In: R. J. Harris (Ed.), Cognitive processing in bilinguals. (pp. 413-426). Amsterdam: Elsevier. (1992).

Auditory and visual speech perception in alphabetic and non-alphabetic Chinese/Dutch bilinguals.

Beatrice de Gelder and Jean Vroomen
Tilburg University

Abstract

Chinese subjects were compared with Dutch controls on a speech-categorisation task in which they had to identify either auditory, visual, or audio-visual speech stimuli. Chinese subjects behave differently from native Dutch speakers on the categorisation of the auditory stimuli from a /ba-da/ continuum. Moreover, they were worse in lipreading /ba/ or /da/. When the Chinese subjects were divided into two subgroups as a function of whether they did master the alphabetic writing system, the results suggest a clear influence of alphabetic skills on speech sound categorisation.

Introduction

Learning to understand and speak a second language involves getting acquainted with a so far unfamiliar phonological system. That the mismatch between familiar and new phonological system can be a major obstacle is evident in one's intuitive and introspective evidence. Very little is known about how the phonology of a second language is acquired and the way native and second language sound systems interact. The categorical perception paradigm offers a tool for investigating an area in which linguistic experience might have effects.

Since its discovery in the late sixties (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967), the phenomenon of categorical perception offers an important source of evidence for arguing that speech is special. Categorical perception phenomena have been observed both with newborn infants and with adults (see Kuhl, 1990 for an overview). Yet, the basis of this well replicated phenomenon remains unclear and the inferences that can be drawn from the phenomenon to the representations and processes underlying speech perception remain hazardous.

The empirical reasons for this theoretical uncertainty are fairly straightforward. Very little is presently known about the influence of linguistic experience on speech processing. One crucial aspect of linguistic experience concerns phonological development itself. Very little is known about the development of the infants's categorisation ability over time. Data that would allow theories to build a bridge between the infant and the adult data are still lacking.

The infants' growing experience with language might affect its early categorisation ability. Clearly, the interpretation of the infants' categorical perception data needs to be put in the context of the development of the lexicon (Eimas, Miller, Jusczyk, 1990). The occurrence of developmental disorders offers a window into phonological development. Young retarded readers show a less steep categorical perception function than chronological and reading age controls (de Gelder and Vroomen, 1988; Werker and Tees, 1987). The poorer performance of this group is not due to poor alphabetic skills but suggests a phonological coding deficit manifesting itself on the occasion of reading acquisition. Nevertheless, experience with written language might be another source of phonological development and variability in categorisation performance.

The matter of the ontogenesis of phonological processing gets more complicated when we look also at visual speech categorisation. Kuhl and Meltzoff (1982) have shown that infants detect the equivalence between spoken and seen phonemes and have a preference for stimuli showing a cross-modal match. Yet, studies by Massaro, Thompson, Barron, and Laren (1986) show that there is a development in the ability to perceive visual speech. This increasing importance of visual speech over time is surprising given data which suggest that the auditory and the visual speech modality extract information from a common, abstract code (e.g., Massaro, 1987).

A second crucial aspect of linguistic experience concerns cross-linguistic influences. Within the categorical perception paradigm, cross-linguistic studies are a valuable source of information on the units of speech processing. At present, very few cross-linguistic studies of categorical perception are available. Here also, only the two extremes of the developmental continuum are represented. Young infants are sensitive to phonetic contrasts not active in their native language (Best et al., 1988; Werker and Tees, 1984), but have lost this ability by the end of the first year. From the available evidence it appears that experience with native language has an influence on phonetic boundaries (for a review, see Repp, 1984). For example, evidence for a cross-language difference between English and Japanese speakers was found for discrimination along the /r/-/l/ contrast (Miyawaki et al., 1975). Japanese subjects show no categorical perception of a contrast without phonemic value in their language. This leads one to ask whether acquiring a second language would modify this situation, or also, whether for the contrast that are active in the native language one would obtain comparable categorisation data across native speakers of two different languages.

