Contribution of a Contralateral Hearing Aid to Perception of Consonant Voicing, Intonation, and Emotional State in Adult Cochlear Implantees

Tova Most*,1, Gal Gaon-Sivan1, Talma Shpak2, Michal Luntz2
1Tel-Aviv University
2Cochlear Implant Program, Department of Otolaryngology, Bnai Zion Medical Center

Received July 20, 2011; revisions received September 21, 2011; accepted September 26, 2011

Binaural hearing in cochlear implant (CI) users can be achieved either by bilateral implantation or bimodally with a contralateral hearing aid (HA). Binaural–bimodal hearing has the advantage of complementing the high-frequency electric information from the CI by low-frequency acoustic information from the HA. We examined the contribution of a contralateral HA in 25 adult implantees to their perception of fundamental frequency-cued speech characteristics (initial consonant voicing, intonation, and emotions). Testing with CI alone, HA alone, and bimodal hearing showed that all three characteristics were best perceived under the bimodal condition. Significant differences were recorded between bimodal and HA conditions in the initial voicing test, between bimodal and CI conditions in the intonation test, and between both bimodal and CI conditions and between bimodal and HA conditions in the emotion-in-speech test. These findings confirmed that such binaural–bimodal hearing enhances perception of these speech characteristics and suggest that implantees with residual hearing in the contralateral ear may benefit from a HA in that ear.

Broadening of the criteria for cochlear implantation has led to increasing numbers of unilateral cochlear implant (CI) users with a significant amount of residual hearing in the nonimplanted ear. Many of these CI users continue to use a hearing aid (HA) in that ear. Thus, using a CI in one ear and a contralateral HA, the individual is exposed to binaural–bimodal hearing (CI + HA).

A number of studies have documented the advantages of bimodal hearing over monaural CI. Some of the advantages are attributed to the fact that the hearing is binaural, which allows better speech perception in quiet and in noise, better localization, better sound quality, and better awareness of sounds in the environment (Ching, Psarros, Hill, Dillon, & Incerti, 2001; Ching, Van Wanrooy, Hill, & Incerti, 2006; Chmiel, Clark, Jerger, Jenkins, & Freeman, 1995; Hamzavi, Pok, Gstoettner, & Baumgartner 2004; Luntz et al., 2003; Luntz, Shpak, & Weiss, 2005; Luntz, Yehudai, & Shpak, 2007; Mok, Grayden, Dowell, & Lawrence, 2006; Morera et al., 2005; Simons-McCandless & Shelton, 2000). Others emphasize the complementary information that bimodal hearing provides. Thus, CI users benefit from an increased audibility of high-frequency information of the speech signal compared to those using HA only (Zeng, 2002), but their detection of low-frequency sounds is poor (Dillier, 2004). In the electrical hearing provided by the CI, the pitch-encoding mechanisms responsible for low-frequency information are not appropriately represented, either by the temporal fine structure of the neural firing pattern or by the place of stimulation (Carroll & Zeng, 2007; Zeng, 2002, 2004). The HA, by contrast, provides better audibility in the low-frequency range than at higher frequencies. Thus, the unique benefit from binaural–bimodal stimulation arises from complementing of the low-frequency acoustic information obtained from the HA with the high-frequency information derived from the electrical stimulation of the CI (Chang, Bai, & Zeng, 2006; Kong, Stickney, & Zeng, 2005; Qin & Oxenham,
2006). Some researchers claim that in most bimodal users the CI provides the major part of the individual’s speech recognition ability while the contralateral HA acts as an adjuvant by providing salient pitch information at low frequencies (Kong et al., 2005; Zeng, 2002, 2004).

Perception of Segmental Speech Characteristics

The advantage of bimodal hearing over CI alone has been demonstrated in the perception of segmental, suprasegmental, and paralinguistic speech characteristics (Kong et al., 2005; Most, Harel, Shpak, & Luntz, 2011; Quadrizius, 2009; Straatman, Rietveld, Beijen, Mylanus & Mens, 2010). In studies of perception of the segmental features of speech in both adults and children, CI users showed advantages over HA users with similar hearing loss (Blamey et al., 2001; Boothroyd & Eran, 1990). However, when the perception of segmental features by CI alone was compared to perception by CI + HA, the bimodal hearing users demonstrated better identification of consonant voicing and manner of articulation (Ching et al., 2001, 2006; Incerti, 2003; Mok et al., 2006; Mok, Galvin, Dowell, & McKay 2007). A study in which bimodal hearing was simulated in adults with normal hearing (Lou & Fu, 2006) showed that the addition of low frequencies (below 500 Hz) to the electrical stimulation in the contralateral ear improved the perception of voicing in Chinese. Zahng, Dorman, and Spahr (2010), in a study of the contribution of acoustic low-frequency information to electrical hearing in the perception of monosyllabic words in quiet by postlingually deafened adults, demonstrated a significant improvement in word recognition when the acoustic information was added. Even when the acoustic information contained frequencies up to 125 Hz only, the bimodal condition showed hearing improvement of 22% over CI alone. Information transmitted up to 125 Hz contained information on the fundamental frequency (F0), some information on the first formant and changes in the amplitude through the signal envelope. The authors claimed that the bimodal contribution was a result of information on consonant voicing and manner of articulation. Ching et al. (2001) and Ching (2005) reported that the contribution of the bimodal condition to the perception of voicing was greater in the presence of noise (10SNR) than in quiet.

