The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/22691

Please be advised that this information was generated on 2017-03-31 and may be subject to change.
Perception of Voicing Cues by Children With Early Otitis Media With and Without Language Impairment

Research on the relationship between early otitis media with effusion (OME), language impairment, and central auditory processing has been equivocal. Identification and discrimination tasks provide us with a sensitive method of assessing speech perception on both an auditory and a phonetic level. The present study examined identification and discrimination of initial bilabial stop consonants differing in voicing by 9-year-old children with a history of severe OME. The groups studied were controlled for language impairment. The ability of these children to perceive major and minor voicing cues was examined using multiple voicing cues. Long-term effects of OME were found for both identification and discrimination performance. Children with OME produced an overall inconsistency in categorization, which suggests poorer phonetic processing. Discrimination was measured by means of "just noticeable differences" (JND). Children with early OME experience demonstrated a greater mean JND than children without early OME experience. Finally, in cases of language impairment with early OME, there was no additional deterioration of auditory or phonetic processing. It appears that either early OME or language impairment can lead to poorer perception.

KEY WORDS: otitis media, language impairment, speech perception, children, voicing
relationship between early OME and later language ability in the same population as the present study. In addition to general language ability, they studied higher-order linguistic constructs such as phonologic awareness and word discrimination. Their conclusion was that a history of OME did not have any negative consequences for language performance. A possible interpretation of the conflicting evidence cited thus far is that long-term effects of OME do not show up in language learning processes but are restricted to lower-order speech perception processes.

Welsh, Welsh, and Healy (1985) demonstrated that a significant number of children with early OME had difficulties in central auditory processing using competing sentences, binaural fusion tasks, filtered speech, and compressed speech. They did not find an effect using rapidly alternating speech. Sak and Ruben (1981) found long-term effects of OME only on the auditory reception subtest of the Illinois Test of Psycholinguistic Abilities. Schilder, Snik, Straatman, and van den Broek (1994) examined children from the same birth cohort as the subjects in the present study. They reported a significant effect of OME on speech-in-noise recognition. No effects were reported for filtered speech, binaural fusion, dichotic speech, or auditory memory.

Hoffman-Lawless, Keith, and Cotton (1981) and Locke (1980a, 1980b) stated that the major methodological shortcomings of auditory studies had to do with the use of insensitive test procedures and confounding of OME with language impairment. Eimas and Clarkson (1986) and Clarkson, Eimas, and Cameron-Marean (1989) used speech continua in combination with identification and discrimination tasks in the assessment of speech processing in children with a history of OME. These tasks have proven to be quite sensitive in assessing central auditory functioning (Repp, 1984). Clarkson et al. (1989) studied the perception of voicing in children age 5.5 years. Their identification and discrimination tasks were based on stimuli varying in one cue: voice onset time (VOT). They found that children with early OME showed poorer discrimination performance. However, only children with language impairments and early OME showed poorer categorization. To further disentangle the effect of OME and language delay a full factorial subject design would be required. However, Clarkson et al. (1989) studied only three groups of subjects: (a) children with OME and language delay, (b) children with OME only, and (c) children without OME or language delay.

To allow for subtle perceptual effects to appear, in the present study we used identification and discrimination tasks with speech continua that varied by small acoustic steps. In an identification task requiring a phonemic judgment, decisions are based on the phonetic properties. In a discrimination task, decisions may be based on both phonetic and auditory properties (Pisoni, 1973; Pisoni & Tash, 1975). To take into account the potentially confounding effect of language impairment, we employed a factorial design consisting of two factors: (a) history of recurrent OME, and (b) language impairment.

Decisions regarding the voiced-voiceless distinction are based on the perceptual integration of several distinct acoustic properties. The major acoustic cue carrying voicing information in Dutch is voice onset time (VOT) (Lisker & Abramson, 1964; Sils & Cohen, 1989a). In addition to the major cue VOT, other minor cues appear to contribute to the voiced-voiceless distinction (Lisker, 1975; Massaro, 1975; Schouten & Pols, 1983; Sils & Cohen, 1969a, 1969b). Multiple cues can combine to signal a phonetic distinction and thereby enhance phonetic clarity and discriminability. Conflicting multiple cues decrease clarity and discriminability (Best, Morrongiello, & Robson, 1981; Fitch, Halwes, Erickson, & Liberman, 1980; Repp, 1981a). Because categorization strategies contribute to discrimination performance, the sensitivity of the discrimination task can be increased by using a cooperating as well as a conflicting cues continuum. In the latter condition the effect of categorization on perceptual discrimination is reduced as compared to the cooperating cues condition, whereas acoustic distances between stimuli are equal.

The topic of auditory versus phonetic processing of speech has been neglected in research on the long-term effects of OME. The importance of auditory and/or phonetic processing abilities for the development of language has been studied in children with normally developing language (e.g., Nittровер & Studdert-Kennedy, 1987; Sussman, 1993a; Sussman & Carney, 1989) and in children with language impairments (e.g., Elliott & Hammer, 1885; Sussman, 1983b; Tallal & Piercy, 1974, 1975).

