

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/87983>

Please be advised that this information was generated on 2017-12-06 and may be subject to change.

# Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid

L. V. Straatman<sup>a)</sup>

Department of Otorhinolaryngology, Head and Neck Surgery, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

A. C. M. Rietveld

Department of Linguistics, Radboud University Nijmegen, Erasmusplein 1, 6525 HT Nijmegen, The Netherlands

J. Beijen, E. A. M. Mylanus, and L. H. M. Mens

Department of Otorhinolaryngology, Head and Neck Surgery, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

(Received 13 June 2008; revised 27 June 2010; accepted 12 July 2010)

Cochlear implants are largely unable to encode voice pitch information, which hampers the perception of some prosodic cues, such as intonation. This study investigated whether children with a cochlear implant in one ear were better able to detect differences in intonation when a hearing aid was added in the other ear (“bimodal fitting”). Fourteen children with normal hearing and 19 children with bimodal fitting participated in two experiments. The first experiment assessed the just noticeable difference in F<sub>0</sub>, by presenting listeners with a naturally produced bisyllabic utterance with an artificially manipulated pitch accent. The second experiment assessed the ability to distinguish between questions and affirmations in Dutch words, again by using artificial manipulation of F<sub>0</sub>. For the implanted group, performance significantly improved in each experiment when the hearing aid was added. However, even with a hearing aid, the implanted group required exaggerated F<sub>0</sub> excursions to perceive a pitch accent and to identify a question. These exaggerated excursions are close to the maximum excursions typically used by Dutch speakers. Nevertheless, the results of this study showed that compared to the implant only condition, bimodal fitting improved the perception of intonation.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3474236]

PACS number(s): 43.64.Me [RYL]

Pages: 1884–1895

## I. INTRODUCTION

Advances in implant technology and speech processing have significantly improved the performance of cochlear implants (CIs). Nevertheless, current CI systems are still largely unable to encode voice pitch information. Following the Continuous Interleaved Sampling (CIS) design principles (Wilson *et al.*, 1991), these systems provide temporal envelope information but discard much of the temporal fine structure required to perceive fundamental frequency (F<sub>0</sub>) (Zeng, 2004). In the CIS sound processing strategy, the pulse rate of the stimulus is fixed and therefore unrelated to F<sub>0</sub>. Earlier speech processing strategies conveyed voice pitch by the rate of pulsatile stimulation, however, the superior speech perception obtained with CIS made these strategies obsolete. Fixed-rate stimulation pulses are amplitude-modulated by low-pass filtered envelope signals extracted from a filter bank. In CIS processors, the lower cut-off frequency for the bandpass filter with the lowest center frequency just reaches the higher F<sub>0</sub> frequencies (Wilson *et al.*, 2005), so most of the temporal fine structure related to F<sub>0</sub> is discarded (Green *et al.*, 2005). Furthermore, for most subjects pitch perception with electri-

cal stimulation saturates at approximately 300 Hz (Shannon, 1983a). Exceptions have been noted, as summarized by Kong *et al.* (Kong, *et al.*, 2009). This saturation imposes an upper limit on the use of any remaining periodicity information and on the use of the envelope as an additional cue to intonation. On the basis of experiments carried out on normal hearing subjects listening to vocoder processed speech simulating the output of CIS, Green *et al.* found that the temporal envelope provided cues to F<sub>0</sub> perception, but only for an F<sub>0</sub> of up to approximately 200 Hz (Green *et al.*, 2002). Spectral cues for F<sub>0</sub> depend on a sufficient number of perceptually independent channels allowing place code for pitch. However, the number of channels in a CIS filterbank is small compared to estimates of the number of independent channels in the normal auditory system (Moore, 2003). Additionally it is further limited to about eight across the speech frequency range, probably due to intracochlear current spread limiting spectral resolution (Berenstein *et al.*, 2008; Friesen *et al.*, 2001; Green *et al.*, 2002; Henry and Turner, 2003; Shannon, 1983b). As a result, CI users show a threshold for spectral envelope differences insufficient to reliably resolve even the first harmonics of F<sub>0</sub> (Green *et al.*, 2002; Henry and Turner, 2003; Berenstein *et al.*, 2008).

