poor readers. If it is the case that poor readers have a speech-specific impairment, one might expect to find an associated deficit in speech reading. On the other hand, if the nature of the impairment is in the general auditory domain, there is no reason to expect that speech reading is impaired. In the latter case, one might find a dissociation between hearing and speech reading in that hearing is impaired while speech reading is preserved or may be even better because speech reading might compensate for the auditory deficit.

Finally, a dissociation would not only be relevant to theories about dyslexia, but it could also give insights in the way auditory and visual speech are combined. One particular question concerning auditory and visual speech perception is what the nature of the information is at the conflux of the acoustic and optical signal. One possibility is that both signals are integrated early in perceptual processing, namely before the acoustic and optic signal are phonetically evaluated (e.g., Summerfield, 1987). The alternative is that integration occurs later, namely, after phonetic feature evaluation (e.g., Massaro, 1987; Vroomen, 1992). The early-integration notion would predict that a phonetic impairment will have similar consequences on hearing and speech reading because both modalities are subserved by the same (deficient) pro-
cessing system. In contrast, the late-integration hypothesis does not exclude that a phonetic impairment might have different consequences for hearing and speech reading because separate processing routes could be involved that can be selectively impaired. Besides exposing the speech-specificity of perception impairments in poor readers, the present study might shed some light on this issue.

METHOD

Subjects

Three groups of subjects were tested, a poor readers group, a group matched on reading age (RA), and a group matched on chronological age (CA). The poor readers group consisted of 11 boys and 3 girls. Mean age was 11.4 years, ranging from 9.4 to 14.2 years. They were referred to special services at the university because of their reading problems and they were diagnosed as poor readers. Nine of the poor readers attended a special school for children with learning problems, two attended a regular primary school, and the others attended secondary school. They all followed remedial teaching in reading. They were considered of normal intelligence on the basis of having WISC Full Scale I.Q.s ranging from 85 to 123 (M = 104, SD = 9.5). All subjects were given a standard reading test (Brus), which required reading aloud Dutch real words in 1 min. Poor readers were lagging 1.3 to 4.6 years (M = 2.7 years) behind their age mates. Subjects in the RA group were selected by individually matching their performances on the reading test with those of the poor readers. This group consisted of 10 boys and 4 girls. They were, of course, younger than the poor readers (M = 8.11 years, ranging from 7.7 to 11.5 years). They attended a regular primary school. The children in the CA group, 8 boys and 6 girls, were individually matched on chronological age with the poor readers. They attended a normal primary or secondary school and did not lag behind in their reading level. All children were native speakers of Dutch and were considered as having normal hearing.

Stimuli

A color videotape prepared by Massaro and Cohen (1983, experiment 2) showed a male speaker seated in front of a neutral background illuminated with ordinary fluorescent light. The speaker’s head filled about two-thirds of the screen. An experimental tape was made by copying the original video tape and replacing the natural soundtrack with synthetic speech. A nine-step /ba/ to /da/ auditory continuum was created as follows: tokens of the speaker’s /ba/s and /da/s were analyzed using linear prediction to derive a set of parameters for driving a software-formant serial-resonator speech synthesizer (Klatt, 1980). By altering the parametric information regarding the first 80 ms of the syllable, a set of nine 400-ms syllables covering the range of /ba/ to /da/ was created. During the first 80 ms, Formant 1 (F1) went from 250 to 700 Hz following a negatively accelerated path. The F2 followed a negatively accelerated path to 1199 Hz from one of nine values equally spaced between 1000 and 2000 Hz from most /ba/-like to most /da/-like, respectively. The F3 followed a linear transition to 2729 Hz from one of nine values equally spaced between 2200 and 3200 Hz. All other stimulus characteristics were identical for the nine auditory syllables.

Procedure

Nine levels along the auditory /ba/ to /da/ continuum were factorially combined with two possible visual articulations, /ba/ or /da/. These 18 trials represent the bimodal condition. There was also an auditory-alone and a visual-alone condition. In the auditory-alone condition
FIG. 1. Proportion of /da/ responses as a function of the auditory and visual information for the dyslexic, chronological age, and reading age groups.

one of the nine auditory stimuli was presented, but the speaker did not move his mouth. In the visual-alone condition, the speaker articulated either /ba/ or /da/, but no auditory speech was presented. In every block of 54 trials, there were 18 bimodal conditions, 18 auditory-alone conditions, and 18 visual-alone conditions. The experiment consisted of 5-1/2 blocks of trials preceded by 10 practice trials for a total of 307 trials. There was a 5-min break after every 80 trials.