The use of two continua with (a) cooperating cues and (b) conflicting cues not only increases sensitivity, but also contributes to the differentiation of auditory from phonetic processing in the discrimination task. Both continua vary from [bak] to [pak] according to differences in VOT. In the case of cooperating cues, the major and minor cues both lead to the same percept and thereby enhance perceptual clarity. In the case of conflicting cues, the major and minor cues lead to percepts in opposite phonemic directions and normally elicit phonetic neutralization. Using both conflicting and cooperating cue conditions enables us to differentiate auditory from phonetic processing. In the discrimination tasks, the absolute acoustic inter-stimulus differences of each pair of the cooperating cues continuum are equal to the counterpart pair of the conflicting cues continuum. If discrimination is based solely on auditory information, then there should be total similarity between the discrimination curves in the cooperating cues and the conflicting cues condition. However, if discrimination is (partly) based on phonetic processing strategies, a difference between conditions should occur.

To summarize, our study extends previous research in several ways by (a) implementing a prospective cohort design, using a reliable OME documentation procedure (tympanometric screening of OME every 3 months between the ages of 2 and 4 years) and controlling for risk factors (e.g., intelligence, grade level in school, ventilating tubes), (b) maintaining a complete four-cell factorial design (early OME and language impairment), (c) using sensitive measurement materials and procedures (speech continua and cooperating and conflicting cues conditions), (d) separating auditory from phonetic processing (by using identification and discrimination tasks with both cooperating and conflicting cues), and (e) studying voicing as a multidimensional feature determined by a major cue (VOT) and a number of minor cues.
We expected that, because of recurrent periods of partial auditory deprivation in early childhood, long-term effects of OME would exist, either in auditory or phonetic processing. In addition, we expected that language impairment also would result in poorer processing of speech.

Experiment 1 focuses on stimulus construction. The conflicting power of combined minor voicing cues was determined. In Experiment 2 the speech perception abilities in children with OME, with and without language impairment, would result in poorer processing of speech.

EXPERIMENT 1

Before generating the speech continua a study was conducted to determine the conflicting power of minor voicing cues. In comparison with the English voicing distinction, the Dutch voicing distinction is differently distributed along the VOT dimension. Whereas the English values fall into a range between 0 and +100 ms, Dutch stop-category values fall into a range between −100 and +10 ms (Lisker & Abramson, 1964).

We made use of four minor cues: (a) the length of the noise burst; (b) the intensity of the noise burst; (c) the formant transition duration of F1, F2, and F3; and (d) the range of the frequency shift of F1.

Method

Subjects

Subjects were 10 Dutch adults with normal hearing—5 men, 5 women; mean age 36.4 (years: months), with a range of 25:4–56:4. All of the subjects met the following selection criteria: (a) normal bilateral pure tone audiometric thresholds (no greater than 20dB HL) at 250, 500, 1000, 2000, and 4000 Hz (ISO, 1985) immediately before testing; (b) Dutch as the native language; and (c) no enrollment in otological medical/surgical treatment.

Stimuli

The starting point for the stimuli was a naturally produced syllable, [buk], spoken by an adult male. This utterance was band-pass filtered between 40 and 5000 Hz (60 dB/octave attenuation), then digitized at a rate of 10 kHz with a DASH-16 data-acquisition card (12 bits resolution). The Interactive Laboratory System (ILS, V6.1, 1989) was used to smooth the spectral structure. For smoothing, the vowel (formant transitions plus steady state vowel) was analyzed with pitch synchronous linear predictive coding (covariance method: pre-emphasis factor .98, Hamming window), yielding 12 reflection coefficients (Markel & Gray, 1976). The locations of the spectral peaks, their bandwidths, and their intensities were estimated by transforming the reflection coefficients into autoregressive coefficients and then performing a fast Fourier transformation (FFT).

The formant frequencies were interactively adjusted. The first formant (F1) started at 400 Hz and linearly increased to a center frequency of 750 Hz by 20 ms. The second (F2) and third (F3) formants started at 1000 and 2150 Hz, respectively, and linearly increased to center frequencies of 1150 and 2500 Hz, respectively, by 52 ms. The sampled data were resynthesized with a pitch synchronous synthesis procedure by transforming the changed reflection coefficients to inverse filter coefficients. The filter was excited using a pulse train. The resynthesized vowel part was spliced onto the initial stop consonant. The temporal structure was adjusted by setting the length of the burst to 10 ms. The intensity of the burst was −11.4 dB (relative to the sound level of the vowel).

If stimuli unambiguously belong to the same phonetic category, phonetic neutralization in discrimination tasks disappears (Repp, 1981b). This suggests that perceptual neutralization occurs near phoneme boundary regions. In the first step, two syllables with VOT values near the phoneme boundary were selected in the following way. We constructed 11 stimuli differing in about 10-ms VOT steps between −71.1 and +24 ms from /buk/ (i.e., box) to /puk/ (i.e., package). The phonemic quality of the tokens was checked in a pilot study by having 10 adults with normal hearing label 10 repetitions of each of the 11 stimuli presented in random order. The mean phonemic boundary occurred at a VOT of −11.31 ms (a value in conformity with the Dutch language). The nearest unambiguously labeled voiced stimulus /buk/ had a VOT of −19.1 ms (mean percentage /b/ of 90%). The nearest unambiguously labeled voiceless stimulus /puk/ had a VOT of 0 ms (mean percentage /p/ of 99.33%). These two speech tokens were then used as stimulus when determining the conflicting potential of the minor cues.