Voice pitch information makes an important contribution to understanding speech. Previous studies suggest that voice

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: l.straatman@kno.umcn.nl

pitch helps a listener segregate speech in noise (Assmann and Summerfield, 1990; Brox and Nootboom, 1982). F0 variation is the physical correlate of intonation, which is a major component of prosody. The perception of pitch movements in non-tonal languages, such as English, Dutch, and German, enables a listener to follow the 'discourse' in terms of focus ('what is important') and to make distinctions between questions and affirmations. It also functions to identify speaker gender and speaker identity. These functions were found to be limited in subjects with a CI (Cleary *et al.*, 2005; Fu *et al.*, 2005; Green *et al.*, 2005). In tonal languages, such as Mandarin and Cantonese, lexical distinctions are based on pitch patterns, implying that the recognition of pitch patterns is essential for a listener to recognize at least a subset of the words in tonal languages. CI recipients are less able to accurately identify lexical tones and consequently lexical meaning, again due to their poor ability to extract voice pitch information (Ciocca *et al.*, 2002; Lee *et al.*, 2002; Peng *et al.*, 2004), even though a minority of subjects perform exceptionally well if natural speech stimuli are used, possibly by using extraneous cues such as duration (Peng *et al.*, 2004).

Several attempts have been made to improve pitch processing in CIs. Geurts and Wouters (Geurts and Wouters, 2004) devised a strategy in which the frequency of the first harmonic was extracted and used to steer the output of two discriminable electrodes in order to produce pitch percepts in between that of each electrode separately. When F0 temporal cues were eliminated from the stimuli, this place-pitch strategy allowed improved F0 detection. This feasibility study did not test speech perception.

Another approach adopted in several studies is to enhance periodicity cues to F0. In one study, the original waveform envelope was replaced with a 100% modulated sawtooth-like envelope on all channels (Green *et al.*, 2004, 2005). Although CI users improved slightly in their ability to identify intonation patterns, vowel perception was worse. In a follow-up study (Hamilton *et al.*, 2007) the enhanced envelope was applied to only one channel to limit the disruption of spectral information thought to have caused the decrement in speech perception found earlier. Contrary to expectations, pitch perception was poorer when the enhancement was applied to the most apical channel, however, for some of the subjects slight improvements were found when the most basal channel was used. Although no significant effects on vowel perception were found, the authors caution that the limited number of subjects precluded definite conclusions. Vandali and associates compared several ways to enhance temporal information (Vandali *et al.*, 2005). One strategy intended to code fine temporal detail in each channel by providing stimuli corresponding in time and amplitude to positive temporal peaks in the band-pass filtered signals. Three other strategies provided deeper F0 modulation cues than the clinical strategy, some of which with extra attention to align the modulation cues across all channels. The CI users better ranked the sung vowels in pitch with the experimental strategies than with the clinical strategy, while no decrement in speech perception was observed. Instead of a sawtooth envelope, Laneau *et al.* imposed a 100% modulated

F0-related sinusoidal envelope across all channels (Laneau *et al.*, 2006). Implant users showed improved musical interval perception with this strategy, but F0 discrimination of single-formant stimuli was the same as with the clinical strategy.

In conclusion, some experimental strategies enhancing F0 place or periodicity cues have shown improved aspects of pitch perception. Although some investigators note that these improvements come with a cost, namely a decrement in speech perception, others do not report such a decrement. This leaves hope for a clinically successful implementation of these experimental strategies. However, as long as the fundamental limits to place coding and rate coding referred to above are not alleviated, it is reasonable to expect only modest improvements in pitch perception.