Perception of Suprasegmental Speech Characteristics

With regard to suprasegmental features (intonation, word emphasis, syllable stress), comparisons between the perception of suprasegmental features by individuals with HA and by individuals with CI have not necessarily favored the latter (Carney, Kienle, & Miyamoto, 1990; Most & Peled, 2007; Peng, Tumblin, & Turner, 2008; Waltzman & Hochberg, 1990; Wu & Yang, 2003). For example, Waltzman and Hochberg (1990) found that not only children who were using a CI (Nucleus 22) but also children with profound hearing loss who were using a HA performed well in perceiving word emphasis and pitch changes. Boothroyd and Eran (1990) reported that perception of syllable number by children with a CI (Nucleus 22) did not differ significantly from that of children with a HA but that the latter group performed better at perceiving intonation. In addition, Most and Peled (2007) reported poorer perception of syllable stress and intonation by children who used a CI than by children with severe or profound hearing loss who used a HA. One suggested explanation for the poorer performance with a CI is that the implant may not provide sufficient information on the changes in F0, which provide an important acoustic cue for the perception of suprasegmental features. For example, O’Halpin, Falkoner, Rosen, and Viani (2006), using a methodology based on synthetic speech stimuli with controlled changes for each of the acoustic parameters (amplitude, duration and F0), suggested that perception of word emphasis in children implanted with the Nucleus 24 (with SPEAK and ACE) does not rely on changes in F0 but rather on duration and intensity cues.

It thus seems that in many of the currently available CI systems the place and the temporal pitch-encoding mechanisms are inadequate, resulting in difficulties in the transfer of information needed for the perception of some suprasegmental features of speech (Carroll & Zeng, 2007). That the information provided by electrical information alone is insufficient was demonstrated in a study in which electrical hearing only and electrical hearing plus low-frequency acoustic hearing were
simulated in young (20- to 37-year old) and older (66- to 82-year old) listeners with normal hearing (Souza, Arheart, Wise Miller, & Muralimanohar, 2011). In all listeners, perception of intonation was better in the bimodal condition than with electrical information only.

In some previous studies, bimodal hearing was found to have an advantage over CI alone for the perception of suprasegmental features (Landwehr, Pyschny, Walger, Von Wedel, & Meister, 2007; Most et al., 2011). For example, when comparing the perception of intonation, word emphasis, and syllable stress by users of CI alone and by experienced bimodal users, Most et al. (2011) demonstrated an advantage of the bimodal mode for all three features.

Perception of Paralinguistic Information

In addition to the segmental and suprasegmental features of speech, paralinguistic information—such as the emotional state of the speaker—is an important aspect of the spoken language communication. Difficulties in perceiving information about the speaker's emotional state may lead to lack of awareness of the individual's impact upon others, lack of empathy, and failure to adapt social skills to specific situations (Mellon, 2000). As with the suprasegmental features, auditory perception of the emotional state is cued through changes in F0 along the utterance, as well as through duration and intensity cues (Banse & Schere, 1996). A comparison of the perception of emotions by adolescents with severe-to-profound hearing loss who were using HAs and by normal-hearing controls revealed that the hearing-impaired group performed significantly worse than the normal-hearing participants, not only through the auditory mode but also through the visual and the combined auditory–visual modes (Most, Weisel, & Zaychik, 1993). Moreover, auditory–visual perception by the hearing-impaired participants was not better than their visual perception, while the normal-hearing participants demonstrated significantly better auditory–visual perception than visual perception alone.

Despite the importance of the perception of emotional content for successful communication, only a few research groups have examined its performance by CI users. Comparisons of auditory perception in postlingually deafened adult implantees and in hearing individuals (Luo, Fu, & Galvin, 2007; Pereira, 2000; Peters, 2006) have revealed that implantees perform worse than hearing controls and show wide variations among participants. In addition, after the stimuli were normalized (via control of intensity cues of the various emotions), the performance of the implantees worsened further, suggesting that they had probably relied on their perception of intensity cues, and once these cues were controlled their performance deteriorated. No such effect of normalization was observed in the performance of the hearing controls (Luo et al., 2007; Pereira, 2000). A comparison of the auditory, visual, and auditory–visual perceptions of emotions by children and adolescents (age range 10–17 years) who used CI, by HA users, and by normal-hearing individuals showed that the auditory and the auditory–visual results of the normal-hearing participants were better than those of the hearing-impaired participants, with no differences reported between CI and HA users (Most & Aviner, 2009).

Cullington and Zeng (2011) compared perception of pitch as well as emotional prosody in participants with bilateral CI and in bimodal (CI + HA) users. The bimodal users achieved higher scores than the bilateral implantees on most tasks, although performance in the two groups did not differ significantly. Interestingly, however, bimodal users achieved similar scores to normal-hearing listeners on the emotional prosody perception test while the scores of the bilateral implantees on this test were significantly lower than those of the normal-hearing listeners.