We incorporated four minor voicing cues into our stimuli: (a) the intensity of the noise burst (low = voiced, high = voiceless); (b) the length of the noise burst (short = voiced, long = voiceless); (c) the formant transition duration of F1, F2, and F3, (long = voiced, short = voiceless); and (d) the range of the frequency shift of F1 (large = voiced, small = voiceless). Four levels were created involving all four minor cues for each of the reference stimuli. This resulted in a total of eight stimuli (see Table 1 for the exact stimulus specifications and Figure 1 for an abstract representation of the stimuli).

Procedure

Stimuli were recorded and played back using an Ampex 467 DAT-tape on a Grundig Fine Arts Digital Audio Tape recorder (type DAT-9000: 16 bit D/A converter, 2-fold oversampling, sampling frequency 48 kHz). Presentation was via a Beyerdynamic closed headphone (Type DT770). Playback level was set at a listening level of 70dB HL, a level judged by all subjects to be comfortable. Subjects were tested in a quiet room.

Each of the 10 subjects was examined for one session of 30 minutes. The identification task was based on a two-alternative forced choice procedure (one stimulus was presented and subjects responded with one of two response alternatives) and consisted of five replications of each of the eight stimuli presented in random order as five series of
TABLE 1. Stimulus specifications for conflicting minor cues.

<table>
<thead>
<tr>
<th>Voiced (VOT = -19.1 ms)</th>
<th>Burst intensity (dB relative to vowel)</th>
<th>Burst length (ms)</th>
<th>Transition duration (ms)</th>
<th>Frequency shift F1 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (original)</td>
<td>-11.4</td>
<td>10</td>
<td>20 52</td>
<td>350 (400–750)</td>
</tr>
<tr>
<td>1</td>
<td>-7.3</td>
<td>14</td>
<td>20 20</td>
<td>0 (750–750)</td>
</tr>
<tr>
<td>2</td>
<td>-3.1</td>
<td>22</td>
<td>20 20</td>
<td>0 (750–750)</td>
</tr>
<tr>
<td>3</td>
<td>1.1</td>
<td>30</td>
<td>20 20</td>
<td>0 (750–750)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Voiceless (VOT = 0 ms)</th>
<th>Burst intensity (dB relative to vowel)</th>
<th>Burst length (ms)</th>
<th>Transition duration (ms)</th>
<th>Frequency shift F1 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (original)</td>
<td>-11.4</td>
<td>10</td>
<td>20 52</td>
<td>350 (400–750)</td>
</tr>
<tr>
<td>1</td>
<td>-14.6</td>
<td>8</td>
<td>52 52</td>
<td>550 (200–750)</td>
</tr>
<tr>
<td>2</td>
<td>-17.7</td>
<td>4</td>
<td>52 52</td>
<td>550 (200–750)</td>
</tr>
<tr>
<td>3</td>
<td>-20.8</td>
<td>0</td>
<td>52 52</td>
<td>550 (200–750)</td>
</tr>
</tbody>
</table>

eight stimuli. The stimuli were separated by an interstimulus interval of 2500 ms. Subjects identified the initial speech sound of the stimulus by writing it down on a form designed for this purpose. All subjects were required to pass the criterion of correctly identifying four out of five presentations of each of the two original stimuli /bak/ and /pak/. Out of a total of 10 subjects, 9 met the criterion. The perceptual data from these 9 subjects (5 men, 4 women) were then analyzed.

Results

The effect of the minor cues is presented in Figure 2. The solid lines indicate the effect of the minor cues on the perception of /b/ (VOT = -19.1 ms). The dashed lines indicate the effect of the minor cues on the perception of /p/ (VOT = 0 ms). The percentage /b/-judgments for the two reference stimuli changed as a function of the minor cues. The slope (b) of the phonemic shift (estimated with linear regression analysis) corresponds to the power of the combined minor cues to elicit phonetic neutralization. Slope values were significantly different from zero for both /buk/ and /pak/ [b = -9.33, t(34) = -2.414, p = .021 and b = -10.89, t(34) = -4.051, p < .001, respectively].

In this experiment the conflicting power of the minor voicing cues was established. The results validated the choice of the amount of conflicting and cooperating voicing

FIGURE 1. Abstract representation of the stimuli in Experiment 1. The waveform at the start of each of the upper graphs represents the amount of glottal pulsing before the opening of the mouth (VOT = -19.1 ms). The filled-in rectangles designate burst properties (width = burst length, and height = burst intensity). The numbers in the upper right corner of each graph correspond to the levels of conflicting voicing information.
EXPERIMENT 2

In this experiment speech perception of voicing in children with OME with and without language impairment was assessed.

Method

Subjects

The factorial design consisted of four cells formed by the presence and absence of language disabilities and by the presence and absence of a history of OME: (a) OME/Language Impairment; (b) OME only; (c) Language Impairment only; and (d) neither OME nor Language Impairment, that is, the control group (C).