A completely different approach to augmenting pitch perception is to combine electrical and acoustic stimulation for subjects with some residual hearing, either with a hearing aid (HA) in the same ear (referred to as Electrical Acoustic Stimulation or "EAS" (von Ilberg *et al.*, 1999; Turner *et al.*, 2004)) or a hearing aid in the non-implanted ear, generally referred to as "bimodal fitting" (Ching *et al.*, 2004; Kong *et al.*, 2005; Kong and Carlyon, 2007). It is assumed that the auditory system is able to combine the low frequency input from the HA with the high frequency information from the CI, even when stimuli are presented in different ears as in bimodal fitting. Acoustic stimulation does not have the same limitations encountered with currently available cochlear implant systems, namely large intra-cochlear channel interactions and temporal saturation of the electrically stimulated nerve. This is not to say that F0 perception is normal once acoustic stimulation is provided. Cochlear hearing loss may cause a decrement in temporal and spectral resolution which alters the processing of voice pitch (Moore and Carlyon, 2005). As a result, a significant increase of the F0 difference limen measured with complex harmonic tones has been found in young moderately hearing-impaired subjects (Moore and Peters, 1992).

In normal hearing, F0 detection primarily depends on resolved lower harmonics. However, for severely hearing-impaired listeners these harmonics may be unresolved, meaning that these listeners may have to rely more on amplitude modulations that occur at a rate equal to the F0, caused by the beating between the unresolved higher harmonics within each analysis filter (Oxenham, 2008).

Studies have shown that speech recognition in noise improves when a CI is combined with a contralateral hearing aid (Ching *et al.*, 2004; Kong *et al.*, 2005; Dorman *et al.*, 2008). Kong *et al.* (2005) investigated the effect of bimodal fitting on listener recognition of English sentences in the presence of a competing talker. A significant bimodal advantage was observed, even when acoustic stimulation alone did not lead to speech recognition (Kong *et al.*, 2005). One hypothesis for this advantage is that temporal F0 cues delivered by the hearing aid enable improved perceptual segregation of speech from competing noises (Oxenham, 2008). However, both the role of F0 periodicity and the segregation account of EAS, and the benefit of bimodal fitting are under investigation. Kong and Carlyon (2007) used vocoder simulations to obtain data on subjects with normal hearing and found an

improvement in simulated bimodal speech recognition in noise even after the F0 cue had been removed from the low-pass filtered speech presented in the other ear. This was taken as evidence that voicing or spectral cues to the first formant may offer “glimpses” into important parts of the running speech signal (Assmann and Summerfield, 1990). Moreover, it was concluded that F0 is neither necessary nor sufficient to explain the improved speech recognition in noise. In contrast, Brown and Bacon (2009) demonstrated that, under vocoded simulated conditions, both F0 and the low-frequency amplitude envelope are useful cues for implant users. Similarly, significant benefit was found when the output of the vocoder was combined with a tone carrying F0 cues, compared to vocoder alone (Brown and Bacon, 2009). This benefit remained, even with a competing tone following F0 of a background speaker, which was interpreted as evidence against segregation and in favor of the “glimpsing” account of EAS and bimodal fitting benefit.

The benefits of bimodal fitting have also been reported for children with CIs. Holt *et al.* showed improved speech perception scores in the bimodal condition compared to CI- and HA-only, particularly for speech understanding in noise with a +5 dB signal-to-noise ratio (Hearing-In-Noise Test Children’s Version) (Holt *et al.*, 2005). Similarly, Lee *et al.* reported improved speech performance in noise in the bimodal condition compared to the CI-only condition (Lee *et al.*, 2008). Obviously, tone-languages heavily depend on the correct reception of F0 information. Luo and Fu showed that in CI vocoder simulations with Chinese adult listeners, low-frequency acoustic information below 500 Hz significantly improved lexical tone recognition in speech-shaped noise (Luo and Fu, 2006).

Until now, it has not been investigated if low-frequency acoustic information also improves the perception of intonation in bimodally fitted children. In the present study, results from two experiments are provided to clarify how additional acoustic information affects F0 detection in bimodally fitted children. Experiment 1 used a two-interval same/different task adapted from O’Halpin *et al.* (2005) to establish the just detectable pitch movement for a bisyllabic nonsense word in which F0 was manipulated. In an attempt to include meaningful material in a more or less ecologically valid task, experiment 2 assessed the children’s ability to distinguish between questions and affirmations in acoustically manipulated simple Dutch utterances. Our hypothesis was that in the bimodal condition (CI plus hearing aid, “CI+HA”) the children would detect smaller pitch movements than in the CI alone condition. For reference purposes, normative data were obtained from a group of children with normal hearing.