The children were tested individually in a quiet room. They viewed a 63-cm television monitor which presented both the auditory and the visual dimensions of the speech stimuli. Subjects sat about 2 meters from the monitor. The audio was set at a comfortable listening level (approximately 71 dB-A in the fast mode). Subjects were told of the three different kind of trials: the bimodal trials, the auditory-alone trials, and the visual-alone trials. They were instructed to report for any of the trial types what the speaker just said, whether /ba/ or /da/. In the case of a visual-only trial they had to report whether the speaker’s mouth made /ba/ or /da/. Further reference to modality was strictly avoided. The experimenter sat next to the child to determine whether the child was watching the screen at the time of the speech presentation. If this criterion was not met, the trial was disregarded. The responses were made by oral report.

RESULTS

Robustness of Auditory Categorization

For each subject in each group and condition, the proportion of /da/ responses was computed. Figure 1 displays the mean proportion of /da/ responses as a function of the nine levels of the auditory continuum and the
three levels of the visual variable, a /ba/, none, or /da/ articulation, separately for the poor readers and the control groups. An analysis of variance (ANOVA) was performed on the proportion of /da/ responses with group as the between-subjects variable and the auditory and visual factors as within-subjects variables. The main effect of the group did not reach significance \(F(2, 39) = 2.39, p > .10\). As expected, the average proportion of /da/ responses increased significantly as the auditory stimulus moved from the /ba/ to the /da/ end of the continuum \(F(8, 312) = 477.0, p < .001\). There was also a large effect of the visual variable \(F(2, 78) = 127.0, p < .001\). The average proportion of /da/ responses increased from .410 for the /ba/ visual stimulus, to .587 for the neutral stimulus, to .683 for the /da/ visual stimulus. The interaction of the auditory and the visual variable was significant \(F(16, 624) = 11.50, p < .001\) because the effect of the visual variable was larger for the ambiguous levels of the auditory continuum. These findings are all in agreement with the studies of Massaro and Cohen (1983).

There was a significant second-order interaction of group by auditory by visual factors \(F(32, 624) = 1.57, p < .025\). All other interactions did not reach significance \((p > .10)\). In order to explore the second-order interaction, a separate ANOVA with group as the between-subjects variable and the auditory factor as the within-subjects variable was performed on the proportion of /da/ responses with a neutral visual articulation. The mean proportions are replotted in Fig. 2. There was a main stimulus effect \(F(8, 312) = 370.9, p < .001\) because subjects identified the first several stimuli as /ba/ and the last several stimuli as /da/. The overall difference between the groups was not significant \(F(2, 40) = 2.62, p > .05\). Of interest was a significant interaction between the groups and the auditory variable \(F(16, 312) = 2.49, p < .001\). Inspection of Fig. 2 suggests that the poor readers were less categorical in the identification of the stimuli. That is, the labeling function for the poor readers is less steep than the functions for their matched control groups. However, the interaction in the ANOVA may be significant for other reasons as well: for example, differences in the place of the phoneme boundary.

To explore this further, the data were analyzed by submitting the proportions of /da/ responses to a logit transformation from which the values of the phoneme boundaries and the slopes were obtained (Finney, 1964). This will allow us to separate the different alternatives. The values are determined by regressively computing the cumulative normal distribution, which is closest, by a maximum likelihood criterion, to the data. The mean of the resulting distribution is the interpolated 50% crossover point (phoneme boundary) and the slope is a measure of the degree of the sharpness with which phoneme categories are distinguished from one another. There were no significant differences in the place of the phoneme boundary (4.25, 3.82, and 4.61 stimulus units for the poor readers, CA, and RA, respectively, \(F(2, 39) = 2.04, p > .10\)). However, the slope of the labeling function was less steep for the poor
FIG. 2. Proportion of /da/ responses as a function of the auditory levels for the dyslexic, chronological age, and reading age groups.

readers than for the controls (Mean = 1.03, 1.40, and 1.43 for the poor readers, RA, and CA, respectively, $F(2, 39) = 3.84, p < .03$). It confirms that poor readers were more variable than their controls in the identification of the speech sounds.

Visual Influence on Auditory Identification

The second-order interaction in the overall ANOVA was also explored for the visual conditions. To get a measure of the influence of vision on speech perception, we computed the visual influence by subtracting the proportion of /da/ responses given a visual /ba/ from the proportion of /da/ responses given a visual /da/, thereby averaging across the auditory levels. The larger this value, the greater the reliance on speech reading in the audio-visual conditions. The visual influence was .225, .282, and .312 for poor readers, RA-, and CA-controls, respectively [$F(2, 39) = 1.51$, NS]. There is thus no significant difference of visual influence, although there is a trend in the direction that poor readers are less influenced by vision than the control groups.
Speech Reading Performance

In the visual-only condition, the proportion of correct responses was determined. The average proportion of correct responses was .67 for poor readers, .77 for reading-age controls, and .77 for chronological-age controls. This difference is significant \[ F(2, 39) = 3.24, p < .05 \], showing that poor readers were poorer speech readers as compared to their controls.