Subjects were selected from the Nijmegen Otitis Media Group, a birth cohort of over 1,400 children born in Nijmegen (The Netherlands) between September 1, 1982, and August 31, 1983. These children were involved in an earlier study of the efficacy of screening preschoolers for OME (Zielhuis, Rach, & van den Broek, 1989). Over 1,300 children were screened for OME using tympanometry every 3 months between the ages of 2 and 4 years, which led to nine sessions per child. The tympanometric results were classified into four types according to a modified version of the method described by Jerger (1970): (a) type A: maximum compliance >0.2 ml at an ear canal pressure of −99 dPa or higher, (b) type C1: maximum compliance >0.2 ml at an ear canal pressure between −100 dPa and −199 dPa, (c) type C2: maximum compliance >0.2 ml at an ear canal pressure between −200 dPa and −399 dPa, and (d) type B: maximum compliance <0.2 ml at an ear canal pressure below −400 dPa. Only a type B tympanogram (i.e., little or no change in compliance of the middle ear) was taken as evidence of OME. At about 8 years of age, over 300 of the children participated in a follow-up study (see Grievesink et al., 1993). This study included several language development tests and questionnaires from the teachers and parents, respectively. The subjects in the present study were selected from this group of children.

The presence/absence of a history of OME was determined by the frequency of type B tympanograms for both ears simultaneously (presence: at least four times; absence: zero times). The presence/absence of language impairment was determined by four variables. Two subtests of the "Language Tests for Children" ("Taaltests voor Kinderen," van Bon, 1984) were used: (a) the "Word Forms Production" test ("Woordvormen produktie" test), which is a productive morphological test concerned with knowledge of word forms, and (b) the "Concealed Meaning" test ("Verzwevend Betekenis" test), which is a receptive test concerned with the child’s understanding of the nonexplicit contents of sentences. Both tests have been shown to be related to different specific language factors (van Bon, 1984). The other two variables used to determine the presence/absence of a language impairment were the data from questionnaires administered to (c) the parents and (d) the teachers. Factor analyses on each questionnaire yielded factor scores of a set of preselected items loading high on general linguistic competence (see the Appendix for details of the analyses used).

In Table 2 the means for the variables used to select the subjects are presented. The means for the OME/Language Impairment group and the Language Impairment group on the "Word Forms Production" test, the "Concealed Meaning" test, the teacher questionnaires, and the parent questionnaire were all more than three standard deviations from the means for the control group. The means for the OME group were all within a single standard deviation from the mean for the control group.

A two-factor analysis of variance (OME [absence/presence] × Language Impairment [absence/presence]) was used to verify the significance of differences in the selection scores. This resulted in a significant effect of OME for frequency of type B tympanogram F(1, 40) = 321.2, p <

### Table 2. Means for variables used to select three experimental groups and one control group.

<table>
<thead>
<tr>
<th></th>
<th>OME/LI</th>
<th>OME</th>
<th>LI</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys/Girls</td>
<td>6/4</td>
<td>2/10</td>
<td>3/7</td>
<td>5/7</td>
</tr>
<tr>
<td>Age</td>
<td>8:11</td>
<td>9:0</td>
<td>9:2</td>
<td>8:11</td>
</tr>
<tr>
<td>OME</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freq. type B</td>
<td>4.9</td>
<td>5.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Language Impairment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;Word Forms Production&quot;</td>
<td>4.0</td>
<td>6.6</td>
<td>3.6</td>
<td>7.3</td>
</tr>
<tr>
<td>&quot;Concealed Meaning&quot;</td>
<td>3.3</td>
<td>7.5</td>
<td>3.5</td>
<td>6.7</td>
</tr>
<tr>
<td>Questionnaire teachers</td>
<td>−0.2</td>
<td>0.6</td>
<td>−0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Questionnaire parents</td>
<td>−0.7</td>
<td>0.5</td>
<td>−0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Note. OME = history of Otitis Media with Effusion; LI = Language Impairment. The scores on the "Word Forms Production" test and the "Concealed Meaning" test are standard scores; a score above 5.0 indicates higher than normal linguistic competence, and a score below 5.0 indicates lower than normal linguistic competence. The scores on the questionnaires are factor scores based on a set of selected items.
.001) and significant effects of Language Impairment for the "Word Forms Production" test, the "Concealed Meaning" test, and the questionnaires filled out by the teachers and parents [F(1, 40) = 106.0, p < .001; F(1, 40) = 107.29, p < .001; F(1, 40) = 14.13, p < .001 and F(1, 40) = 22.36, p < .001, respectively]. There was no significant interaction between OME and Language Impairment for either frequency of Type B tympanogram [F(1, 40) = 0.58, p = .45], "Word forms Production" test [F(1, 40) = 3.49, p = .07], the "Concealed Meaning" test [F(1, 40) = 1.84, p = .18], and the questionnaires filled out by the teachers and parents [F(1, 40) = 0.0, p = .98 and F(1, 40) = 0.63, p = .43, respectively]. These control statistics confirmed that the cells of the factorial design were appropriately filled.

Further, all of the children met the following selection criteria: (a) normal bilateral pure tone audiometric thresholds (no greater than 20 dB HL) at 250, 500, 1000, 2000, and 4000 Hz (ISO, 1985) immediately before testing; (b) normal speech-in-noise recognition (S/N ratio 0 dB, presented at 70 dB HL), measured on two series of 10 monosyllabic words immediately after testing (scored was the percentage of correctly perceived phonemes; a score within 1 standard deviation from the mean of the control group was considered normal; none of the experimental groups differed significantly from the control group); (c) sufficient intellectual capacities (as measured by the Coloured Progressive Matrices for Children, Raven, 1965); a standard score of at least 50 was considered sufficient; (d) no bilingualism; (e) Dutch as the native language; (f) no ventilating tubes inserted in the tympanic membrane at the time of testing nor during earlier phases of OME; (g) no enrollment in medical/surgical treatment; (h) no severe speech production problems; and (i) no missing values on any of the selection variables. All subjects attended regular schools and were in grade levels appropriate for their age.