## II. MATERIALS AND METHODS

### A. Subjects

Children with normal hearing (NH) and implanted children were recruited for the study. The two groups were matched for age. All participants were Dutch native speakers except for two CI users. Although Dutch was not considered their first language, they both had sufficient command of the language and could understand the task. Approval for the

study protocol was received from the Ethical Committee of the Radboud University Nijmegen Medical Centre (number 2007/090). Written informed consent was obtained from the subjects or from their parents if the child was under 16 years of age.

### 1. Subjects with normal hearing

To obtain reference values and to establish whether children in the age range of our test group were able to perform the task, 14 NH children aged between 6.8 and 16.7 years (mean age 10.1 years) were recruited among colleagues and friends of the researchers. Normal hearing was defined as the presence of otoacoustic emissions or, in the absence of otoacoustic emissions, as pure-tone air conduction thresholds of  $\leq 15$  dB HL at octave frequencies from 125 to 8000 Hz.

### 2. Subjects using a cochlear implant and a hearing aid

The experimental group comprised 19 implanted children whose age at implantation was below 16 years and whose age at testing ranged from 6.0 to 19.8 years (mean 12.1 years). All the children had been using their implant for more than one year. Experience with bimodal fitting varied from 1.0 to 8.7 years. Demographic data on each subject and details about the implant and hearing aid use are shown in Tables I and II. A hearing aid had been fitted to the non-implanted ear of each child. The free-field unaided and aided hearing thresholds in the implant and the non-implanted ear are also given in Tables I and II. Fitting of both CI and HA was evaluated by the same audiologist. Word recognition based on the number of correct phonemes identified by a listener was tested using the “Nederlandse Vereniging voor Audiologie” (Dutch Society of Audiology) open-set monosyllable word recognition test (Bosman, 1989). The mean percentage of correct phonemes in the three conditions was: CI-only 69% (SD 23%), equivalent to a word score of 47%, HA-only 24% (SD 25%), equivalent to a word score of 10%, and CI+HA 75% (SD 19%), equivalent to a word score of 56%.

### B. General procedure of Experiments 1 and 2

Each child was tested in a sound-treated double-walled room. At the beginning of each experiment, the task was explained and examples of the stimuli were presented. Stimuli were presented via a loudspeaker placed one meter in front of the subject. The presentation level was chosen to be comfortably loud. It was checked if this level exceeded the infinite compression threshold of the child’s speech processor.

We selected stimulus with the highest RMS level and looped the stationary part of the vowel that received either sentence accent, word stress, or both. The loudest stimuli were found to be 68 dB SPL as measured with a Bruel & Kjaer Investigator 2260 sound level meter held at the level of the subject’s ear. Assuming an dynamic input range of 30 dB for the older processors and 40 dB for the newer processors (Davidson *et al.*, 2009), and given the aided thresholds in

TABLE I. Demographic data and details on the implants and unaided and aided thresholds of the implanted ear.