It is not clear whether these data suggest a loss of the discrimination ability or a reorganisation within the phonological system in such a way that non-native language distinctions are no longer easily accessed. Werker and colleagues (1984) have investigated this question. Their findings suggest that when given enough practice and adapting the testing procedures adult listeners can regain their ability to distinguish non-native contrasts and return to a phonetic or even an auditory mode of responding as opposed to a phonemic one. Such data would favour the reorganisation view.

The present study adds new data to the cross-linguistic aspect of categorical perception. Indirectly it also throws light on the issue of development. Our aim was to find out (1) whether Chinese bilinguals speaking Chinese having acquired Dutch at adult age would differ from native Dutch speakers in auditory categorisation, visual identification of lipread speech, and the visual influence on auditory categorisation, and (2) whether acquaintance with an alphabetic script would make any difference above and beyond the differences observed between bi-lingual Chinese and native Dutch speakers.

There are major differences between the sound systems of Dutch and Chinese. Spoken Chinese (Cantonese, Mandarin) is a syllabic language consisting of combinations of consonants, diphthongs, consonantic vowels and consonants. Yet, there is a major degree of overlap between the Chinese and the Dutch sound systems. For example, the syllables /ba/ and /da/ figure prominently in Dutch as well as in Chinese. Against that background there is no reason to expect major differences in auditory categorisation.

Visual speech identification performance offers a complementary source of information on speech sound categorisation as vision and audition represent two autonomous but very closely linked input modalities for speech. Adults with normal hearing combine the auditory and visual speech information in normal circumstances as well as under impoverished conditions (e.g., Massaro, 1987). Individual differences in lipreading ability have been noted but should not been taken as an indication that lipreading skill is under the control of such general factors as intelligence or verbal ability (Jeffers and Barley, 1971). No comparative data are available on visual speech identification as a function of the specific phonology of the native language. A final aspect of bimodal input situations concerns the influence of the input of one modality on the processing of the information in the other modality. Available studies predict that the visual influence will be a function of the subjects' lipreading skills. In populations where there is a suggestion of a deficit in phonological processing we have observed reduced influence of the visual input on the auditory processing (de Gelder and Vroomen, 1988; 1990; 1991). Thus there is no reason to expect reduced visual influence purely as a consequence of non-native language experience.

However, a critical aspect of the present study concerns the possible influence of the scripts the subject masters on his performance in the categorical perception task. Of the few available studies engaging in cross-linguistic comparison none has paid attention to the differences in orthographies between the languages and the possible influence of these differences. In contrast with Chinese, Dutch has an alphabetic and surface orthography. It is well known that alphabetic reading skills go hand in hand with phonemic awareness skills. Non-alphabetic readers cannot perform phonemic segmentation tasks (Read, Yun-Fei, Hong-Yin, and Bao-Qing, 1986) while their Chinese bi-scriptal counterparts manage well (de Gelder, Vroomen, and Bertelson, 1990). Of course, performance on meta-phonological tasks as examined in the above studies is a different matter than speech categorisation. The present study makes a beginning with the study if this complex issue. On this occasion we wanted to check whether differences in orthographic skills between the two groups of native Chinese speakers might be a source of variability in the categorisation task. If so, bi-scriptal Chinese subjects would differ from monoscriptal ones on the auditory categorisation and on the visual identification task. The differences would be less though than the one observed between monoscriptal Chinese and Dutch speakers.

Method

Subjects. Two groups of subjects were tested, a group of 18 Chinese subjects and a group of 17 Dutch controls. The Chinese subjects (14 female and 4 male) came from Mainland China and a few had spend some years in HongKong. Most of them worked in Chinese restaurants. The mean age of the Chinese subject was 30;9 year (range from 16;0 to 53;0 year). They had spend about 5;6 year in the Netherlands. The control group consisted of 9 male and 8 female native speakers of Dutch (mean age = 33;6 year). None of the subjects reported any hearing or seeing deficit.