Stimuli: Generating the Two Continua

Two eight-step /b-p/ continua were generated. By manipulation of the linear predictive coding parameters and resynthesis in combination with modifying parts of the oscillographic waveform, the consecutive stimuli of both continua were constructed.

The cooperating-cues continuum consisted of stimuli with major and minor voicing cues that represented the same voicing state. This continuum started with /b/ and moved in the direction of /p/. The conflicting-cues continuum consisted of stimuli with major and minor cues working in opposite phonemic directions.

In order to construct appropriate cooperating cues and conflicting cues voicing continua, stimulus specifications have to be chosen carefully. Two major considerations constrain the choice of the stimulus specifications: (1) the specifications must be within the limits defined by the acoustic effects of natural articulation, i.e. the parameters chosen to synthesize speech must reflect normal human articulation, and (2) the specifications should permit perceptual neutralization to take place. We chose reference stimuli with conflicting minor cues specifications that elicited 20%-25% /p/-responses for the /buk/ reference stimulus and 20%-25% /b/-responses for the /pak/ reference stimulus. In Figure 2 the levels indicated by '2' were judged to offer the appropriate stimulus conditions. These levels were fixed at VOT-values outside the values of the reference stimuli and linearly interpolated at VOT-values between the values of the reference stimuli. Hence, the minor cues were varied in a block. Table 3 lists the stimulus specifications for both the cooperating cues and the conflicting cues continua.

### TABLE 3. Stimulus specifications for the cooperating-cues and conflicting-cues continua.

<table>
<thead>
<tr>
<th></th>
<th>VOT (ms)</th>
<th>Burst intensity (dB rel. to vowel)</th>
<th>Burst length (ms)</th>
<th>Transition duration (ms)</th>
<th>Frequency shift F1 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperating cues</td>
<td></td>
<td></td>
<td></td>
<td>F1</td>
<td>F2 + F3</td>
</tr>
<tr>
<td>1</td>
<td>-52.7</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>-40.9</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>-29.1</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>4</td>
<td>-19.1</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>5</td>
<td>-10.8</td>
<td>-10.4</td>
<td>13</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>0.0</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>8.0</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>16.0</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Conflicting cues</td>
<td></td>
<td></td>
<td></td>
<td>F1</td>
<td>F2 + F3</td>
</tr>
<tr>
<td>1</td>
<td>-52.7</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>-40.9</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>-29.1</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>-19.1</td>
<td>-3.1</td>
<td>22</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>-10.8</td>
<td>-10.4</td>
<td>13</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>0.0</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>8.0</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>8</td>
<td>16.0</td>
<td>-17.7</td>
<td>4</td>
<td>52</td>
<td>52</td>
</tr>
</tbody>
</table>

Note. VOT = Voice Onset Time.

*A Due to the use of a pitch synchronous LPC procedure, the appropriate value of 36 ms was not possible. Instead we chose a value of 31 ms.*
Procedure

The stimuli were recorded as in Experiment 1 and played back using a portable AIWA Digital Audio Tape Recorder (Type AIWA HD-S1: bit-stream D/A converter). They were presented via the same headphones as in Experiment 1. The playback level was set at the level judged as comfortable by the subject (always close to 70 dB HL). The subjects were tested in a quiet room at the school they were attending.

Each child was examined in a one-hour session. In order to accustom the child to the manipulated speech, he or she first heard four repetitions of the endpoint stimuli from the two different continua without having to respond. This was followed by a training trial of a series of 12 repetitions of the endpoint tokens of the cooperating-cues continuum. The subject had to meet the criterion of identifying 10 out of 12 correctly. All subjects met this criterion.

The identification task consisted of a two-alternative forced choice response to a single auditory stimulus. Eight repetitions of each of the 16 stimuli (total of the two continua) were presented in a random order consisting of eight blocks of 16 stimuli each. The stimuli were separated by an interstimulus interval of 3500 ms. Subjects could identify the stimulus by pointing to one of two pictures; a picture of a box, representing the stimulus /buk/, and a picture of a package, representing the stimulus /pak/.

The AX discrimination task required a response of "same" or "different" on each trial. In order to obtain a bias-free measure of discriminability, the tasks were set up in such a way that signal detection measures could be applied (Coombs, Dawes, & Tversky, 1970). For this, each task contained physically different as well as identical pairs. There were two separate discrimination tasks, one for the cooperating-cues and one for the conflicting-cues continuum. In both tasks the subjects heard two series of 27 discrimination pairs. Each series contained two repetitions of the physically identical pairs 2-2, 3-3, 4-4, 5-5, 6-6, 7-7, and three repetitions of the physically different pairs consisting of stimulus 2 (/buk/), the so-called "anchor" stimulus for which the JND was being measured; this resulted in pairs 2-3, 2-4, 2-5, 2-6, and 2-7. The anchor stimulus was always in first position in the pair. All pairs in one series were randomly ordered with an intrapair interval of 400 ms and an interpair interval of 3500 ms.