Subject	Gender	Age (years)	Age at implantation (years)	Ear implanted	Implant	Strategy	Unaided threshold implanted ear (dB SPL)					Aided threshold with CI (dB SPL)				
							0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz
1	Female	6.0	4.0	Right	Cochlear N24 rca	ACE	110	110	120	110	120	30	35	35	30	35
2	Male	6.1	5.0	Left	Cochlear Freedom ca	ACE	100	100	110	120	120	35	35	30	40	50
3	Female	6.2	1.5	Right	Cochlear N24 rcs	ACE	120	120	120	120	120	40	35	40	35	45
4	Female	7.3	6.3	Left	Cochlear Freedom ca	ACE	Unknown	35	90	110	110	30	30	35	35	35
5	Female	7.3	3.3	Left	Cochlear N24 rcs	ACE	110	120	120	120	120	30	25	30	30	35
6	Female	8.6	7.5	Left	Cochlear Freedom ca	ACE	90	90	100	110	110	20	15	20	15	25
7	Female	10.2	9.2	Right	Cochlear Freedom ca	ACE	90	95	105	105	100	30	30	30	30	30
8	Female	10.2	8.1	Right	Cochlear N24 rca	ACE	80	85	100	95	80	30	35	35	35	40
9	Female	10.6	4.6	Right	Cochlear N24 <sup>a</sup>	ACE	95	105	110	110	110	25	25	45	35	40
10	Male	11.2	10	Left	Cochlear Freedom ca	ACE	95	95	110	115	120	25	25	25	25	30
11	Female	11.3	6.8	Right	Cochlear N24 rcs	ACE	40	100	110	120	110	45	45	40	40	40
12	Female	14.7	11.3	Left	Cochlear N24 rcs	ACE	100	100	115	120	120	50	45	55	50	50
13	Male	14.8	6.9	Right	Cochlear N24 <sup>b</sup>	ACE	65	70	90	115	110	25	25	30	25	30
14	Male	15.3	11.3	Right	Cochlear N24 rcs	ACE	85	110	115	120	120	35	50	45	45	50
15	Male	15.9	12.3	Right	Cochlear N24 rcs	ACE	120	120	120	120	120	30	30	20	20	30
16	Male	16.7	11.0	Right	Cochlear N24 rcs	ACE	85	105	115	115	110	50	55	50	45	50
17	Female	18.7	14.3	Left	Adv. Bionics C2HiFo2	SAS	100	105	115	120	120	50	40	45	45	40
18	Female	18.8	10.0	Right	Cochlear N24 <sup>b</sup>	ACE	90	95	110	120	120	15	25	25	50	25
19	Female	19.8	15.5	Right	Cochlear N24 rcs	ACE	85	100	105	120	120	30	25	30	30	30

<sup>a</sup>N24 straight.

Table I, it follows that in nine subjects the loudest stimulus may have reached the threshold for infinite compression.

All the children completed experiments 1 and 2 and experiment order was randomized across subjects. In addition to random experiment presentation, the order in which the CI and CI+HA conditions were tested was randomized in each experiment.

### C. Experiment 1: Just noticeable difference of fundamental frequency (F0) of the non-word “baba”

#### 1. Speech materials

The stimuli consisted of the bisyllabic nonsense-word “baba” ([ˈbaːbaː]) manipulated according to O’Halpin *et al.* (O’Halpin *et al.*, 2005). Two Dutch native speakers (a male aged 26 years and a female aged 25 years) produced the

TABLE II. Demographic data and details on the hearing aids and unaided and aided thresholds of the non-implanted ear fitted with a hearing aid.

Subject	Provider of hearing aid	Duration of hearing aid use (years)	Unaided thresholds the non-implanted ear (dB SPL)					Aided thresholds with hearing aid (dB SPL)				
			0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	0.25 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz
1	Oticon sumo xp	3.7	110	100	110	120	120	Unknown	45	50	65	85
2	Phonak supero 412	3.8	85	85	95	95	120	80	60	60	60	85
3	Phonak pp-c-l-p	4.8	90	90	95	120	120	60	40	30	25	55
4	Phonak maxx211	2.3	Unknown	30	100	110	110	30	35	45	50	80
5	Phonak sf pp-c-l-4+	5.7	110	100	95	90	80	45	40	35	35	45
6	Phonak maxx411	6.4	Unknown	100	90	85	65	50	50	45	45	50
7	Widex p38	7.1	90	95	100	100	100	30	40	35	65	65
8	Phonak supero 411	6.8	85	85	100	95	80	35	40	50	55	75
9	Oticon digifocus 2	9.9	85	100	110	110	110	75	60	60	55	80
10	Phonak supero 412	10.1	90	95	105	100	90	40	40	55	45	65
11	Phonak pp-c-l-p2	7.4	50	95	100	100	95	70	65	65	65	75
12	Phonak picoforte pp-c-p	13.1	90	90	100	90	120	60	55	50	55	80
13	Phonak supero 412	13.0	60	75	90	110	120	60	50	60	65	100
14	Siemens phoenix 313	14.5	90	110	110	115	120	50	55	50	65	95
15	Resound viking	14.4	75	100	105	105	90	40	40	50	60	80
16	Phonak pp-c-l-p	14.2	90	95	110	115	110	65	70	70	75	75
17	Philips s460	18.0	100	95	105	120	120	50	55	55	75	90
18	Siemens phoenix 313	14.8	85	95	105	120	115	35	35	40	75	90
19	Phonak superfront pp-c-l	17.8	65	85	115	115	115	15	20	30	50	70