Stimuli. Subjects were presented a colour videotape prepared by Massaro (see Massaro and Cohen, 1983, experiment 2) showing a speaker seated in front of a wood panel. The speaker’s head filled about two thirds of the screen. The tape was made by copying the original video tape and replacing the natural sound track with synthetic speech. A nine-step /ba/ to /da/ auditory continuum was used. It 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 msec of the syllable, a set of nine 400-msec syllables covering the range of /ba/ to /da/ was created. During the first 80 msec, Formant 1 (F1) went from 250 Hz 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, 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 this case, the subject had to rely entirely on lip-reading. 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 160 trials. Subjects 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. Subject sat about 2 meter from the monitor. The audio was set at a comfortable listening level. 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 orally whether the speaker said /ba/ or /da/. The experimenter sat next to the subject and noted the responses and determined whether the subject was watching the screen at the time of the speech presentation. If this criterion was not met, the trial was disregarded.

Results

The results show that Chinese subjects are less categorical in their labelling of the /ba-da/ continuum and that they are worse in lipreading than the controls. The effect of the visual information on the bimodal labelling was not significantly different between the two groups.

We first consider the robustness of auditory categories. Figure 1 displays the 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 a /da/ articulation, separately for the two groups. An analysis of variance (ANOVA) was performed on the proportion of /da/ responses with group as between subjects variable and the auditory and visual factors as within subjects variables. As expected, the average proportion of /da/ responses increased as the auditory stimulus moved from the /ba/ to the /da/ end of the continuum [$F(8,264) = 271.0$, $p < .001$]. There was also a large effect of the visual variable [$F(2,66) = 135.3$, $p < .001$]. The average proportion of /da/ responses increased from .290 for the /ba/ visual stimulus, .529 for the neutral stimulus, to .772 for the /da/ visual stimulus. The interaction of the auditory and the visual variable was significant [$F(16,528) = 15.3$, $p < .001$], because the effect of the visual variable is larger at the ambiguous levels of the auditory continuum. These findings are in agreement with earlier studies (cf. Massaro, 1987).

---

figure 1 about here

The only effect in which the group variable was significant was a second order interaction among the group, auditory, and visual factors [$F(16,528) = 2.37$, $p < .002$]. In order to explore this, a separate ANOVA with group as between subjects factor was performed on the proportion of /da/ responses in the neutral visual articulation condition. These data are replotted in figure 2. There was a main effect of the auditory variable [F(8,264 = 356.8, $p < .001$], and a significant interaction between the groups and the auditory variable [$F(8,264) = 6.32$, $p < .001$]. Inspection of figure 2 suggests that the Chinese subjects were less categorical in the identification of the stimuli. That is, the labelling function for the Chinese subjects is less steep than that of the control group.

figure 2 about here

To analyze this, the individual proportions of /da/ responses were submitted to a logit transformation from which the values of the phoneme boundaries and the slopes are obtained (Finney, 1964). These 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. In a two-tailed *t-test*, there was a marginally significant difference in the place of the phoneme boundary [t(33) = 2.03, p = .06]. The mean boundaries were located at 4.34 and 4.93 stimulus units for the Chinese and the controls, respectively. The slopes of the labelling functions were significantly different [t(33) = 2.46, p < .02]. The mean slope (in degrees) was 40.8 for the Chinese subjects and 49.9 for the controls. Thus, Chinese subjects were less categorical in the identification of the /ba-da/ continuum.

We now turn to the results on visual speech identification. In the visual-only condition, the proportion of correct responses was determined. The Chinese subjects were worse in lip-reading than the controls [t(33) = 2.93, p < .01]. The average percentage of correct responses was 80.3 for the Chinese and 93.7 for the controls.

In order to measure the contribution of the visual source to the bimodal speech events, the individual percentage of /da/ responses given a visual /ba/ was subtracted from the percentage of /da/ responses given a visual /da/. The mean contribution of the visual source was 41.6 for the Chinese group and 54.0 for the controls. This difference felt short of significance [t(33) = 1.68, p = .10].

