Speech Prosody Perception in Cochlear Implant Users With and Without Residual Hearing

Mathieu Marx,[1,2,3] Christopher James,[3,4] Jessica Foxton,[1,2] Amandine Capber,[1,2] Bernard Fraysse,[3] Pascal Barone,[1,2] and Olivier Deguine[1,2,3]

Objectives: The detection of fundamental frequency (F0) variations plays a prominent role in the perception of intonation. Cochlear implant (CI) users with residual hearing might have access to these F0 cues. The objective was to study if and how residual hearing facilitates speech prosody perception in CI users.

Design: The authors compared F0 difference limen (F0DL) and question/statement discrimination performance for 15 normal-hearing subjects (NHS) and two distinct groups of CI subjects, according to the presence or absence of acoustic residual hearing: one "combined group" (n = 11) with residual hearing and one CI-only group (n = 10) without any residual hearing. To assess the relative contribution of the different acoustic cues for intonation perception, the sensitivity index d' was calculated for three distinct auditory conditions: one condition with original recordings, one condition with a constant F0, and one with equalized duration and amplitude.

Results: In the original condition, combined subjects showed better question/statement discrimination than CI-only subjects, d' 2.44 (SE 0.3) and 0.91 (SE 0.25), respectively. Mean d' score of NHS was 3.3 (SE 0.06). When F0 variations were removed, the scores decreased significantly for combined subjects (d' = 0.66, SE 0.51) and NHS (d' = 0.4, SE 0.09). Duration and amplitude equalization affected the scores of CI-only subjects (mean d' = 0.34, SE 0.28) but did not influence the scores of combined subjects (d' = 2.7, SE 0.02) or NHS (d' = 3.3, SE 0.33). Mean F0DL was poorer in CI-only subjects (34%, SE 15) compared with combined subjects (8.8%, SE 1.4) and NHS (2.4%, SE 0.05). In CI subjects with residual hearing, intonation d' score was correlated with mean residual hearing level ($r = -0.86$, n = 11, $p < 0.001$) and mean F0DL ($r = 0.84$, n = 11, $p < 0.001$).

Conclusion: Where CI subjects with residual hearing had thresholds better than 60 dB HL in the low frequencies, they displayed near-normal question/statement discrimination abilities. Normal listeners mainly relied on F0 variations which were the most effective prosodic cue. In comparison, CI subjects without any residual hearing had poorer F0 discrimination and showed a strong deficit in speech prosody perception. However, this CI-only group appeared to be able to make some use of amplitude and duration cues for statement/question discrimination.

Key words: Cochlear implant, F0 discrimination, Prosody perception, Residual hearing.

(Ear & Hearing 2015;36;239–248)

INTRODUCTION

Speech provides meaning through its semantic content, syntactic arrangement in the sentence, and the way words are uttered. Indeed, the meaning of a sentence can be altered using intonation variation, syllable, or word emphasis and pauses in discourse (Wingfield 1975; Speer et al. 1996). The acoustic correlates of these phenomena are variations in voice fundamental frequency (F0), amplitude, and duration (Nagel et al. 1994). These specific characteristics of language are referred to as suprasegmental information and gathered under the term "prosody."

Prosody yields information regarding the emotional status of the speaker, which can be inferred based on specific acoustic parameter changes. Primary emotions such as anger, fear, or happiness have been the focus of many studies of prosody. Anger in speech is, for example, associated with a higher voice fundamental frequency (F0) and a higher amplitude level on emphasized syllables (Fónagy 1966; Williams & Stevens 1972). Fonagy (1958) reported that happiness also leads to an increase in F0 and F0 range, with an increased speaking rate.

Prosody may also convey linguistic content, relying on the relative arrangement of lexical units and intonation variations in the oral discourse. The way words are grouped together can facilitate the extraction of the meaning of the message. For instance, it helps distinguish a command from a simple observation. Linguistic prosody, or speech prosody, also exerts its influence on the recognition of the sentential mode, that is, the discrimination between a question and a statement or an exclamation. In French, as in English, an interrogative sentence typically ends with an ascending intonation that reflects a rising F0, whereas intonation remains stable or decreases in a statement sentence (Patel et al. 1998).

Prosody Perception in Cochlear-Implanted Deaf Patients

For more than 25 years, cochlear implants (CIs) have restored excellent speech recognition (Rouger et al. 2007; Wilson & Dorman 2008). However other aspects of auditory perception remain problematic for many patients. Music perception for CI recipients is generally limited to rhythmic cues (Nimmons et al. 2008) and following speech in songs (Gfeller & Lansing 1991; El Fata et al. 2009) due to limitations in sound processing and the means of stimulation. Instrumental passages are hardly perceived and recognized because the representation of melody and timbre relies on too many acoustical cues that do not appear to be efficiently transmitted through the implant (Gfeller et al. 2002b). Similarly for prosody perception, many studies highlight poor recognition of speech intonation (Green et al. 2002; Peng et al. 2008) and lexical tone (Peng et al. 2004) for both pre- and postlingually deafened CI recipients. Nakata et al. (2012) assessed the perception of affective prosody by CI children and their normal-hearing counterparts and showed a strong deficit in the recognition of emotions in speech using CI. Their scores were above chance level for the recognition of "happy" and "sad" stimuli (mean 61 and 64% correct, respectively) but were significantly lower (mean 39%) for "angry."

[1]Laboratoire CerCo, Université Paul Sabatier, Toulouse, France; [2]Centre National de la Recherche Scientifique, UMR 5549, Toulouse, France; [3]Department of Otology-Neurotology and Skull Base Surgery, Purpan University Hospital, Toulouse, France; and [4]Cochlear France SAS, Toulouse, France.

The poor perception of music and affective speech prosody perception observed in CI recipients is often attributed to a deficit in the representation of F0. Indeed, resolution for both place and temporal information for F0 is lacking in current CIs. Place F0 encoding is limited by the number of electrodes (from 12 to 22 electrodes in current CIs) and possibly where limited insertion depth obtained with some electrode arrays may not allow stimulation of locations associated with low frequencies in the cochlea. Temporal coding of F0 in most current strategies is available via modulation rate cues up to the practical limit of perhaps 300 Hz corresponding to maximum rate discrimination (Shannon 1983). However, temporal coding of F0 does not appear to be always effective in this range (Milczynski et al. 2009) and certainly appears absent for higher modulation rates (Geurts & Wouters 2001; Vandali et al. 2005). Sound-coding strategies using fixed-rate carriers essentially discard temporal fine structure in the input signal except for low F0s which are represented by envelope modulation. Speech recognition in quiet can be largely supported by fluctuations in the amplitude envelope (e.g., <20 Hz), but speech recognition in noise (Stickney et al. 2005) and other challenging tasks such as prosody perception (Chatterjee & Peng 2008) require the correct perception of much higher amplitude modulation rates.