Quadrizius (2009) assessed the perception of emotions by a 9-year-old girl under three conditions: CI only, CI + HA in the same ear, and unilateral CI with contralateral HA. In both conditions in which the HA was used, the perception of emotions was better than with the CI alone.

In this article, we will further examine the auditory perception of emotions by bimodal users, in an attempt to determine whether the addition of the contralateral HA contributes to the user's perception of the speaker's emotional state.

In summary, the purpose of this research was to evaluate the contribution of a contralateral HA in adult CI users to their perception of speech...
characteristics, namely initial consonant voicing, intonation, and emotional content, which are cued to the listener mainly by information transmitted through the F0. Perception of these three speech characteristics was examined here in participants with HA alone, with CI alone, and in the bimodal condition (CI + HA). We postulated that the perception of these characteristics would be better under the bimodal condition than with each of the sensory aids alone. Our expectation was that the use of tests that specifically examine the complementary bimodal benefit would provide evidence indicating a gain from the use of the contralateral HA by CI implantees.

Methods

Participants

Included in the study were 25 adults (9 males and 16 females) whose ages ranged from 15 to 67 years (mean ± SD, 39.6 ± 17.0 years). All participants were experienced bimodal users. All were native speakers of Hebrew and used spoken language as their mode of communication. Hearing loss (HL) was prelingual in 19 participants and postlingual in 6. The participants used various kinds of CIs (see Table A1 for details).

Age at implantation ranged from 10 to 66 years (mean ± SD, 35.5 ± 17.0 years). Experience with bimodal hearing ranged from 11 months to 10 years (mean ± SD, 4.0 ± 28.9 years). Digital HAs were used by 23 participants and analogic HAs by 2. All participants had used a HA in the nonimplanted ear prior to implantation. All used their HAs for at least 75% of waking time regardless of the amount of residual hearing in the nonimplanted ear, and all had well-established and stable CI maps.

Residual hearing in the nonimplanted ears was measured with a GSI-16 audiometer. The unaided pure-tone average (PTA) at 500 Hz, 1000 Hz, and 2000 Hz) in the nonimplanted ears ranged from 66.66 to 116.66 dB HL (mean ± SD, 94.63 ± 11.50 dB HL) and the aided PTA from 25 to 61.66 dB HL (mean ± SD, 42.08 ± 11.33 dB HL). At 250 Hz, the unaided pure-tone threshold in the nonimplanted ear ranged from 55 to 100 dB HL (mean ± SD, 70.86 ± 9.96 dB HL) and the aided threshold from 20 to 70 dB HL (mean ± SD, 37.29 ± 13.51 dB HL).

Demographic details of the participants are presented in Table A1.

Stimulus Materials

Three speech perception tests were used. Initial consonant voicing was used to assess the perception of a segmental feature, intonation for the perception of a suprasegmental feature, and emotional content for the perception of a paralinguistic feature. A closed set paradigm was used for all tests.

Perception of initial consonant voicing was assessed by the use of the relevant subtest in the second version of the Hebrew Speech Pattern Contrasts (HESPAC) (Rich, Wisper, & Most, 2009). This subtest incorporates 12 items. Each item includes a minimal pair comprised of meaningful CVC (consonant–vowel–consonant) words differing in voicing of the initial consonants (e.g., kir [wall] vs. gir [chalk]). The initial consonants consisted of voiced and voiceless plosives and fricatives accompanied by different vowels. Perception of voicing was assessed by 24 items (2 lists of 12 items each). The lists of words were recorded by a young woman with normal voice and speech articulation using Sony Sound Forge 7 software in an IAC sound booth. The intensity levels of all stimuli were normalized. The recordings were tested by presentation to normal-hearing young adults and were found to be valid.

Perception of intonation (statement vs. yes/no question) was assessed by the use of the relevant subtest in the first version of the HESPAC (Kishon-Rabin, Eran, & Boothroyd, 1990). This subtest incorporates 12 items. Each item comprises two identical sentences differing in their intonation contours: a statement with a falling intonation contour, for example, cham po (it’s hot here), and a yes/no question with a rising intonation contour, for example, cham po? (is it hot here?). This feature was assessed by 24 items. The sentences were recorded, intensity levels of all stimuli were normalized, and recordings were validated as described above for perception of voicing.

The 24 items were subjected to acoustic analysis with Praat software. The parameters analyzed were F0 (average, maximum, minimum, F0 at the beginning and at the end of the utterance), amplitude (average,
Perception of emotional content was assessed by means of an emotion identification test (Most & Michaelis, 2011). This test was previously presented to young children with various degrees of hearing loss. It incorporates 24 items, comprising six presentations of each of the following emotions: anger, fear, sadness, and happiness. Each emotion is expressed through the use of the same nonsense words: “bado minu gana.” Thus, the verbal content remains the same throughout the test and differences in emotional content are expressed by nonverbal auditory cues. The emotions were produced by a professional actress using normal voice and speech articulation. Recording was done in an IAC booth using an Audio-Technica AT-892 head microphone connected to a computer with Line 6 Tone Port and an external voice card. The use of a head microphone made it possible to maintain a constant distance between the actress’s mouth and the microphone. The items were tested by presentation to young adults with normal hearing and were found to be valid.