A correlation between subjects' proportion correct in the visual-only condition and the steepness of the labelling function in the audio-visual condition with a neutral articulation was computed. The correlation between these variables was significant for the combined results, r = .521, p < .001. For the two groups separately, the correlation was positive but felt short of significance in the control group (Chinese, r = .519, p < .02; control group, r = .255, p = .16). In general, the results show a positive relationship between the ability to lipread and the degree of categorical speech perception.

As already argued, we wanted to know whether alphabetic reading skills has an influence on speech categorisation. For that reason, the Chinese group was divided into two sub-groups: one group consisted of non-alphabetic Chinese subjects who could not read an alphabetic script (Dutch or any other alphabetic script), but who were fluent readers of the Chinese logographic script; the other group of Chinese subjects could read logographic Chinese as well as the Dutch alphabetic script (See table 1 for details of the subjects). The group was split on the basis of an alphabetic reading test which required reading aloud 20 mono and bisyllabic pseudo words. Pseudo words were created by changing one letter from real words. Subjects who could not read more than 50 percent correct were assigned to the non-alphabetic group; subjects who read more were assigned to the alphabetic group. The mean percentage correct was 13.6 for the non-alphabetic group and 82.0 for the alphabetic group. Subjects were also given a reading test in which they had to read 20 Dutch mono- and bisyllabic high frequency words. The mean percentage correct on this reading test was for the non-alphabetic group 49.3 percent and for the alphabetic group 96.5 percent. In a logographic reading test subjects had to read 20 logographs of nouns. Both Chinese groups performed errorless on this task. The non-alphabetic group consisted of 1 male and 7 female with a mean age of 29;10 year. They had been living for about 2;6 year in the Netherlands. The alphabetic Chinese group consisted of 3 male and 7 female subjects (mean age = 31;6). These subjects had been to Dutch reading classes and had acquired a basic alphabetic reading skill. To assess the reading level of the control group, these subjects were given a standardized speed reading test (Brus and Voeten, 1973). On average they read 100.1 real words in 1 minute and were all within normal range.

Table 1

Details of the Chinese Sub-Groups.

			Reading % correct		
Group	N	Age	Real	Pseudo	Logographic
Non-alphabetic	8	29;10	49.3	13.6	100
Alphabetic	10	31;6	96.5	82.0	100

In the following analyses, we compared the three groups (non-alphabetic Chinese, alphabetic Chinese, and controls) on their labelling of the stimuli. Figure 3 displays the proportion of /da/ responses separately for the three groups. Note that the control group is the same as in the previous analyses. An ANOVA was performed on the proportion of /da/ responses with group as between subjects variable and the auditory and visual factors as within subjects variables. As in the previous analyses, the auditory [$F(8,256) = 246.4$, $p < .001$] and visual [$F(2,64) = 109.7$, $p < .001$] stimulus effects were significant, as was the interaction between these two variables [$F(16,512) = 13.0$, $p < .001$]. There was also a significant second order interaction of the group by auditory by visual factor [$F(32,512) = 1.71$, $p < .01$].

figure 3 about here

A separate ANOVA with group as between subjects variable was performed on the proportion of /da/ responses in the neutral visual articulation condition. These data are presented in figure 4. There was a main effect of the auditory variable [F(8,256 = 305.1, $p < .001$], and a significant interaction between the groups and the auditory variable [$F(16,256) = 3.27$, $p < .001$]. Submitting the data to a logit transformation revealed that there were no significant differences in the place of the phoneme boundary [$F(2,32) = 2.05$, $p = .14$]. The mean boundaries were at 4.27, 4.40, and 4.93 stimulus units for the non-alphabetic Chinese, alphabetic Chinese and the controls, respectively. The slopes of the labelling functions were different [$F(2,32) = 3.83$, $p < .05$]. The mean slopes (in degrees) were 37.3, 43.6 and 49.9 for the non-alphabetic Chinese, alphabetic Chinese, and controls, respectively. *Post hoc* analyses (Fisher's LSD) showed that the non-alphabetic Chinese were less categorical in their labelling than the controls ($\alpha = .05$), but they did not differ from the alphabetic Chinese. Alphabetic Chinese did not differ significantly from the controls.