F0 representation might not be impaired where acoustic residual hearing coexists with CI. The perception of sound via residual acoustic hearing is generally fused with that of electric stimulation to form a unique and integrated auditory percept (Tyler et al. 2002). The residual acoustic hearing can be either contralateral to the CI as in "bimodal" stimulation, or both contralateral and ipsilateral to the CI as is the case for "hybrid" or electric acoustic stimulation (EAS). Bimodal and hybrid stimulation have shown advantages in auditory tasks that rely on F0 perception. Thus, CI patients with substantial residual hearing have better scores for speech recognition in noise (Armstrong et al. 1997; Adunka et al. 2010) and melody identification (Kong et al. 2004; El fata et al. 2009) than traditional CI subjects. Straatman et al. (2010) found that bimodal CI children discriminated more accurately questions from statement if they wore their hearing aid.

Using a statement/question discrimination paradigm, Peng et al. (2009) pointed to the necessity of co-operating cues in the signal to improve the performance of CI patients. More specifically, CI subjects relied on the congruency of intensity and F0 cues to perform the task, and these findings suggested the role of intensity cues in intonation recognition.

We hypothesized that speech prosody perception would be better in CI patients with bimodal stimulation or EAS because these patients have access to acoustic F0 cues through their residual hearing. Our corollary assumption was that better the residual hearing, the better prosody and F0 perception, as has been shown for song recognition (El Fata et al. 2009). Our second main hypothesis was that CI patients without residual hearing would rely on cues other than F0 to discriminate questions from statements. Therefore, the objectives of our study were to evaluate question/statement discrimination abilities in two distinct groups of cochlear-implanted patients according to the presence or absence of residual hearing and to identify which acoustic cues were more important for any one group. We also tested the pitch perception for simple harmonic tones to measure their overall capacity to perceive F0 without a speech context.

MATERIALS AND METHODS

Participants

Twenty-one postlingually deafened adult unilateral CI users participated in the study. Subject characteristics are given in Table 1. Four of them had postoperative bilateral residual hearing (subjects S1 to S4) and seven had significant levels of residual hearing in the ear contralateral to the CI (S5 to S11). These first 11 subjects were included in the group with combined electric and acoustic stimulation, called the "combined" group. Four of these subjects had bilateral residual hearing and were equipped with a Nucleus Freedom "Hybrid" sound processor. This sound processor additionally incorporates a receiver-in-the-ear acoustic amplification component. The Hybrid processors were programed using the manufacturer's own Cochlear Hybrid Prescription for gain and maximum power output incorporated into the Custom Sound software. The default 80 dB HL cutoff level was used to determine the nominal crossover frequency between acoustic and electric stimulation (Lenarz et al. 2013). For all other subjects, the manufacturers' default frequency to electrode allocations were used.

All 11 subjects with residual hearing wore hearing aids in the contralateral ear: these had been verified by an audiologist within 2 months before testing. Unaided pure-tone average air-conduction thresholds for the low frequencies (125–250–500 Hz) for the best (contralateral) ear are given in Table 2 for these subjects.

The other 10 subjects had no residual acoustic hearing ("CI-only" group) and did not use hearing aids in contralateral ears. No acoustic pure-tone thresholds better than 110 dB HL were found in any of the 10 subjects. The two groups were comparable with respect to aetiology of deafness. All CI users were assessed in their daily listening mode, that is, using unilateral or bilateral hearing aids, to measure optimal performance and to avoid learning and/or experience effects. We intended to focus our study on speech prosody perception and thus only included subjects with good speech perception scores in quiet (dissyllabic word recognition >70% correct).

Fifteen normal-hearing subjects (NHS) served as controls for the modified stimulus conditions. These 15 subjects had no history of otologic pathology or neurological disorder. Pure-tone thresholds were between 0 and 20 dB HL, from 125 Hz to 8 kHz. All subjects were required to give informed consent to participate in the study, and an institutional ethics committee approval was obtained (n°AFSSAPS:B90183-20).

Speech Prosody Testing

We evaluated speech prosody perception using a statement–question intonation discrimination task. A sentence containing no syntactic indices was presented to the subject who had to determine whether this sentence was a question or a statement. Sentences were chosen from the study reported by Patel et al. (2008) where sentence pairs were acoustically similar. To evaluate the relative contribution of different acoustic cues in this task, three different conditions were tested.

In the first condition ("original" condition), new original and natural recordings were presented. A total of 24 sentences were recorded by two professional actors (one male and one female) who uttered them once in an interrogative mode and once in a statement mode.

TABLE 1. Demographic and device characteristics of subjects

Subject Number	Group	Age (yrs)	Aetiology of Deafness	Sex	Time Since Activation (mos)	Model of Implant	Residual Hearing	Processor	Strategy	Implant Side
1	Combined	43	Unknown	F	68	Cochlear CI 24	R+L (EAS)	Freedom	ACE	R
2	Combined	51	Unknown	F	60	Cochlear Hybrid L	R+L (EAS)	Freedom	ACE	L
3	Combined	47	Unknown	F	51	Cochlear Hybrid L	R+L (EAS)	Freedom	ACE	R
4	Combined	70	Unknown	M	7	Cochlear MRA	R+L (EAS)	Freedom	ACE	L
5	Combined	68	Unknown	M	60	Hi-Res 90 K HiFocus 1J	R (Bim)	Harmony	HiRes-S	L
6	Combined	63	Otosclerosis	F	24	Cochlear CI 24	L (Bim)	Freedom	ACE	R
7	Combined	62	Unknown	M	7	Cochlear CI 24	R (Bim)	Freedom	ACE	L
8	Combined	42	Unknown	M	18	Cochlear CI 24	R (Bim)	Freedom	ACE	L
9	Combined	43	Unknown	F	50	Cochlear CI 24	L (Bim)	Freedom	ACE	R
10	Combined	21	Unknown	F	10	Cochlear CI 24	R (Bim)	Freedom	ACE	L
11	Combined	68	Unknown	M	13	Hi-Res 90 K HiFocus 1J	R (Bim)	Harmony	Fidelity 120-S	L
12	CI only	76	Trauma	M	15	Hi-Res 90 K HiFocus 1J	No	Harmony	Fidelity 120-S	R
13	CI only	54	Unknown	M	44	Hi-Res 90 K HiFocus 1J	No	Harmony	HiRes-S	L
14	CI only	73	Meniere	F	7	Hi-Res 90 K HiFocus 1J	No	Harmony	Fidelity 120-S	R
15	CI only	76	Meniere	M	26	Hi-Res 90 K HiFocus 1J	No	Harmony	Fidelity 120-S	R
16	CI only	44	Unknown	M	60	Cochlear CI 24	No	Freedom	ACE	R
17	CI only	70	Unknown	F	84	Cochlear CI 24	No	Freedom	ACE	L
18	CI only	69	Otosclerosis	M	36	Cochlear CI 24	No	Freedom	ACE	R
19	CI only	66	Genetic	F	84	Cochlear CI 24	No	Freedom	ACE	L
20	CI only	69	Unknown	M	55	Cochlear CI 24	No	Freedom	ACE	L
21	CI only	72	Unknown	M	43	Hi-Res 90 K HiFocus 1J	No	Harmony	HiRes-S	R

ACE, Advanced Combinational Encoder; Bim, bimodal stimulation; CI, cochlear implant; EAS, Electric-Acoustic Stimulation.