The 24 items were subjected to acoustic analysis with Matlab and Praat softwares. The parameters analyzed were pitch (average, maximum, minimum, range), amplitude (average, maximum, minimum, range), duration (total utterance), and voice quality (jitter, shimmer, noise-to-harmonic ratio, and autocorrelation). Detailed information appears in Most and Michaelis (2011).

Procedure

The participants were recruited from an implant center and through an organization of individuals with hearing loss. All signed their informed consent to participate. The examiner met with each participant individually in a quiet room for a session that lasted about an hour. During that time, the three tests were presented on a portable computer that was connected to two loudspeakers located on either side of the computer at an angle of 45 degrees from the participant’s seat. The computer screen was turned to face the examiner. Stimuli were presented at conversation level (72 dB SPL). The intensity level was measured with a TES-1350A Sound Level Meter (TES Electrical Electronic Corp). The participants were asked whether that level was comfortable, and it was adjusted accordingly if not. Participants used their own HA and CI volume and sensitivity settings during the entire session.

The three tests were presented to each participant under three conditions: (a) using only the CI, (b) using only the HA, and (c) using both CI and HA. The order of presentation and the order of tests within each condition were randomized. After each condition, there was a break, of about 10 min, during which time the participant switched to the next condition (involving a change in the aid to be used), and the examiner chatted with the participant in order to allow time for adjustment to the next condition. In general, the same level of presentation was used across the three conditions; in only a few cases, a participant requested adjustment to a more comfortable level. Participants listened to the stimuli and were instructed to mark what they heard on a response form. Prior to each test, two items that were not included in the test were presented for practice. At the end of the session, the participant completed a questionnaire on demographic details including age, age at onset of hearing loss, age at implantation, and types of CI and HA. Audiograms including details of each participant’s aided and unaided degree of hearing loss in the non-implanted ear were collected as well.

Results

Each participant was awarded three perception scores (initial consonant voicing, intonation, and emotion) for each of the three presentation conditions: HA alone, CI alone, and CI + HA. Because we used closed-set materials, we corrected the scores for guessing according to the formula of Boothroyd (1988), which accounts for the number of possible alternatives as follows:

\[ Sc = \frac{(Su - Sg)}{(100 - Sg)} \times 100, \]

where \( Sc \) is the corrected score expressed as a percentage, \( Su \) is the uncorrected score expressed as a percentage, and \( Sg \) is the mean score expected
Table 1  Corrected mean percent scores and SDs for each of the tests (initial consonant voicing, intonation, and emotion) in each of the three conditions: HA alone, CI alone, and CI + HA

<table>
<thead>
<tr>
<th></th>
<th>HA</th>
<th>CI</th>
<th>CI + HA (bimodal)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (%)</td>
<td>SD</td>
<td>Mean (%)</td>
</tr>
<tr>
<td>Initial consonant voicing</td>
<td>77.33</td>
<td>27.48</td>
<td>90.00</td>
</tr>
<tr>
<td>Intonation</td>
<td>72.66</td>
<td>27.79</td>
<td>62.00</td>
</tr>
<tr>
<td>Emotion</td>
<td>34.44</td>
<td>24.32</td>
<td>33.55</td>
</tr>
</tbody>
</table>

Note. HA, hearing aid; CI, cochlear implant.

from guessing (allowing 25% for emotion and 50% each for intonation and initial voicing). Note that after correction, a score equal to that expected from guessing becomes 0%, while a perfect score remains at 100%.

Table 1 presents the corrected mean percentage scores and their standard deviations for each test in each of the three conditions.

For each test, ANOVA with repeated measures was conducted (using a mixed model). In the ANOVA for initial consonant voicing, the independent variable was the level of voicing (voiced/voiceless), and the repeated measure was the condition (HA alone, CI alone, CI + HA). The dependent variable was the correct perception. The analysis revealed a significant effect of condition, \(F(2,23) = 5.70, p = .002\). There was no significant main effect of voicing. Thus, there were no differences between the correct scores of the voiced and the voiceless items. There was no significant interaction between condition and level of voicing. Multiple comparisons using the studentized maximum modulus adjustment (Hochberg, 1974) were conducted among the scores for the three conditions. These comparisons revealed a significant difference between the bimodal condition and the CI alone condition \(p = .009\). There were no significant differences between the HA alone and the CI alone conditions or between the HA alone and the bimodal conditions.

In the ANOVA for the intonation test, the independent variable was the emotion (anger, fear, sadness, and happiness) and the repeated measure was the condition (HA alone, CI alone, CI + HA). The dependent variable was the correct perception. The analysis revealed a significant main effect of emotion, \(F(3, 264) = 9.99, p < .0001\), and a significant effect of condition, \(F(2, 264) = 7.78, p = .0005\). There was no significant interaction between condition and emotion. Multiple comparisons using the studentized maximum modulus adjustment (Hochberg, 1974) were conducted among the scores for the three conditions. These comparisons revealed significant differences between the CI alone and the bimodal conditions \(p = .001\) and between the HA alone and the bimodal conditions \(p = .002\). There was no significant difference between the HA alone and the CI alone conditions.