The subjects were asked to point to a picture containing a triangle and a circle when the words in the pair they heard sounded different and simply not to respond when the words in the pair they heard sounded the same. Half of the subjects started with the stimuli from the cooperating-cues continuum, and half of the subjects started with the stimuli from the conflicting-cues continuum. The children were motivated to respond by randomly verbally reinforcing responses throughout the experiment. In addition the subjects knew they were to receive a small present for cooperation after finishing the tasks. Subjects never received differential feedback for particular responses.

All subjects first performed the identification task with four series of 16 stimuli and then one of the discrimination tasks. After a short break, the subjects performed the identification task with the remaining four series of 16 stimuli and the other discrimination task.

Results

Identification

Figures 3 and 4 display the mean identification curves for the four groups in the cooperating-cues and the conflicting-cues conditions, respectively. Each individual identification curve was submitted to probit transformations (Finney, 1971), yielding slope values and phoneme boundary values. A high slope value indicates a small uncertainty range and suggests a highly consistent ability to categorize a speech contrast, whereas a low slope value indicates a large range and suggests difficulty in the identification of a speech contrast. Table 4 shows the mean phoneme boundary and slope scores for the four groups. Figure 5 displays the...
individual slope scores for the subjects in the cooperating-cues and the conflicting-cues conditions.

For both the phoneme boundary scores and the slope scores a three-factor analysis of variance was used to test for significant differences. The three factors were OME (levels: presence versus absence), Language Impairment (levels: presence versus absence), and Stimulus Type (levels: cooperating versus conflicting cues), with the levels of Stimulus Type treated as repeated measures. Hence, in the factor OME, the subject groups OME/LI and OME represented the presence of early OME, whereas the subject groups LI and Control represented the absence of early OME. In the factor Language Impairment, the subject groups OME/LI and LI represented the presence of language impairment, whereas the subject groups OME and Control represented the absence of language impairment. There was a significant mean slope difference between stimulus types [Stimulus Type: F(1, 40) = 21.9, p < .001], which indicates that stimuli in the cooperating cues continuum were less ambiguously perceived when compared to the stimuli in the conflicting cues continuum. In addition, there were significant slope effects of OME, Language Impairment, and the interaction term OME × Language Impairment [F(1, 40) = 6.20, p = .017; F(1, 40) = 6.77, p = .013; F(1, 40) = 6.06, p = .018, respectively]. A post hoc Tukey (HSD) range test (p = .05) showed that the control group had significantly higher slope scores than either experimental group in both the cooperating-cues and the conflicting-cues conditions. The three experimental groups were not significantly different from each other. The results and the data of Table 4 indicate that children with either early OME experience or language impairment identified the speech tokens less consistently. Thus, a severe history of OME appears to result in poorer phonetic processing, irrespective of language impairment. However, a combination of early OME experience and language impairment did not further increase the perceptual problems as indicated by the significant interaction between OME and Language Impairment, which suggests nonadditivity for the levels of phonetic processing associated with OME and Language Impairment.

Finally, there was no significant interaction between OME and Stimulus Type, Language Impairment and Stimulus Type, and OME and Language Impairment and Stimulus Type [F(1, 40) = 1.42, p = .241; F(1, 40) = 0.92, p = .342; F(1, 40) = 1.18, p = .283, respectively]. Thus, the processing of major and minor cues was largely similar for the control group and the three experimental groups.

With regard to the mean phoneme boundaries, there was a significant main effect of Language Impairment [F(1, 40) = 4.76, p = .035]. The mean phoneme boundary of children with language impairments was shifted to the right, which indicates a higher number of /b/ responses. In addition, there was a significant main effect of Stimulus Type [F(1, 40) = 4.54, p = .039], which was the result of the subjects’ tendency to perceive the voiced sounds from the conflicting cues continuum less ambiguously than the voiceless sounds in the conflicting cues continuum.

**Discrimination**

Discrimination results for each pair was expressed with the nonparametric estimate of d', yielding −ln eta scores (discriminability) and ln beta scores (response bias) (Wood, 1976). The −ln eta results, as a function of stimulus pair, are shown in Figures 6 and 7 for the cooperating-cues and the conflicting-cues continua, respectively. Discriminability (−ln eta) equals zero when performance is at chance. It increases with greater accuracy of discrimination, without influences of bias to respond “same” or “different.” Discriminability is maximal at the value of −ln eta of 4.6. This 4.6 value is obtained when the probabilities of correct “different” and correct “same” responses are both .99, which was the value assigned (for computational purposes) when the actual probabilities were 1.00.

From the overall level of the curve, it seems that the OME/LI group shows the poorest sensitivity of any of the groups in both Figure 6 and Figure 7. Especially in the conflicting-cues condition, this group has the shallowest discrimination function. The Control group shows the steepest discrimination function in both the cooperating- and the conflicting-cues condition. This indicates highest sensitivity of all groups. The overall discrimination levels and the
A 3-way ANOVA (OME × Language Impairment × Stimulus Type) was performed on ln beta scores (response bias). This did not result in significant differences for either OME \( F(1, 40) = 0.55, p = .462 \), Language Impairment \( F(1, 40) = 0.23, p = .632 \), or OME × Language Impairment \( F(1, 40) = 1.11, p = .299 \). Hence, there were no differences in tendencies to favor one response over the other (independent of stimulus discriminability) between the groups.