In a second condition (flat F0), we removed F0 variations from the original stimuli to evaluate the role of F0 variations in statement–question discrimination. Using Praat® (Boersma & Weenink 2014), the mean F0 in the original recording was extracted and then maintained constantly through the whole sentence. To perform well under this condition, subjects would have to rely on time and amplitude variations as cues.

We used the French sentence stimuli created by Patel et al. (2008) for the third "synthetic" condition: Two versions, question and statement, of each of the 12 sentences were recorded. The final 24 stimuli were constructed using a cross-splicing technique so that members of a question/statement pair were acoustically identical until the final syllable. The timing of syllables was identical and the acoustic waveform amplitude and perceived intensity were approximately equal, leaving pitch as the only salient cue for discrimination.

For each of the three prosody conditions, the 24 sentences were presented in random order. Each sentence was presented first visually in large type on a laptop screen for 2 sec, then acoustically, so that subjects focused on the prosody discrimination task and not on semantic meaning. Four separate practice trials were presented before each test for each condition.

F0 Detection Threshold

This test aimed to determine the smallest perceptible difference in F0 between two stimuli, which is often referred as the F0 difference limen (F0DL). We used a classic adaptive procedure (Green 1993) to estimate F0DL for various standard F0s implemented using the MATLAB MLP toolbox (Grassi & Soranzo 2009) (available online for download http://www.psy.unipd.it/~grassi/mlp.html). A three-alternative forced-choice task was used, as suggested by Grassi and Soranzo (2009) because it provides a good trade off both in terms of the accuracy and the duration of the experiment. In MLP, several a priori hypotheses are generated according to the psychometric function of the subjects and, therefore, several hypothetical F0DLs are compared with the current set of responses. Then the hypothesis having the highest likelihood of resembling the actual subject's psychometric function is selected. The next stimulus is automatically presented at the predicted threshold level, here at the hypothesized F0DL. The likelihood of each hypothesis was recalculated after each trial, and here 30 trials were used.

Stimuli were five-tone harmonic complexes starting at F0, with constant spectral level and duration of 250 msec. They were presented at 60 dB SPL via Sennheiser HD 219 circum-aural ear phones. The test was run using MATLAB® code (the Mathworks, Inc., Natick, MA). Several F0DLs were measured with standard F0s being: 110, 220, 400, 500, and 750 Hz. The test used a three-alternative forced-choice procedure with 30 presentations of three stimuli. In each trial, two of the sounds had the same F0 and were called the F0 standards. The other one had a higher F0 and was called the F0 target. Between those three stimuli, subjects had to determine the stimulus with the highest F0. If the subject correctly identified the sound with the highest F0, the difference between the sounds' F0 was

TABLE 2. Unaided pure-tone thresholds (dB HL) for the better, contralateral ear in CI subjects with residual hearing (combined group)

Subject Number	PTA 125–500 Hz	125 Hz	250 Hz	500 Hz	750 Hz	1 kHz	2 kHz	4 kHz	8 kHz
1	30	20	30	40	90	90	95	100	100
2	55	40	50	65	75	85	x	x	x
3	31.6	10	30	55	55	100	110	x	x
4	51.6	45	50	60	65	80	90	110	105
5	45	na	50	40	45	35	45	x	x
6	71.6	70	75	70	na	65	75	80	x
7	78.3	65	85	85	90	90	80	80	100
8	81.6	65	90	90	95	95	95	95	x
9	31.6	25	25	45	na	110	105	x	x
10	73.3	45	80	95	na	100	100	115	x
11	78.3	65	85	85	85	85	85	110	x

PTA125-500 Hz, pure-tone average of 125, 250, and 500 Hz; na, not available; x, no measurable threshold.

adaptively reduced until the smallest discriminable difference in hertz between the frequencies was determined.

Speech Recognition

All subjects were tested on auditory-only open-set recognition of French dissyllabic words using the commonly used Fournier word lists, presented through loudspeakers at 60 dB SPL. Also, single lists of 15 everyday French sentences containing a total of 100 words each were presented at 65 dB SPL in quiet and in multitalker babble with a 10 dB signal-to-noise ratio (SNR) (MMBA2 lists). The thirty-six lists are scored by words correct and are a standard measure of speech recognition in noise for French CI users. Lists were drawn exhaustively at random.

Test Set-Up

All subjects were assessed sitting comfortably in a sound attenuated audiometric booth. All auditory stimuli were presented through Sennheiser® HD 219 circumaural ear phones at an intensity level of 60 dB SPL to match with conventional speech recognition measurements. Speech prosody tests were performed using Presentation® software and F0 threshold was assessed using the MLP procedure as described above. The three speech prosody conditions and the F0 threshold determination order were balanced across subjects to avoid training effects. No repetition of any test stimulus was permitted. The examiner was also inside the booth, recording subjects' responses and running the test. The whole procedure took one to one and a half hours including four separate training items for the speech prosody testing.

Data Analysis

In each question–statement discrimination condition, we used as a primary measure the hit rates for questions and statements converted to d' values (Creelman 1962). d' is a criterion of perception sensitivity independent of decision bias, relying on hit rate and false alarm rate according to Signal Detection Theory (Tanner & Swets 1954; Swets et al. 1966). The F0 detection thresholds (F0DL) were expressed as Weber fractions: smallest difference detected [(F0 in hertz − F0 standard in hertz)/F0 standard], expressed as a percentage.

Separate one-way analyses of variances (ANOVA) were performed on subjects' age, speech prosody d' scores, and F0DLs for each condition, with group (CI-only versus combined versus NHS) as factor. Post hoc Fisher's Least Significant Difference (LSD) tests were used to perform multiple comparisons and control the risk of type I error. Student's *t* tests were used to perform paired comparisons between the three conditions of the speech prosody perception procedure for each group. Pearson correlation coefficients were computed for each correlation analysis along with significance levels. All statistical analyses were performed using Statview® software.

RESULTS

Results for speech prosody tests, F0 discrimination, and speech recognition scores are summarized in Table 3.

Group Characteristics

Twenty-one postlingually deafened adult cochlear-implanted subjects were included in two different groups according to the absence (CI-only, n = 10) or presence of residual hearing (combined group, n = 11). Mean age was 52.5 years (SE 4.58) for the combined group and 66.9 years (SE 3.23) for the CI-only group. Mean age was significantly higher in the CI-only group compared with the combined group (Fisher's LSD test, $p = 0.012$). NHS were significantly younger than both CI groups (mean 27.9 years, SE 2.78, Fisher's LSD test, $p < 0.001$).

Mean delay postimplantation was 33.5 months (SE 7.31) for the combined group and 45.4 months (SE 8.27) for the CI-only group. This difference was not statistically significant (*t* test, $t[19] = 1.09$, $p = 0.29$).