Multiple comparisons among the emotions revealed significant differences between happiness and sadness \(p = .0003\), between anger and fear \(p = .04\), and between fear and sadness \(p < .0001\). There were no significant differences between anger and happiness, between anger and sadness, or between happiness and fear.

Table 2 presents the mean scores and their standard deviations for each of the emotions in the three conditions. ANOVA with repeated measures was conducted for each emotion. The repeated measure was the condition (CI alone, HA alone, CI + HA). The analysis for anger revealed significant differences.
among the conditions, $F(2, 48) = 3.25, p = .047$. Multiple comparisons revealed a significant difference between the HA and the bimodal conditions ($p = .042$). There were no significant differences between the HA and the CI conditions or between the CI and the bimodal conditions. The analysis for happiness revealed significant differences among the conditions, $F(2, 48) = 3.77, p = .030$. Multiple comparisons revealed a significant difference between the CI and bimodal conditions ($p = .026$). There were no significant differences between the CI and the HA conditions or between the HA and the bimodal conditions. The analyses for sadness and fear revealed no significant differences among the conditions.

Confusion matrices among the emotions revealed that, in general, the confusions were similar in the three conditions. Sadness was identified best and fear was identified the worst. Participants tended to show confusion between anger and happiness and to identify sadness as fear. Different confusions in the different conditions were seen only for happiness: in the CI condition, participants identified happiness as anger more than as happiness, whereas in the HA alone and the bimodal conditions participants identified happiness more successfully and confused it much less with anger. All four emotions were identified most successfully in the bimodal condition.

Individual Data

Although the group patterns were clear, descriptive information on individual participants may be of importance because of the substantial heterogeneity known to characterize this population. Descriptive data on the individual participants in the HA, CI, and CI + HA conditions are shown in Figures 1–3 for the perception of initial consonant voicing, intonation, and emotion, respectively.

The figures reflect marked variation in the scores of individual participants for each of the three tests (initial voicing, intonation, and emotions) and in each condition (HA only, CI only, and CI + HA). Some implantees appeared to derive some benefit from the addition of a contralateral HA, some derived none, and for some the CI alone was better than the bimodal condition. It should be noted, however, that the information presented in the figures is only descriptive; hence, this variability should be further examined in future research.

Pearson correlations among performances in the three tests for each condition revealed significant correlations between the intonation and the initial voicing scores in the CI condition ($r = .71, p < .0001$) and between the emotion and the intonation scores in the bimodal condition ($r = .44, p < .02$). There were no other significant correlations.

Bimodal Advantage and Residual Hearing in the Nonimplanted Ear

The bimodal advantage (the difference between the bimodal score and the CI alone score) was computed for each participant in each of the three tests. Pearson correlations were conducted between the bimodal advantage values and the aided and unaided PTA scores in the nonimplanted ear as well as between these values and each of the aided and unaided thresholds at 250, 500, 1000, and 2000 Hz. The analyses revealed a significant negative correlation between the aided threshold at 250 Hz and the bimodal advantage for intonation ($r = -.4, p < .04$),

### Table 2

Mean percent scores and SD obtained for each of the emotions (anger, happiness, sadness, fear) in the three conditions: HA alone, CI alone, and CI + HA

<table>
<thead>
<tr>
<th></th>
<th>HA (Mean %)</th>
<th>SD</th>
<th>CI (Mean %)</th>
<th>SD</th>
<th>CI + HA (bimodal) (Mean %)</th>
<th>SD</th>
<th>$F(2, 48)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>32.44</td>
<td>33.62</td>
<td>40.44</td>
<td>33.77</td>
<td>52.00</td>
<td>24.57</td>
<td>3.25*</td>
</tr>
<tr>
<td>Happiness</td>
<td>31.55</td>
<td>40.02</td>
<td>22.66</td>
<td>40.63</td>
<td>45.77</td>
<td>34.59</td>
<td>3.77*</td>
</tr>
<tr>
<td>Sadness</td>
<td>47.55</td>
<td>37.88</td>
<td>52.00</td>
<td>27.72</td>
<td>60.88</td>
<td>32.88</td>
<td>2.7</td>
</tr>
<tr>
<td>Fear</td>
<td>27.11</td>
<td>32.40</td>
<td>20.88</td>
<td>34.59</td>
<td>36.88</td>
<td>34.94</td>
<td>1.41</td>
</tr>
</tbody>
</table>

*Note: HA, hearing aid; CI, cochlear implant. *$p < .05$. 

$p = .047$. 

$F(2, 48) = 3.25, p = .047$.
that is, the lower the threshold, the greater the advantage. There was also a significant correlation between the bimodal advantage for the initial voicing and the unaided threshold at 2000 Hz ($r = .42$, $p < .04$); the higher the threshold, the greater the advantage.

Figure 1  Perception scores of the individual participants in the initial voicing test.