---

figure 4 about here

There was also a significant difference in lipreading performance between the three groups [$F(2,32) = 10.9$, $p < .001$]. The average percentage of correct responses was 70.0 for non-alphabetic Chinese, 88.5 for alphabetic Chinese, and 93.7 for the controls. *Post hoc* analyses showed that non-alphabetic Chinese performed worse than alphabetic Chinese and the controls (Fisher's LSD, $\alpha = .01$).

The mean contribution of the visual source in the bimodal conditions was 36.0 for the non-alphabetic group, 46.1 for the alphabetic Chinese, and 54.0 for the controls. This difference felt short of significance [$F(2,32) = 1.88$, $p = .16$]. Table 2 summarizes the results.

Table 2

Performance on the Speech Identification Test for the Three Groups.

Group	Steepness auditory categorization	Lipreading % correct	Visual dominance
Non-alphabetic Chinese	37.3	70.0	36.0
Alphabetic Chinese	43.6	88.5	46.1
Controls	49.9	93.7	54.0

Discussion

The present study suggests that experience with native language leaves its mark on the way speakers make speech sound categorisations, irrespective of whether speech sounds are presented auditorially, visually, or bimodally. Speakers learning a second language in adulthood remain different from the native speakers of that language. At the same time, the data suggest a second kind of effect of experience with language, this time occurring on the occasion of learning to read the alphabetic script of the non-native language. This second experience appears to narrow the gap between native and non-native speakers, creating an intermediate category between the two other groups. If so, the possibility that orthographic skills exercise an influence on speech categorisation must be taken seriously.

The most surprising fact suggested by these data is that the phoneme boundary is less marked for Chinese subjects than it is for the native Dutch speakers. What aspects of linguistic experience might explain this fact? How might linguistic experience affect speech sound perception and what factors might drive it? A central factor so far appears to be the role of a non-native contrast in the speaker's native language. A recent proposal by Best, McRoberts, and Sithole (1988) distinguishes four alternatives: (a) the two stimuli are assimilated to a single category; (b) the two stimuli are assimilated to opposing native language categories; (c) one stimulus is privileged because taken as representing a better example of the category; (d) neither member is assimilated to a native category. Unlike in the study by Miyawaki et al. (1975) the present study did not present the subjects with a stimuli without phonemic value in their native language. The syllables /ba/ vs. /da/ are as common in Chinese as they are in Dutch. Moreover, the subjects were all familiar with spoken Dutch.

Is the categorical perceptual difference observed here suggestive of a difference in phonemic vs. phonetic or acoustic discrimination difference between the groups? It is clear that the audiovisual presentation of the stimuli strongly induces a speech mode of listening even in the auditory-only condition where the speakers' face was not moving. But more convincingly, the fact that categorisation differences between the three groups show up in the auditory as well as in the visual presentation modality strongly suggests that cognitive-linguistic and not peripheral differences exist between the groups.

Two other aspects of the data are straightforward. Steepness of auditory categorisation is a good predictor of success on the lipreading task, at least in neurologically unimpaired subjects. The data from visual speech identification offer new support for that view. The results on auditory speech categorisation do not lead us to expect a very high performance on the visual discrimination task in the two groups of Chinese subjects. Likewise, the results on the influence of visual information on auditory categorisation are predicted by earlier studies. Werker (1991) mentions the finding of a similar cross-linguistic difference in audio-visual speech perception and an effect of the visual information that is different as a function of the native language of the subjects (French vs. English). The representation activated as a consequence of a mismatch between the auditory and the visual input, the so called McGurk illusion (McGurck and MacDonald, 1976) is thus different for English and for French speakers watching the same stimuli.