To avoid effects due to overall auditory performance, only subjects with dissyllabic words recognition scores better than 70% were included in both CI groups. Comparisons between the combined group and the CI-only group showed no significant difference in dissyllabic word recognition (85.9%, SE 3.0 for the combined group versus 79.4%, SE 3.58 for the CI-only group, *t* test, $t[19] = -1.39$, $p = 0.18$) or sentence recognition in quiet (93%, SE 5.34 for the combined group versus 95.1%, SE 1.82 for the CI-only group, *t* test, $t[19] = 0.34$, $p = 0.34$). The difference in mean recognition scores for sentences in noise between the combined group (89.7%, SE 4.03) and the CI-only group (77.1%, SE 5.07) approached statistical significance (*t* test, $t[19] = -1.97$, $p = 0.06$).

Speech Prosody Perception

Results of the three different speech prosody conditions are illustrated in Figure 1.

Original Condition

This condition was tested to evaluate speech prosody perception for “naturally” produced sentences, when all normal acoustic cues are present. Globally, subjects with acoustic hearing (combined or NHS group) performed better than subjects without any residual hearing (CI-only group). The mean score for the combined group was 87.1% (SE 3.7) and for the CI-only group was 65.4% (SE 4.07). The mean score for NHS was 97.9% (SE 0.66). Raw scores were converted to d’ scores (Fig. 1) using the equation by Creelman (1962). A one-way ANOVA on d’ scores revealed a significant effect for group ($F[2,33] = 33.37$, $p < 0.001$): Mean d’ was 2.44 (SE 0.3) for the combined group, 0.91 (SE 0.25) for the CI-only group, and 3.27 (SE 0.06) for the NHS group. Paired comparisons showed lower scores for the CI-only group compared with that of the combined group (Fisher’s LSD test, $p < 0.001$) and of NHS (Fisher’s LSD test, $p < 0.001$). Performance of the combined group was significantly poorer than that of NHS (Fisher’s LSD test, $p = 0.006$). Correlation analyses were performed to examine the impact of experience with CI within the two groups. d’ scores for the combined group were significantly and positively related to time since activation ($r = 0.89$, n = 11, $p < 0.001$). For the CI-only group, there was no statistically significant relationship between time since activation and prosody perception performance ($r = −0.1$, n = 10, $p = 0.79$).

Constant F0 Condition

This condition allowed us to evaluate the effect of removing F0 variations on statement/question discrimination. There was no significant effect of group on d’ scores for this condition ($F[2,33] = 0.73$, $p = 0.49$). The mean score was 61.7% (SE 2.68) for the combined group, 58.7% (SE 1.45) for the CI-only group, and 58.1% (SE 1.45) for NHS. Corresponding d’ values were 0.66 (SE 0.51) for the combined group, 0.53 (SE 0.19) for the CI-only group, and 0.44 (SE 0.09) for the NHS group.

Paired comparisons for d’ in combined group showed a significant decrease in performance ($t[10] = 4.14$, $p = 0.002$) between original and constant F0 conditions. The decrease was not significant for the CI-only group ($t[9] = 2.38$, $p = 0.55$).

NHS performance was also significantly affected by the absence of an F0 contour. Their scores dropped from near perfect in the original condition (mean score = 97.9%, SE 0.66; d’ = 3.27, SE 0.06) to near chance (mean score = 58%, SE 5.63, with d’ = 0.44, SE 0.09).

Synthetic Condition

The French language stimuli of Patel et al. (2008) were used for this condition. Starting from natural recordings of affirmative and interrogative forms of the same sentence, the pairs of stimuli only differed in their F0 contour for the last syllable/word (Patel et al. 2008), and not in duration or intensity (see Materials AND Methods). As the F0 contour is the prominent cue in speech prosody perception, we expected that scores of NHS would not be affected by this acoustic manipulation.

There was a statistically significant effect of group on d’ scores for the synthetic condition ($F[2,33] = 69.35$, $p < 0.001$). NHS had high scores (mean correct 98.2%, SE 3.72, d’ 3.3, SE 0.33). Similarly for combined group subjects, d’ scores were not significantly different from those for the original condition (synthetic condition mean d’ 2.7, SE 0.02 versus original condition d’ 2.44, SE 0.3, $t[10] = −1.8$, $p = 0.12$). In contrast, CI-only subjects showed a significant drop in scores for the synthetic condition versus the original condition ($t[9] = 2.3$, $p = 0.047$) with d’ = 0.91 (SE 0.25) for the original condition compared with 0.34 (SE 0.28) for the synthetic condition.

F0 Perception

F0DLs were measured for several frequency ranges (110–220–400–500–750 Hz) for each subject and an aggregate, mean F0DL was then calculated. An ANOVA with group as factor was performed on mean F0DL for the CI groups and NHS (Fig. 2). There was a significant effect of group ($F[2, 33] = 4.66$, $p = 0.016$). Mean F0DL was significantly higher (poorer) for the CI-only group compared with the combined group (34.0%, SE 15.0 versus 8.8%, SE 1.4, Fisher’s LSD test, $p = 0.03$) and compared with the NHS group (2.2%, SE 0.05, Fisher’s LSD test, $p = 0.004$). When the scores were analyzed frequency by frequency, the combined group had significantly better/lower F0DLs than the CI-only group when F0 standard was 110 and 750 Hz (Fisher’s LSD test, respectively, $p = 0.0009$ and $p = 0.007$). Otherwise, no differences could be identified due to the wide variability in CI-only subjects’ performance.

These results showed that intonation perception and F0 discrimination were better in CI subjects with residual acoustic hearing. CI-only subjects displayed a significant deficit in speech prosody perception and variable F0DLs.

Impact of Residual Hearing Level

The relationship between speech prosody perception, F0DL perception, and residual hearing level in the combined group was further explored. The best hearing thresholds for each subject were found in the nonimplanted ear; pure-tone average low-frequency hearing losses using 125, 250, and 500 Hz are presented in Table 2 for these ears.

Mean residual hearing level and speech prosody d’ scores in the original condition were highly correlated ($r = −0.86$, n = 11, $p = 0.001$) (Fig. 3B); as were mean residual hearing level and mean F0DL ($r = −0.84$, n = 11, $p < 0.001$) (Fig. 3A). It appeared therefore that the level of residual hearing had a large impact

TABLE 3. Summary of auditory performance for speech prosody conditions (original condition, constant F0 condition, synthetic condition), F0 discrimination (mean F0DL), and speech recognition scores

	D’ Original Condition*	D’ Constant F0	D’ Synthetic Condition*	Mean F0DL*	Words in Quiet	Sentences in Quiet	Sentences in Noise*
Combined, n = 11	2.44 (0.3)	0.66 (0.14)	2.75 (0.2)	8.8% (1.9)	85.9% (3.0)	93% (5.34)	89.7% (4.03)
CI only, n = 10	0.91 (0.25)	0.53 (0.19)	0.34 (0.28)	34% (16.0)	79.4% (3.58)	95.1% (0.82)	77.1% (5.07)

Values are group means with the standard error of means in parentheses.
**Statistically significant differences were found between groups for the measure.*