Figure 2  Perception scores of the individual participants’ in the intonation test.
Discussion

The purpose of the present research was to evaluate the contribution of a contralateral HA to the perception of initial consonant voicing, intonation, and emotional state of the speaker by adult cochlear implantees. We tested the hypothesis that the additional acoustic information in the low-frequency range provided by the HA yields a better perception of these speech characteristics than that obtained with the CI alone.

In general, our findings demonstrated the advantage of the bimodal condition in the identification of initial consonant voicing, intonation, and emotional state of the speaker. These results support previous findings demonstrating the benefit of the bimodal condition (CI + HA) (Ching, Van Wanrooy, & Dillon, 2007; Dorman, Gifford, Spahr, & McKarns, 2008; Fitzpatrick, Seguin, Schramm, Chenier, & Armstrong, 2009; Kong et al., 2005; Most et al., 2011).

With respect to perception of initial consonant voicing, our results demonstrated better performance in the bimodal condition than in the HA alone condition. There was no difference in performance, however, between the bimodal and the CI alone condition. These results are in contrast to previous findings obtained in English-speaking participants by Incerti (2003) in Ching et al. (2007), who reported better perception of voicing when CI users were provided with a contralateral HA than with the CI alone. It should be noted, however, that the results of voicing recognition with the CI alone in this article were high (90% on average), and therefore, the significant contribution of the HA to the CI could not be demonstrated. Future research in which the test is presented under more difficult conditions, such as in the presence of noise, might make it possible to demonstrate the benefit of the bimodal state. Our finding of high voicing identification scores with the CI alone supports previous findings in Hebrew-speaking postlingual implantees (Kishon-Rabin, Taitelbaum-Sweed, Muchnik, & Hildesheimer, 2010), who also achieved high scores in voicing perception (80% [on average]). Taken together, these findings support the notion that the CI transmits the voicing information successfully. While the F0 is an important cue in the perception of voicing (Rosen, 1992), evidently there are other temporal and intensity cues on which the listener can rely for voicing perception (Raphael, Borden, & Harris, 2007). Because intensity cues are transmitted efficiently through the CI, it is possible that our

![Figure 3](image-url)
participants used these other perceptual cues, and thus, the advantage of the bimodal condition relative to the CI alone condition was not observed. Future research should provide a larger number of voiced and voiceless stimuli representing different manners of articulation, such as plosives and fricatives, which their voicing perception relies on different acoustic cues. This larger pool of stimuli would make it possible to explore the interaction between voicing and manners of articulation and to further investigate the contributions of different perceptual cues to the perception of voicing. The present results demonstrating good perception of voicing by CI alone and thus the failure to demonstrate the advantage of the bimodal condition might also be explained by the use of different stimuli and tasks in comparison to other studies. In this research, we used minimal pairs differing by initial consonant voicing in a closed set format, whereas Ching’s study utilized an open set of words. Although the presentation of stimuli in a closed set format is more controlled, it is possible that it was less challenging and did not allow us to demonstrate the relative advantage of the bimodal condition.

Perception of intonation was found in this article to be better in the bimodal condition than with the CI alone. There was no significant difference, however, between the bimodal perception and the perception with the HA alone. The crucial perceptual cue for discriminating between a statement and a yes/no question is the change in F0 along the utterance (Rosen, 1992). It seems that acoustic information in the low-frequency range provided by the HA complemented the electrical information provided by the CI, resulting in better perception of intonation in the bimodal condition. These results support previous findings with Hebrew-speaking as well as with German-speaking bimodal users (Most et al., 2011; Straatman et al., 2010). As in previous studies, the present results demonstrated that statements with falling pitch contours were perceived better than yes/no questions, in which the contours rise (Most & Peled, 2007; Most et al., 2011).

With respect to the perception of emotions, our results demonstrated that the bimodal perception was better than the perception with each aid alone. Perception of emotions with the HA and with the CI was similar. These findings in our group of bimodal users support the results of Quadrizius (2009), who studied the perception of emotions by a single bimodal user, a 9-year-old girl. Also, in a study by Cullington and Zeng (2011), bimodal (CI + HA) participants performed similarly to normal-hearing participants in their perception of emotional prosody, whereas the scores of bilateral (CI + CI) participants were significantly lower. It thus seems that addition of the HA to the CI contributes to the perception of emotions. Further research will be needed, however, to determine whether this advantage is a result of additional acoustic information provided by the HA or a result of information derived from the binaural–bimodal mode.

The hierarchical order of emotions perceived in this test was similar in all three hearing conditions. Sadness was perceived the best and fear the worst. Similar findings have been reported previously (Most & Aviner, 2009; Most & Michaelis, 2011; Pereira, 2000). The success in identifying sadness may be explained by the fact that the intensity of this emotion was significantly lower than that of the other emotions. In other words, sadness was acoustically more salient in comparison to the other emotions that led to the ability to discriminate it from the other emotions.

In addition, emotion confusion matrices were similar in all three hearing conditions for all the emotions except happiness. In the CI condition, happiness was perceived more as anger, whereas in the HA and HA + CI conditions, it was perceived with more success and without much confusion. Acoustic analysis of the different emotions revealed significant differences between happiness and the other emotions in the pitch parameters. Thus, it is possible that in the CI condition, the pitch information was not adequately transmitted, leading to many confusions, whereas in the CI + HA and HA conditions changes in pitch were transmitted more successfully, resulting in better identification.