DISCUSSION

The present study examined auditory and visual speech perception in poor readers and normal readers. Its purpose was to add to the evidence in support of a speech-specific basis of subtle speech perception anomalies in poor readers. In the speech-specific notion, reading problems are associated with impairments in the phonological domain, while the auditory view seeks the origin of the deficit in the processing of brief auditory cues. Performance on speech reading should allow us to differentiate among these very different accounts since speech reading does not rely on the auditory channel. The results showed that poor readers were worse than their controls in processing auditory and visual speech. It thus supports the notion that difficulties in reading are associated with speech-specific problems and not with a general auditory deficit.

These findings are consistent with as well as extend previous results showing that reading-disabled children have problems in categorical perception of some consonants. Godfrey et al. (1981) used synthetic stop consonant continua varying in place of articulation (/ba/–/da/ and /da/–/ga/) and found that poor readers were less consistent in the identification of the stimuli as their identification slopes were less steep than those of the better readers. The same findings were obtained independently by Werker and Tees (1987) and Reed (1989). Irasquin, de Gelder, and Vroomen (1997) showed an anomaly in the categorical perception of contrastive consonants where place of articulation is critical. It thus seems to be a consistent finding that the phonological speech categories of poor readers are less robust. This claim is compatible with other results showing that poor readers have a speech deficit.

For instance, Brady, Shankweiler, and Mann (1983) found that poor readers had more difficulty than good readers in identifying words presented in noise, while there was no difference between the groups when environmental sounds were used.

A difficult issue with age-matched comparisons is whether the poor readers’ difficulties with speech processing are a cause or a consequence of their reading skills. It could be that their perceptual problems stem from their low reading performance. This suggestion is not unreasonable since it is known that alphabetic reading skill has influence on speech processes. A well-known example is that rhyme judgement is influenced by reading abilities (Seidenberg & Tanenhaus, 1979), but even more perceptual tasks such as
identification, as used here, are influenced by reading performance. For instance, de Gelder and Vroomen (1992) found that Chinese subjects with alphabetic knowledge had steeper identification functions on a /ba/-/da/ continuum than nonalphabetic Chinese controls. It is therefore instructive to equate reading experience and to compare poor readers with normal readers of the same reading level. Despite the fact that reading-age match is of lower age, poor readers were found to be worse in auditory and visual speech processing. This suggests that the perceptual deficit of the poor readers is rather independent of the attained reading level. The relevance of a speech perception deficit must thus not be sought in a direct link between actual reading skill achieved by poor readers and speech perception. Rather, phonological deficits related to speech processing may result in poor phonological representation as involved in the acquisition of grapheme/phoneme conversion skills.

Two other observations underscore the speech-specific nature of the observed deficit in contrast with the general auditory account of the poor readers’ difficulties. The very fact that the deficit is observed in the visual speech modality rules out that the phonological deficit might ultimately reduce to a deficit in auditory processing, simply because visual speech does not have an auditory component. Next, one might expect that if the handicap of poor readers was in general auditory processing, other nonauditory information would be employed to compensate for impaired auditory processing. For instance, Snowling, Goulandris, Bowlby, and Howel (1986) obtained evidence that poor readers relied more than normals on lexical information in identifying auditorily presented words. Poor readers, CA, and RA controls could repeat high-frequency words equally well. However, poor readers were worse than CA controls in low-frequency words, and they were worse than CA and RA controls in repeating nonwords. Thus, these results suggest that poor readers have problems in nonlexical processes which can be compensated for by lexical information. Similarly, Reed (1989) found that poor readers had larger lexical-status effects in phoneme identification. She presented syllables varying on a /b/-/d/ continuum between “badge” and “dadge.” When the initial phoneme was ambiguous, poor readers tended more than normals to report the word “badge” than the nonword “dadge.” In the present study, one might expect that if poor readers were having a problem in the general auditory domain, they would rely more than the controls on concurrent information from speech reading. On the other hand, if poor readers do have a speech-specific perception problem that is manifest in the auditory as well as in the visual modality, speech reading is unlikely to compensate for the auditory speech deficits. Our results show that there is no such compensatory effect of vision on audition since the visual influence in audition was not larger for the poor readers than the controls. If anything, it was even lower. Poor readers did not rely more than controls on visual information to compensate for their poor auditory skills. This result makes sense if
one assumes that for poor readers the speech-read information is less informative than for normals and it thus strengthens the idea that the poor readers' difficulties are located in the speech domain.

To summarize our findings, poor readers have difficulties in the perception of auditory and visual speech. Hence, a theory which is solely based on impairments in general auditory processing cannot account for these results. Rather, it seems that a speech-processing impairment is underlying reading difficulties. The deficit is probably located at a stage after the auditory and visual speech signals are combined. The comparison with the RA controls suggests that the perceptual deficit cannot be compensated for by attained reading level, at least in the restricted range that was used here.

REFERENCES