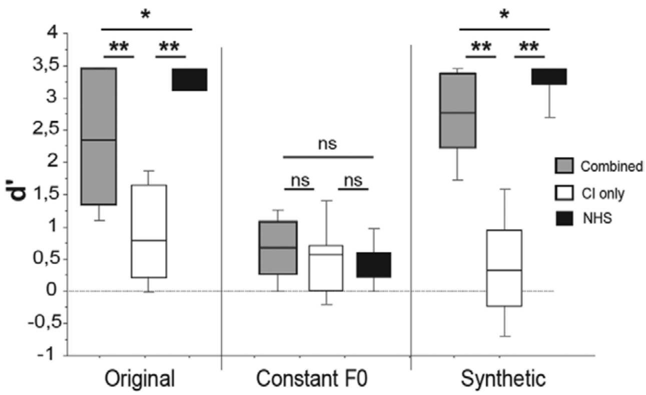

Fig. 1. Box plots summarizing d' scores obtained by cochlear-implanted subjects in the combined group (grey) and in the cochlear implant (CI)–only group (white) and by normal-hearing subjects (NHS, black). *Statistically significant differences were found between the two groups for the measure (*$p < 0.05$; **$p < 0.0001$). ns indicates no statistically significant difference was found between groups for the measure.

on intonation perception. From observation of Figure 3, one can identify six subjects (S1, S2, S3, S4, S5, and S9) with high question–statement discrimination scores (d'>2) with a residual hearing level better than 60 dB and the other five subjects with poorer residual hearing who had correspondingly lower d' scores (S6, S7, S8, S10, and S12). The six best performing subjects were the four "EAS" subjects with bilateral residual hearing and two of the "bimodal" subjects with residual hearing contralateral to the implant.

Relation Between F0 Perception and Speech Prosody Perception

We may conclude that F0 cues were relevant for CI subjects with residual hearing because d' scores decreased significantly between the original and the constant F0 condition. In contrast, performance of the CI-only subjects was not affected when F0 variations were removed. It was therefore plausible that the relationship between the speech prosody perception performance in the original condition and the level of F0 discrimination would not be significant in this group of subjects.

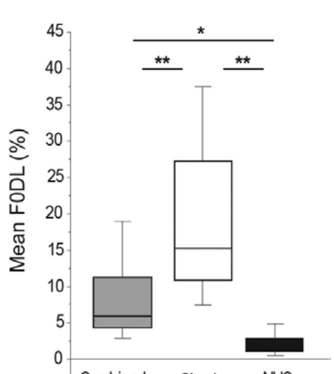

Fig. 2. Box plots summarizing mean F0 difference limen (F0DL) in cochlear-implanted subjects in the combined group (grey) and in the cochlear implant (CI)–only group (white) and by normal-hearing subjects (NHS, black). *Statistically significant differences were found between groups for the measure (*$p < 0.05$; **$p < 0.0001$). ns indicates no statistically significant difference was found between groups for the measure.

Considering all 21 implanted subjects, a global significant negative correlation between F0DL and question/statement discrimination in original condition was found ($r = -0.47$, n = 21, $p = 0.03$). Considering only those subjects with residual hearing (combined group), a higher correlation was obtained between the two variables ($r = -0.67$, n = 11, $p = 0.03$). Within the CI-only group, no significant correlation was found ($r = 0.34$, n = 10, $p = 0.38$).

DISCUSSION

Residual Hearing Advantages in F0-Related Tasks

Our results highlight how residual hearing may help CI patients with prosody perception. In the present study, CI subjects with residual hearing (combined group), using either bimodal stimulation (residual hearing in the ear contralateral to the implant) or electric-acoustic stimulation (residual hearing in both ears and CI), achieved better performance than subjects using a CI alone (CI-only group) in discriminating questions from statements in the "real-life," original condition. In our study, CI subjects with residual acoustic hearing better than 60 dB HL for the range 125 to 500 Hz achieved high d' scores in the original condition, comparable with that of NHS. Furthermore, there were strong correlations both between residual hearing level and F0DL ($r = 0.84$, n = 11, $p = 0.001$) and between residual hearing level and d' scores for question/statement discrimination (original condition, $r = -0.86$, n = 11, $p < 0.001$). Altogether, these results indicate that residual hearing, when sufficient, provided the F0 information necessary for the perception of this form of intonation. The relative role of acoustic and electric cues in combined subjects warranted

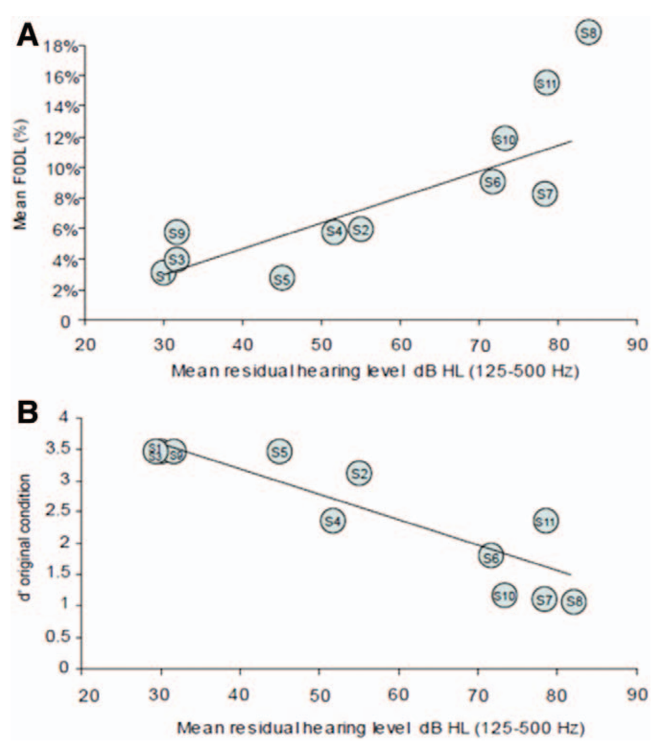

Fig. 3. A, Correlation between mean F0 difference limen (F0DL) and mean residual hearing level 125, 250, and 500 Hz, for combined subjects. B, Correlation between d' score in original condition and mean residual hearing level 125, 250, and 500 Hz for combined subjects.

further investigation: four additional subjects who used a CI and contralateral hearing aid with thresholds better than 60 dB HL from 125 to 500 Hz were tested in three different listening modes, such as acoustic-only, electric-only, and combined electric-acoustic. Subjects were presented with the "original" stimuli in the question–statement discrimination task. When these combined subjects used their daily listening condition, CI and contralateral hearing aid, all of them obtained for the highest d' score possible (3.46). d' scores decreased for three subjects to 0.43, 2.35, and 1.93, respectively, when listening via electric stimulation only (no contralateral hearing aid and ear canal plugged). The fourth subject maintained a high d' score (3.46). Finally, an acoustic-only evaluation was performed, where the four subjects wore the contralateral hearing aid but not the CI. The d' scores were 3.46, 3.11, 3.11, and 3.46. Thus in this limited sample, acoustic hearing appeared to be sufficient to support high scores on the task. However, d' scores in the electric-only condition appeared surprisingly good for three subjects (2.35, 1.93, and 3.46), and better than the maximum score obtained for CI-only subjects (i.e., 1.64 by S19). In our ongoing research into temporal pitch, we have since found a few individuals who use electrical stimulation alone and have no residual hearing who can also achieve high d' scores on the intonation test, this appears to be related to a greater ability to use temporal pitch cues. In addition, combined subjects may be able to generalize learning gained in their habitual "bimodal" electric-acoustic condition when processing speech in a single modality, as it has been shown in audiovisual interactions (Shams & Seitz 2008; Barone & Deguine 2011). Similar mechanisms could be engaged in prosody processing when performed by combined subjects. Indeed, for the combined group, there was a significant correlation between the time since activation and d' scores for the original condition. This effect might be explained by subjects needing some time to fuse acoustic and electric hearing (Tyler et al. 2002). This area deserves further research to determine under what conditions limitations exist for pitch processing with CI for individuals and whether training can improve performance for CI-alone (Fu & Galvin 2007).