Finally, we turn to the influence from experience with written language, e.g. learning to read Dutch. It is clear from the present data that the categorisation performance of bi-scriptal Chinese speakers differs from that of the mono-scriptal Chinese. Yet, at present there is little known about whether reading skill or orthography influences speech perception. But there is reason to believe that cross-modal influences between spoken language representation and written language representation do occur. In contrast, what is well documented is that alphabetic scripts promote phonemic awareness (Bertelson and de Gelder, 1990 for and overview). The syllabic phonology of Chinese does not encourage an explicit segmental representation of speech. Whether we are dealing with an effect of having acquired basic reading skill or with an effect of the teaching methods is not clear. In the domain of metaphonological skills the same question crops up and is currently still debated in research on reading acquisition with young children (Bertelson and de Gelder, 1990). It is likely that in the course of reading instruction a good deal of attention has been devoted to building up awareness of internal sounds structures. Although the speech categorisation task is not a metaphonological awareness task, subjects may have profited from their recently acquired skills at segmental analysis and they may have become able to focus attention in a more structured way on phonemic differences.

Acknowledgements. We are indebted to D. Massaro for use of the test materials and to E. Vianen and M. van Zon for assistance with testing the subjects. This work was partly supported by the Human Frontiers of Science Program. We are grateful to O. Tzeng for comments on the manuscript.

References

Bertelson, P., and de Gelder, B. (1990). The emergence of phonological awareness. In I. Mattingly and M. Studdert-Kennedy (Eds.), *The motor theory of speech perception* (pp. 393-412). Hillsdale, NJ: Lawrence Erlbaum Associates.

Best, C. T., McRoberts, G. W., and Sithole, N. N. (1988). The phonological basis of perceptual loss of nonnative contrasts: Maintenance and discrimination among Zulu Clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance,* ***14***, 245-260.

Brus, B. T., and Voeten, M. J. M. (1973). *Een minuut test. Vorm A en B*. Nijmegen: Berkhout.

de Gelder, B., and Vroomen, J. (1988, August). *Bimodal speech perception in young dyslexics*. Paper presented at the 6th Australian Language and Speech Conference. Sydney.

de Gelder, B., and Vroomen, J. (1990, July). *Poor readers are poor lipreaders*. Paper presented at the Applied Psychology Conference, Kyoto.

de Gelder, B., and Vroomen, J. (1991). Phonological deficits beneath the surface of reading- acquisition problems. *Psychological Research,* ***53***, 88-97.

de Gelder, B., Vroomen, J., and Bertelson, P. (1990, january). *Segmental abilities in bi-scriptal Chinese*. Paper presented at the Experimental Psychology Society, London.

Finney, D. J. (1964). *Probit analysis*. Cambridge: Cambridge University Press.

Jeffers and Barley (1971). *Speechreading*. Springfield, Illnois: Thomas

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America,* ***67***, 905-917.

Kuhl, P. (1990). On babies, birds, modules, and mechanisms: A comparative approach to the acquisition of vocal communication. In R. J. Dooling and S. H. Hulse (Eds.), *The comparative psychology of audition* (pp. 379-419). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, **74***, 431-461.

Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry.* Hillsdale, NJ: Lawrence Erlbaum Associates

Massaro, D. W., Thompson, L. A., Barron, B., and Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology, **41***, 93-113.

Massaro, D. W., and Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, **9***, 753-771.

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

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (1975). An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception and Psychophysics, **17***, 9-16.

Read, C., Yun-Fei, Z., Hong-Yin, N., and Bao-Qing, D. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition, **24***, 31-44.

Repp, B. H., (1984). Categorical perception: Issues, methods and findings. In N.J. Lass (Ed.), *Speech and language: Advances in basic research and practice, Vol. 10* (pp. 310-330). New York: Academic Press.

Werker, J. F. (1991). Ontogeny of speech perception. In I. G. Mattingly and M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (91-110). Hillsdale, NJ: Lawrence Erlbaum Associates.

Werker, J. F. and Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganisation during the first year of life. *Infant Behavior and Development, **7***, 49-63.

Werker, J. F. and Tees, R. C. (1987). Speech perception in severely disabled and average reading disabled children. *Canadian Journal of Psychology, **41***, 48-61.