Examination of the individual scores in each of the tests in the three hearing conditions revealed considerable heterogeneity. Some individuals derived benefit from the bimodal condition, whereas others did not, and some participants performed better with the CI alone. We found that the benefit from the HA to the CI, that is, in the bimodal condition, was more clearly
evident in the perception of intonation and emotions than in the perception of voicing. Wide variation among individuals with respect to the benefit obtained from the bimodal condition has also been demonstrated in previous studies in which the perception of different speech aspects was examined (Luntz et al., 2005; Mok et al., 2006; Most & Aviner, 2009; Most & Peled, 2007).

The results of this article demonstrated a significant relationship between bimodal advantage and residual hearing in the nonimplanted ear. For the perception of intonation, for example, the better the residual hearing at 250 Hz, the greater the bimodal advantage. For the perception of voicing, on the other hand, the poorer the residual hearing at 2000 Hz, the greater the bimodal advantage. The correlation between bimodal advantage with better residual hearing at low frequency and the correlation between the bimodal benefit and poorer residual hearing at 2000 Hz were demonstrated previously (Callington & Zeng, 2011; Mok et al., 2006; Mok, Galvin, Dowell, & Mckay, 2010; Most et al., 2011). The acoustic information at low frequencies, which is provided by the HA, adds significantly to the information that is provided through the electrical hearing mainly at middle and higher frequencies. By contrast, information at the middle and high frequency ranges (1000, 2000, and 4000 Hz) that is transmitted via the HA may interfere with the information that is transmitted by the CI at these frequencies.

As mentioned earlier, all of our study participants were experienced bimodal users who used the HA in addition to the CI throughout the day. It is worth noting that all of them reported experiencing subjective benefit during their everyday lives from the HA in the contralateral ear. They reported that although they mostly relied on the CI for communication, the HA enhanced the sound by adding a natural quality to it. These impressions support previous reports (Ching, Incerti, & Hill, 2004; Ching et al., 2007; Fitzpatrick et al., 2009; Hamzavi et al., 2004).

In conclusion, the present findings support previous research findings demonstrating the benefit from bimodal fitting (Beijen, Mylanus, Leeuw, & Snik, 2008; Ching et al., 2006; Dorman et al., 2008; Lou and Fu, 2006; Most et al., 2011; Quadrizius, 2009; Straatman et al., 2010). Taken together, all these findings suggest that professionals should encourage implantees with residual hearing in the contralateral ear to continue using their HA in that ear. Our results further suggest that the use of task-specific tests to assess the complementary bimodal benefit can provide evidence-based clinical information on the benefit that bimodal users can expect to gain from the contralateral HA and may lead to a more reliable decision in terms of supportive contralateral implantation.

Conflicts of Interest
No conflicts of interest were reported.

References


Quadrizius, S. (2009). Effects of combined electric and acoustic hearing on speech perception of a pediatric cochlear implant user (PhD dissertation). Washington University School of Medicine, Washington, DC.


Zahng, T., Dorman, M. E., & Spahr, A. J. (2010). Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. Ear and Hearing, 31, 63–69. doi: 0196/0202/10/3101-0063/0.