Benefits for bimodal stimulation or electric acoustic hearing have been shown in other F0-related tasks, as for speech recognition in noise (Armstrong et al. 1997; Fraysse et al. 2006; Adunka et al. 2010) or for music perception (Leal et al. 2003; El Fata et al. 2009). Straatman et al. (2010) showed a bimodal advantage for speech prosody in 17 CI children with residual hearing. They compared their question/statement discrimination performance when children were wearing or were not wearing their contralateral hearing aid. Using their residual hearing, correct recognition was improved between 6 to 11 percentage points. Cullington and Zeng (2010) compared bimodal stimulation to bilateral CI in different tasks, including affective prosody (Ross et al. 1997). Although no significant difference in this specific test was found between bilateral and bimodal CI users, performance tended to be better in bimodal patients, who displayed comparable scores to NHS controls, in contrast with bilateral CI subjects.

Although our experiments did not include bilateral cochlear-implanted subjects, we studied CI patients with comparable speech in quiet recognition scores to focus on the hypothetical benefit of residual hearing in F0 and speech prosody perception. The positive impact of residual hearing was demonstrated by both EAS and bimodal subjects showing better speech prosody perception than CI-only subjects. However, CI subjects in the combined group displayed better d' scores than CI-only subjects in original condition and six of them with the best residual hearing had particularly high d' scores, comparable with NHS. It appears that average low-frequency residual hearing better than 60 dB HL allows for excellent speech prosody perception.

Information Provided by Residual Acoustic Hearing

Residual hearing allows for acoustic transmission of F0 and thereby more accurate perception of intonation in the sentence. Indeed, mean F0DL was significantly better in the combined group subjects with the best acoustic hearing. The acoustic transmission of F0 may help CI subjects in music perception (Gfeller et al. 2002a, Gantz et al. 2006) or for recognizing speech in noise (Nooteboom & Terken 1982; Assmann & Summerfield 1990). But residual hearing may convey additional information to F0, plausibly relevant for speech prosody perception, like first formant transitions. Kong and Carlyon (2007) showed that the EAS advantage for speech in noise recognition persisted at low SNRs even when F0 cues were removed from the low-passed stimulus. Furthermore, this advantage disappeared at high SNR levels when F0 cues were preserved but low-frequency phonetic cues were eliminated. Besides F0, low-frequency hearing could provide useful information to target the talker in background noise. Li and Loizou (2008) suggested that low-frequency hearing was used to glimpse speech information in modulated background noise and to detect and integrate acoustic cues characterizing the speaker. Speech harmonics falling in the low frequencies are perceived and provide reliable F0 information—the better residual hearing the better the resolution of harmonics. In addition, residual hearing may also contribute to the representation of the first formant (F1) frequency, which is critical for vowel and stop-consonant identification (voicing information), even at low SNRs. To get CI-only subjects to perform as well as EAS or bimodal users in F0-related situations (speech in noise recognition, music perception, prosody perception, talker identification), F0 coding and ideally spectral resolution in the low-frequency region should be improved.

F0 Encoding in CIs

Our F0DL results for the CI-only group are consistent with the literature, both quantitatively and qualitatively. It has been reported that the best scores for CI subjects reach F0DLs around 10% of a reference F0 (Geurts & Wouters 2001). Here the average F0DL was 34% (SE 0.28) in the CI-only group. Also, as seen here, substantial variability in F0 discrimination has been consistently reported in the literature: Pressnitzer et al. (2005) reported F0DLs between 2 and 7 semitones, that is, between 16.5 and 58.5%. Geurts and Wouters (2001) found F0DLs between 4 and 13% at 150 Hz for modulated pulse trains presented on single channels. In their study, at 250 Hz, one subject was not able to perform the task and in the three others F0DL varied between 5 and 12%.

Although most CI subjects have impaired F0 discrimination, they are still able to perform gross F0-related tasks, such as sex identification. But finer F0 discrimination is required for more difficult F0 tasks, such as melody perception or sex categorization of ambiguous voices. In another study by our group (Massida et al. 2013), performance for sex identification was similar between CI users and NHS for typical male and female voices at the extremes of a continuum (CI >90% of score of NHS), but scores dropped by 63 percentage points in the case of more ambiguous voices. This is similar to the findings by Kovacić and Balaban (2009) where sex categorization was generally only accurate for CI users outside of the overlapping male–female F0 region between 138 and 163 Hz.

One of the main approaches that has been adopted to improve F0 perception is to enhance the temporal representation. Among the different clinical strategies currently available, the Fine Structure Processing (FSP) and Fine Structure 4 processing strategies from Med-El® attempt to represent fine structure information on one to four apical electrodes by stimulating very short bursts of biphasic pulse trains at each zero crossing of the output of the corresponding filter bank. As a result, apical electrodes provide timing information via a dynamically changing stimulation rate. The clinical performance of this signal processing has been assessed in various F0-related tasks, such as Mandarin speech recognition (Chen et al. 2013), music perception (Looi et al. 2011; Landwehr et al. 2014), and speech recognition in noise (Vermeire et al. 2010; Riss et al. 2011). Benefit for FSP was shown for tone perception (Chen et al. 2013), music perception, and appreciation (Looi et al. 2011; Landwehr et al. 2014). In addition, FSP gave equivalent or better performance than the standard Continuous Interleaved Sampled (CIS) strategy for speech recognition in noise (Vermeire et al. 2010; Riss et al. 2011). Studies by Green et al. (2004, 2005) assessed a 100% modulated sawtooth-like envelope distributed to all electrodes and found not only an improvement of intonation recognition but also a reduction in vowel recognition. Vandali et al. (2005) evaluated four different experimental strategies that enhanced temporal information. The Peak Derived Timing strategy was intended to code fine temporal structure in each channel using positive temporal peaks in the band-pass filtered signal, which might improve the perception of individual harmonics in voiced speech or musical sounds. The three other strategies used modifications of the Advanced Combinational Encoder (ACE) traditionally used in Nucleus® CI sound processors. The Modulation Depth Enhancement (MDE), the F0-synchronized ACE (F0sync), and the Multi-channel Envelope Modulation provided more F0-modulation cues, based on deeper amplitude modulation. These modulations were aligned in time across all channels to reduce phase differences between neighboring electrodes. Three experimental strategies (MDE, F0sync, and Multi-channel Envelope Modulation) provided better scores than the clinical strategy (ACE) for sung vowels, but two of them (MDE and F0sync) affected speech recognition in noise. They should be further explored for musical stimuli and/or tonal languages. Another approach, aimed to improve F0 place coding, was devised by Geurts and Wouters (2004), using an original triangle filter bank designed to resolve the first harmonic. Results showed improved F0 detection, but speech recognition was not assessed. F0 place coding might be facilitated with the "current steering" design, applied in Advanced Bionics® CIs using Fidelity 120 strategy (as used by some Advanced Bionics® subjects in our study). The output current is here steered between two distinct electrodes to induce an intermediate F0 percept. This strategy implements up to 120 theoretically distinct virtual channels, but has not, to date, shown a strong benefit over 16-channel CIS for overall auditory performance (Buechner et al. 2010).