We focused analyses on JND measures of sensitivity.\(^1\) Linear regression analyses were performed on the individual discrimination functions. JNDS could be determined by computing the interpair difference that provided a discriminability of 50% of the maximum discriminability value (i.e., \(-\ln \eta = 2.3\)). Table 5 presents the mean JNDS for the four groups.

A 3-way ANOVA (OME × Language Impairment × Stimulus Type) of JNDS, with Stimulus Type as a repeated measure, resulted in significant effects of OME, Language Impairment, and Stimulus Type \( F(1, 40) = 4.39, p = .042; F(1, 40) = 13.83, p < .001; F(1, 40) = 55.51, p < .001 \), respectively. Children with early OME experience or language impairment required a greater auditory difference between two stimuli in order to differentiate between them. The significant Stimulus Type effect indicates that the cooperating-cues stimuli resulted in smaller JNDS than the conflicting-cues stimuli.

Identification results from the current investigation, specifically slope results, suggest that children with early OME or language impairment have difficulties with categorization. The significant interaction between OME and Language Impairment demonstrates that coexistence of language impairment does not further deteriorate phonemic perception. Further, discrimination results demonstrate that children with early OME, irrespective of language impairment, also have poorer sensitivity to voicing cues than children without early OME. Children with early OME or language impairment appear to need more redundancy of auditory information. From the insignificant interaction terms OME × Stimulus Type and OME × Language Impairment × Stimulus Type on the JNDS, it can be deduced that the lower sensitivity is an overall effect and not particularly related to one or a few specific voicing cues.

Our results add new information to the results of Eimas and Clarkson (1986) and Clarkson et al. (1989), who also used identification and discrimination tasks in assessing the

---

\(^{1}\)We were interested in the sensitivity to voicing cues. JND measures sensitivity as a function of increasing physical differences between pairs. As such, it reflects the smallest acoustical difference that can be heard. In many studies on speech perception, overall discriminability is used as an index of discrimination. Overall discrimination is an average of the discriminability across all pairs (both within and across phoneme pairs). As a consequence, it is a less precise measure of sensitivity. In addition, using overall discriminability scores is most appropriate when working with a fixed-step AX discrimination task. We used an anchor procedure with variable stepizes between pairs. A JND procedure is most appropriate in an anchor discrimination task. Therefore, our statistical analyses concern JND data only.

---

**TABLE 5. Mean discrimination results for the four groups.**

<table>
<thead>
<tr>
<th>JND Type</th>
<th>OME/LI</th>
<th>OME</th>
<th>LI</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperating</td>
<td>4.73</td>
<td>4.02</td>
<td>4.22</td>
<td>3.33</td>
</tr>
<tr>
<td>Conflicting</td>
<td>7.63</td>
<td>5.68</td>
<td>6.04</td>
<td>5.12</td>
</tr>
</tbody>
</table>

Note. OME = history of Otitis Media with Effusion; LI = Language Impairment; JND = Just Noticeable Difference.
long-term effects of OME. However, they did not separate the contribution of language impairment from the contribution of early OME in a complete factorial design.

Eimas and Clarkson (1986) reported significant differences in overall discriminability (i.e., averaged across pairs) that were due to OME. In the current investigation, discrimination was assessed using JNDs instead of a fixed interval n-step AX discrimination task. From the correspondence across studies it can be deduced that the current discrimination procedure using the JND measure provided a sensitive means for assessing auditory processing of subtle acoustic differences.

However, unlike the present study, Eimas and Clarkson (1986) did not find that significant differences in identification were due only to OME. We found significant differences in phonetic identification that was due to both OME and language impairment. It is likely that the use of both major and minor cues enhanced the sensitivity of the identification task to subject differences.

In addition, there are differences in design between the study of Clarkson et al. (1989) and our own. Clarkson et al. studied 5-year-old children; we studied 9-year-olds. Thus, maturational changes in the 4-year interval could have affected results differentially. We do partly agree with Eimas and Clarkson’s (1986) interpretation of the results, however. That is, recurrent OME may be considered a form of early sensory deprivation, and information relevant to the perception and categorization of speech may be less consistent during the years when the child should be acquiring the sound system of the native language. This inconsistency may aggravate discovery of how the specific structures of relevant acoustic information are mapped onto the sound categories of the language. Furthermore, earlier OME episodes may be related to the currently observed poorer sensitivity of children with OME to voicing cues. Differing sensitivity may underlie their less-consistent phonetic identification abilities.

The manipulation of conflicting and cooperating cues in the current study provided a means for studying potential differences in weighting of the major and minor cues to voicing in Dutch. As shown in Table 4, all groups showed a higher phoneme boundary in the conflicting-cues condition as compared to the cooperating-cues condition. If the experimental groups attributed more weight to the minor cues, then one would expect their shift in phoneme boundary to be larger than that of the control group. Although there was a tendency for the OME/LI group to make such a larger shift, it was not significant. We, therefore, conclude that the perceptual weighting of major and minor cues in the experimental groups is similar to the weighting in the control group.