## Table A1  Demographic details of the participants

<table>
<thead>
<tr>
<th>No.</th>
<th>Age</th>
<th>Pre/post</th>
<th>CI type</th>
<th>Processor type</th>
<th>Duration with CI (years)</th>
<th>Duration with HA (years)</th>
<th>Thresholds with HA (dB HL) in the ear contralateral to CI</th>
<th>Thresholds without HA (dB HL) in the ear contralateral to CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>Post</td>
<td>Nucleus Freedom</td>
<td>3.4</td>
<td>29</td>
<td>5</td>
<td>20 30 25 40 55</td>
<td>70 85 95 95 100</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>Pre</td>
<td>ABC Harmony</td>
<td>4</td>
<td>27</td>
<td>17</td>
<td>30 35 35 80 130</td>
<td>80 90 130 130 130</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>Pre</td>
<td>ABC 90k Harmony</td>
<td>3.8</td>
<td>17</td>
<td>17</td>
<td>25 25 25 25 25</td>
<td>70 85 90 90 90</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>4.1</td>
<td>18</td>
<td>18</td>
<td>40 40 40 45 45</td>
<td>75 80 95 85 85</td>
</tr>
<tr>
<td>5</td>
<td>54</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>5.3</td>
<td>38</td>
<td>38</td>
<td>40 40 30 45 95</td>
<td>80 95 90 110 120</td>
</tr>
<tr>
<td>6</td>
<td>39</td>
<td>Pre</td>
<td>ABC 90k Harmony</td>
<td>6</td>
<td>51</td>
<td>51</td>
<td>20 30 30 35 65</td>
<td>65 75 110 105 105</td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>Pre</td>
<td>ABC HiRes 90k/HiFocus Harmony</td>
<td>1.4</td>
<td>29</td>
<td>29</td>
<td>35 35 40 50 65</td>
<td>65 75 100 95 80</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>1.5</td>
<td>23</td>
<td>23</td>
<td>30 30 40 35 35</td>
<td>— — — — —</td>
</tr>
<tr>
<td>9</td>
<td>54</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>1</td>
<td>51</td>
<td>51</td>
<td>20 40 25 25 35</td>
<td>60 95 95 85 90</td>
</tr>
<tr>
<td>10</td>
<td>23</td>
<td>Pre</td>
<td>Medel Opus2</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>35 30 35 25 40</td>
<td>60 60 85 95 130</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>Pre</td>
<td>ABC 90k Harmony</td>
<td>4</td>
<td>46</td>
<td>46</td>
<td>25 30 30 65 90</td>
<td>75 80 100 105 120</td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>4.2</td>
<td>19</td>
<td>19</td>
<td>40 45 40 35 60</td>
<td>65 75 85 90 105</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>25 20 30 50 75</td>
<td>65 80 95 105 120</td>
</tr>
<tr>
<td>14</td>
<td>49</td>
<td>Pre</td>
<td>ABC 90k Harmony</td>
<td>3</td>
<td>38</td>
<td>38</td>
<td>30 50 45 50 75</td>
<td>65 90 115 130 130</td>
</tr>
<tr>
<td>15</td>
<td>67</td>
<td>Post</td>
<td>Nucleus Freedom</td>
<td>10.3</td>
<td>26</td>
<td>26</td>
<td>25 40 45 75 90</td>
<td>70 85 105 130 130</td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>2.8</td>
<td>45</td>
<td>45</td>
<td>30 35 40 40 85</td>
<td>65 85 105 95 95</td>
</tr>
<tr>
<td>17</td>
<td>56</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>5.3</td>
<td>20</td>
<td>20</td>
<td>25 40 45 70 90</td>
<td>75 80 105 110 120</td>
</tr>
<tr>
<td>18</td>
<td>37</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>5.2</td>
<td>36</td>
<td>36</td>
<td>45 45 50 130 130</td>
<td>100 100 110 110 115</td>
</tr>
<tr>
<td>19</td>
<td>22</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>4.7</td>
<td>20</td>
<td>20</td>
<td>35 40 35 40 90</td>
<td>85 100 100 100 115</td>
</tr>
<tr>
<td>20</td>
<td>38</td>
<td>Pre</td>
<td>Nucleus Freedom</td>
<td>2</td>
<td>34</td>
<td>34</td>
<td>45 50 45 90 90</td>
<td>75 100 110 130 130</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>Pre</td>
<td>ABC c-1 BTE</td>
<td>10.4</td>
<td>19</td>
<td>19</td>
<td>55 45 40 50 60</td>
<td>60 80 100 100 85</td>
</tr>
<tr>
<td>22</td>
<td>67</td>
<td>Post</td>
<td>ABC HiRes 90k/HiFocus Harmony</td>
<td>1</td>
<td>22</td>
<td>— — — — — —</td>
<td>— — — — — —</td>
<td>— — — — — —</td>
</tr>
<tr>
<td>23</td>
<td>64</td>
<td>Post</td>
<td>Nucleus Freedom</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>60 35 35 30 70</td>
<td>70 65 70 65 75</td>
</tr>
<tr>
<td>24</td>
<td>61</td>
<td>Pre</td>
<td>ABC Harmony</td>
<td>4</td>
<td>36</td>
<td>36</td>
<td>55 35 40 65 130</td>
<td>55 65 85 105 115</td>
</tr>
<tr>
<td>25</td>
<td>33</td>
<td>Post</td>
<td>Nucleus Freedom</td>
<td>4</td>
<td>12</td>
<td>12</td>
<td>70 55 50 45 55</td>
<td>80 95 90 75 70</td>
</tr>
</tbody>
</table>

Appendix A

Bimodal Stimulation and Speech Perception 257
## Appendix B: Acoustic Analysis of the Intonation Test

### Table B1  Fundamental frequency (F0) parameters

<table>
<thead>
<tr>
<th></th>
<th>Mean F0 (Hz)</th>
<th>F0 at the beginning of the utterance (Hz)</th>
<th>F0 at the end of the utterance (Hz)</th>
<th>Minimum F0 (Hz)</th>
<th>Maximum F0 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement</td>
<td>230.06</td>
<td>246.97</td>
<td>193.43</td>
<td>171.06</td>
<td>307.05</td>
</tr>
<tr>
<td>Yes/no question</td>
<td>265.12</td>
<td>235.67</td>
<td>445.03</td>
<td>195.90</td>
<td>449.91</td>
</tr>
</tbody>
</table>

### Table B2  Amplitude and duration parameters

<table>
<thead>
<tr>
<th></th>
<th>Total utterance duration (s)</th>
<th>Last word duration (s)</th>
<th>Average amplitude (dB)</th>
<th>Amplitude of last word (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement</td>
<td>1.17</td>
<td>0.57</td>
<td>63.12</td>
<td>61.75</td>
</tr>
<tr>
<td>Yes/no question</td>
<td>0.99</td>
<td>0.52</td>
<td>62.4</td>
<td>62.94</td>
</tr>
</tbody>
</table>