Thus, although the improvement of F0 transmission using refined coding strategies is necessary, most of the results reported to date remain disappointing. Indeed, these experimental and/or new strategies provided only limited benefits for F0-related tasks when compared with the potentially high contribution of residual hearing. In children, these results raise the issue of bilateral implantation and the need for a proper preoperative hearing thresholds evaluation. They suggest caution more specifically in cases of doubtful low-frequency hearing thresholds and might support sequential bilateral cochlear implantation rather than simultaneous implantation after an attentive follow-up of the acoustic hearing in the nonimplanted ear.

F0 is undeniably the prominent cue supporting prosodic information, but the other acoustic cues might play a facilitating role. Peng et al. (2009) demonstrated the facilitating role of intensity for intonation labeling when its variations are congruent with F0 contour information. In their study, the proportion of "question" recognition was higher when the stimulus ended with an ascending F0 and an increment of intensity. This proportion decreased in cases of conflicting cues, that is, when an ascending F0 and a decreased intensity were present at the end of the stimulus. Our results obtained in the CI-only group also lead us to consider intensity as a prosodic cue. In the synthetic condition of the question–statement test in the present study, CI-only subjects showed a significant decrease in recognition scores compared with their performance in the original condition (mean d' dropped from 0.91 to 0.34; mean correct recognition scores from 65.4 to 55.4% or very near chance level). Therefore, the absence of intensity cues significantly and negatively influenced the performance of CI-only subjects in statement–question recognition. However, it should be emphasized that their overall performance was very poor even for the original condition.

Limitations

This study was conducted on a small number of subjects, but the differences between groups were large and statistically significant. One of the main limitations is the difference between mean ages across groups, which might have affected CI-only group's performance. Several studies demonstrated the effect of aging on F0-related tasks (Souza et al. 2011; Schvartz & Chatterjee 2012). In the present study, however, there was no correlation between age and mean F0DL ($r = 0.24$, n = 21, $p = 0.3$) or between age and d' in the original condition ($r = -0.37$, n = 21, $p = 0.09$).

CONCLUSIONS

Residual hearing combined with a CI provided F0 via low-frequency harmonic information that is required for speech prosody perception. In our study, residual hearing level was highly correlated to both F0 detection and question–statement discrimination. Furthermore, mean acoustic residual hearing better than 60 dB HL for 125 to 500 Hz in CI subjects provided near-normal performance in speech prosody recognition, which emphasizes the potential complementarity between acoustic and electric cues. The significant correlation between time since CI activation and d' scores for the combined group for the original prosody condition suggests that this complementarity might be optimized over time.

ACKNOWLEDGMENTS

The authors thank the cochlear-implanted subjects and normal-hearing subjects for their participation in this study, M.-L. Laborde for her help in collecting the data, and C. Marlot for her help with the bibliography.

This work was supported by Direction de la Recherche Clinique et de l'Innovation, CHU Toulouse, the ANR Plasmody (P.B.) and CNRS (M.M.).

The authors declare no other conflict of interest.

Address for correspondence: Mathieu Marx, Department of Otology-Neurotology and Skull Base Surgery, Purpan University Hospital, Place du Docteur Baylac, 31059 Toulouse, France. E-mail: marx.m@chu-toulouse.fr

Received February 23, 2013; accepted August 7, 2014.

REFERENCES

Adunka, O. F., Pillsbury, H. C., Adunka, M. C., et al. (2010). Is electric acoustic stimulation better than conventional cochlear implantation for speech perception in quiet? *Otol Neurotol*, *31*, 1049–1054.

Armstrong, M., Pegg, P., James, C., et al. (1997). Speech perception in noise with implant and hearing aid. *Am J Otol 18*, S140–S141.

Assmann, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *J Acoust Soc Am*, *88*, 680–697.

Barone, P., & Deguine, O. (2011). Multisensory processing in cochlear implant listeners. In F. G. Zeng, R. Fay, A. Popper (Eds.), *Springer Handbook of Auditory Research. (Shar 39)*. *Auditory Prostheses: Cochlear Implants and Beyond* (pp. 365–382). New York, NY: Springer-Verlag.

Boersma, P., & Weeninck, D. (2014). Praat: Doing phonetics by computer [Computer program]. Version 5.3.77. Retrieved May 18, 2014, from http://www.praat.org/.

Buechner, A., Frohne-Büchner, C., Gaertner, L., et al. (2010). The advanced bionics high resolution mode: Stimulation rates up to 5000 pps. *Acta Otolaryngol*, *130*, 114–123.

Chatterjee, M., & Peng, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hear Res*, *235*, 143–156.

Chen X., Liu B., Liu S., et al. (2013). Cochlear implants with fine structure processing improve speech and tone recognition in Mandarin-speaking adults. *Acta Otolaryngol*, *133*, 733–738

Creelman, C. (1962). Human discrimination of auditory duration. *J Acoust Soc Am*, *34*, 582–593.

Cullington, H. E., & Zeng, F. G. (2010). Bimodal hearing benefit for speech recognition with competing voice in cochlear implant subject with normal hearing in contralateral ear. *Ear Hear*, *31*, 70–73.

El Fata, F., James, C. J., Laborde, M. L., et al. (2009). How much residual hearing is "useful" for music perception with cochlear implants? *Audiol Neurootol*, *14*(Suppl 1), 14–21.

Fonagy, I. (1958). Phonetics & psychoanalysis; psychological remarks on the stress accent. *Psyche*, *12*, 63–72.

Fónagy, I. (1966). Electrophysiological and acoustic correlates of stress and stress perception. *J Speech Hear Res*, *9*, 231–244.

Fraysse, B., Macías, A. R., Sterkers, O., et al. (2006). Residual hearing conservation and electroacoustic stimulation with the nucleus 24 contour advance cochlear implant. *Otol Neurotol*, *27*, 624–633.

Fu, Q.-J., & Galvin J. J. (2007). Perceptual learning and auditory training in cochlear implant recipients. *Trends Amplif*, *11*, 193.