Our factorial design provided the opportunity to study the effects of OME, irrespective of language impairment. Early OME is often accompanied by language impairment. An important question is whether language impairment can further interfere with auditory and phonetic processing when it coexists with early OME. Our results indicate that either early OME or language impairment is related to perceptual problems. A combination of the two, however, did not make speech perception significantly worse, although children with OME and language impairments did have the most shallow slopes in identification, the most shifted phoneme boundaries, and the poorest discrimination abilities of any group. Of course, the nonadditivity of the effects of OME and language impairment may be the result of a floor effect. Thus, the nonadditivity of perceptual problems in cases of co-existence of a history of otologic problems and language impairment certainly is an interesting subject for future research.

Most psycholinguistic models of speech perception assume both auditory and phonetic levels in processing of speech. In a hierarchical dual-coding strategy (e.g., fitting the dual process model of Fujisaki & Kawashima, 1969, 1970), both processing stages show interdependency. Only in cases of insufficient information for phonemic decisions is acoustic information in memory consulted (see Macmillan, 1987). Classical dual-coding models of speech perception do not provide a rationale for independence of processing stages. Earlier studies (Groenen, Maassen, & Crul, 1994; Groenen, Maassen, Crul, & Hulsmans, 1994) suggest that discrimination and identification tap different perceptual processes. Auditory processing can be affected, whereas phonetic processing is intact. Although this was not the case in the current investigation, our results do fit a nonhierarchical structure of speech processing, involving an auditory stage and a phonetic stage partly allowing for stage-independent output, without the integrity of phonetic processing being dependent on the outcome of auditory processing.

In Groenen, Thoonen, Maassen, and Crul (1995) it was suggested that reduction of the redundancy in speech stimuli may increase their diagnostic value only when the reduction pertains to linguistically relevant dimensions. One of the criteria used in selecting the subjects for the present study was normal speech-in-noise recognition. Speech-in-noise tasks typically aim at assessing central perception. The four groups employed in this study did not differ in their speech-in-noise recognition abilities, whereas they did differ in our identification and discrimination tasks. This suggests that degradation of the stimuli with nonlinguistic information (e.g., adding noise) may have less value for assessing psycholinguistic difficulties because of its marginal relationship to the speech signal and the specific processes of speech perception.

Research on the developmental sequelae of OME has been equivocal. The complex long-term effects of OME demand detailed attention to the different psycholinguistic levels of speech perception. The present study demonstrates that children with early OME experience (but normal hearing at present) show poorer sensitivity to voicing cues and less distinctive phonetic categorization, similar to children with language impairments (e.g., Thibodeau & Sussman, 1979). Disturbances in lower-level perception may have diverse effects on higher-order language learning processes. We suggest that these disturbances in lower-level perception processes form the basis for higher-order linguistic problems in some children and thereby form an intermediate between OME and the diversity of outcome in language learning.
Acknowledgments

The Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged for funding this project. This research was conducted while Paul Groenen was supported by a PSYCHON- grant of this organization (560-256-047), awarded to Dr. Ben Maas- sen and Dr. Thom Cruil. We would also like to thank Esmeralda Brunia for her cooperation in testing the children, Eefje Grevink and Sylvia Peters for providing the data concerning the language skills of the children, and Anne Schilder for providing the otologic data. Finally, we want to thank all the children and teachers of over 35 schools in the region of Nijmegen who participated in this study.

References


Appendix

Factor Analyses on Questionnaires

Principal factor analysis was used on each of two questionnaires (a questionnaire filled out by the children's parents and a questionnaire filled out by the children's teachers). This was done in order to obtain factor scores for linguistic competence. Each questionnaire consisted of 36 items concerning speech, language, and communicative behavior. All items had three response alternatives. The exact contents of both questionnaires can be found in Grievink, Peters, van Bon, and Schilder (1993).

Principal factor analysis on the responses to the parent questionnaires yielded two orthogonal factors with considerable eigenvalues (greater than 2.0). We used a factor loading criterion of .65 to select relevant items. Factor 1 and factor 2 consisted of 11 and 0 items, respectively, with loadings above .65. Only the first factor was therefore used. The 11 items selected for factor 1 were best described in terms of general linguistic competence. To test for unidimensionality of the first factor, we repeated principal factor analysis on the 11 selected items. This resulted in a first factor with an eigenvalue above 5.5. The remaining factors had eigenvalues below 1.0, which indicated unidimensionality for the selected items.

For each child, factor scores were computed on the basis of the 11 items. Two examples of the 11 selected items are (a) "Has a faulty pronunciation in sentences," and (b) "Understands only simple sentences."

The same procedure was followed with the questionnaire filled out by the teachers. Principal factor analysis yielded two orthogonal factors with eigenvalues greater than 2.0. Factor 1 and factor 2 consisted of 16 and 3 items, respectively, with loadings above .65. Only the first factor was used. As was the case with the questionnaire filled out by the parents, the 16 items selected for factor 1 were best described in terms of general linguistic competence. A repeated-principal factor analysis on the selected 16 items resulted in a first factor with an eigenvalue above 9.0, whereas the remaining factors had eigenvalues below 1.0. This indicates unidimensionality of the selected items. Factor scores were computed on the basis of these 16 items. Two examples of the 16 selected items are (a) "Uses few different words," and (b) "Merely understands sentences consisting of very common and frequently heard words."