Gantz, B. J., Turner, C., Gfeller, K. E. (2006). Acoustic plus electric speech processing: Preliminary results of a multicenter clinical trial of the Iowa/Nucleus Hybrid implant. *Audiol Neurootol*, *11* (Suppl 1), 63–68.

Geurts, L., & Wouters, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *J Acoust Soc Am*, *109*, 713–726.

Geurts, L., & Wouters, J. (2004). Better place-coding of the fundamental frequency in cochlear implants. *J Acoust Soc Am*, *115*, 844–852.

Gfeller, K., & Lansing, C. R. (1991). Melodic, rhythmic and timbral perception of adult cochlear implant users. *J Speech Hear Res 34*, 916–920.

Gfeller, K., Turner, C., Mehr, M., et al. (2002a). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int*, *3*, 29–53.

Gfeller, K., Witt, S., Woodworth, G., et al. (2002b). Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Ann Otol Rhinol Laryngol*, *111*, 349–356.

Grassi, M., & Soranzo, A. (2009). MLP: A MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behav Res Methods*, *41*, 20–28.

Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes–no task. *J Acoust Soc Am*, *93*(4 Pt 1), 2096–2105.

Green, T., Faulkner, A., Rosen, S. (2002). Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants. *J Acoust Soc Am*, *112*(5 Pt 1), 2155–2164.

Green, T., Faulkner, A., Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *J Acoust Soc Am*, *116*, 2298–2310.

Green, T., Faulkner, A., Rosen, S., et al. (2005). Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification. *J Acoust Soc Am*, *118*, 375–385.

Kong, Y. Y., & Carlyon, R. P. (2007). Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *J Acoust Soc Am*, *121*, 3717–3727.

Kong, Y. Y., Cruz, R., Jones, J. A., et al. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear Hear*, *25*, 173–185.

Kovacić, D., & Balaban, E. (2009). Voice gender perception by cochlear implantees. *J Acoust Soc Am*, *126*, 762–775.

Landwehr, M., Fürstenberg, D., Walger, M., et al. (2014). Effects of various electrode configurations on music perception, intonation and speaker gender identification. *Cochlear Implants Int*, *15*, 27–35.

Leal, M. C., Shin, Y. J., Laborde, M. L., et al. (2003). Music perception in adult cochlear implant recipients. *Acta Otolaryngol*, *123*, 826–835.

Lenarz, T., James, C., Cuda, D., et al. (2013). European multi-centre study of the Nucleus Hybrid L24 cochlear implant. *Int J Audiol*, *52*, 838–848.

Li, N., & Loizou, P. C. (2008). A glimpsing account for the benefit of simulated combined acoustic and electric hearing. *J Acoust Soc Am*, *123*, 2287–2294.

Looi, V., Winter, P., Anderson, I., et al. (2011). A music quality rating test battery for cochlear implant users to compare the FSP and HDCIS strategies for music appreciation. *Int J Audiol*, *50*, 503–518.

Massida Z., Marx M., Belin P., et al. (2013). Gender categorization in cochlear-implanted deaf subjects. *J Speech Lang Hear Res*, *56*, 1389–1401.

Milczynski, M., Wouters, J., van Wieringen, A. (2009). Improved fundamental frequency coding in cochlear implant signal processing. *J Acoust Soc Am*, *125*, 2260–2271.

Nagel, H. N., Shapiro, L. P., Nawy, R. (1994). Prosody and the processing of filler-gap sentences. *J Psycholinguist Res*, *23*, 473–485.

Nakata, T., Trehub, S. E., Kanda, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. *J Acoust Soc Am*, *131*, 1307–1314.

Nimmons, G. L., Kang, R. S., Drennan, W. R., et al. (2008). Clinical assessment of music perception in cochlear implant listeners. *Otol Neurotol*, *29*, 149–155.

Nooteboom, S. G., & Terken, J. M. (1982). What makes speakers omit pitch accents? An experiment. *Phonetica*, *39*, 317–336.

Patel, A. D., Peretz, I., Tramo, M., et al. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain Lang*, *61*, 123–144.

Peng, S. C., Lu, N., Chatterjee, M. (2009). Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiol Neurootol*, *14*, 327–337.

Peng, S. C., Tomblin, J. B., Cheung, H., et al. (2004). Perception and production of mandarin tones in prelingually deaf children with cochlear implants. *Ear Hear*, *25*, 251–264.

Peng, S. C., Tomblin, J. B., Turner, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear Hear*, *29*, 336–351.

Pressnitzer, D., Bestel, J., Fraysse, B. (2005). Music to electric ears: Pitch and timbre perception by cochlear implant patients. *Ann N Y Acad Sci*, *1060*, 343–345.

Riss, D., Hamzavi, J. S., Selberherr, A., et al. (2011). Envelope versus fine structure speech coding strategy: A crossover study. *Otol Neurotol*, *32*, 1094–1101.

Ross, E. D., Thompson, R. D., Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang*, *56*, 27–54.

Rouger, J., Lagleyre, S., Fraysse, B., et al. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc Natl Acad Sci U S A*, *104*, 7295–7300.

Schvartz, K. C., & Chatterjee, M. (2012). Gender identification in younger and older adults: Use of spectral and temporal cues in noise-vocoded speech. *Ear Hear*, *33*, 411–420.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends Cog Sci*, *12*, 411–417.

Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear Res*, *11*, 157–189.

Souza, P., Arehart, K., Miller, C. W., et al. (2011). Effects of age on F0 discrimination and intonation perception in simulated electric and electroacoustic hearing. *Ear Hear*, *32*, 75–83.

Speer, S. R., Kjelgaard, M. M., Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *J Psycholinguist Res*, *25*, 249–271.

Stickney, G. S., Nie, K., Zeng, F. G. (2005). Contribution of frequency modulation to speech recognition in noise. *J Acoust Soc Am*, *118*, 2412–2420.

Straatman, L. V., Rietveld, A. C., Beijen, J., et al. (2010). Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid. *J Acoust Soc Am*, *128*, 1884–1895.

Swets, J. A., Harris, J. R., McElroy, L. S., et al. (1966). Computer-aided instruction in perceptual identification. *Behav Science*, *11*, 98–104.

Tanner, W. P. Jr., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychol Rev*, *61*, 401–409.

Tyler, R. S., Parkinson, A. J., Wilson, B. S., et al. (2002). Patients utilizing a hearing aid and a cochlear implant: Speech perception and localization. *Ear Hear*, *23*, 98–105.

Vandali, A. E., Sucher, C., Tsang, D. J., et al. (2005). Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. *J Acoust Soc Am*, *117*, 3126–3138.

Vermeire, K., Punte, A. K., Van de Heyning, P. (2010). Better speech recognition in noise with the fine structure processing coding strategy. *ORL J Otorhinolaryngol Relat Spec*, *72*, 305–311.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *J Acoust Soc Am*, *52*, 1238–1250.

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hear Res*, *242*, 3–21.

Wingfield, A. (1975). Acoustic redundancy and the perception of time-compressed speech. *J Speech Hear Res*, *18*, 96–104.