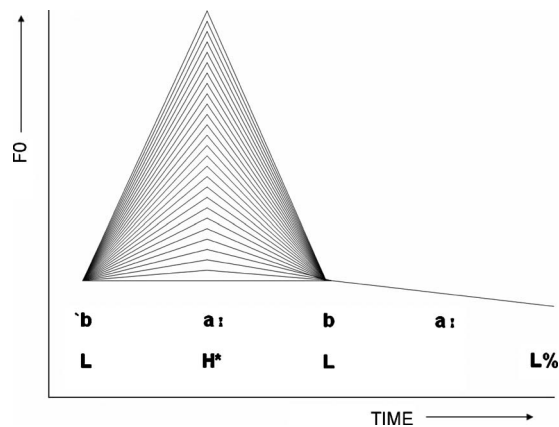


FIG. 1. Examples of manipulated F0 contours of the ‘baba’-bisyllables, recorded from native Dutch speaker (male). The black bold line represents the unstressed or base stimuli. The gray lines represent the different contours of the peak of the accented stimuli.

word 15 times, each time attempting a flat intonation. For each speaker, recordings were selected that sounded the most monotone. From this preselection, we selected recordings with only minimal differences in length between the first and second syllable to minimise word stress cues and sentence accent. The recordings were digitized at a sampling frequency of 44 kHz (16 bit resolution). Pitch contours of these bisyllabic words were manipulated using the Pitch Synchronous Overlap and Add (PSOLA) technique implemented in the speech processing software PRAAT (Boersma and Weenink, 2006). PSOLA provides high quality manipulated speech signals (Moulines and Charpentier, 1990). The onset F0 values of the first syllable produced by the male and the female speaker were set at 100 Hz and 200 Hz respectively. To replicate the F0 declination in natural speech, a linear F0 fall of 2.8 semitones was added to the second syllable. This is in accordance with the rules incorporated into a Text-To-Speech-system for Dutch that are based on typical Dutch pitch movements (Gussenhoven *et al.*, 1999; Gussenhoven and Rietveld, 2000; Kerkhoff and Rietveld, 1995; Hart *et al.*, 1990).

A pitch accent was induced in the first syllable by creating a peak in the F0 contour at the mid-point of the vowel (known as H\* in the “autosegmental description of intonation” (Gussenhoven *et al.*, 1999). From the F0 at onset (“L”), the pitch contour rose linearly to the midpoint (“H\*,” followed by a linear fall to the end of the first syllable (“L”). The difference in F0 at onset and at H\* ranged from 0.85 semitones to 22.1 semitones (almost two octaves), with a step size of 0.85 semitones (Fig. 1). The duration of the first syllable as produced by both speakers was approximately 200 ms. To create the pitch accent, the first 200 ms of the stimuli were manipulated. Total durations of the male utterance and the female utterance were 540 ms and 590 ms, respectively. Introducing pitch movements may have altered the loudness of the stimuli and thus introduced a cue into the discrimination task. To reduce this possibility, the peak amplitude of each manipulated utterance was scaled to the same value (95% of the maximum amplitude of the sound buffer of the computer system).

## 2. Procedure

An adaptive two-alternative forced-choice same/different task was used. In each trial, two “baba” non-words were presented, separated by a silent interval of 500 ms. In half of the trials, only one of the words (randomly chosen) contained an accent, while in the other half of the trials neither of the words had a pitch accent. Each test was run twice in a random order across the subjects: once with the male version and once with the female version. Subjects were asked to press the left or right arrow button on the keyboard that had been labeled with the word “same” and a drawing of two identical elephants (left) or “different” and a drawing of two dissimilar elephants (right). Visual feedback was provided. At the beginning of each test, a practice run of 16 trials was presented. All the implanted children used their cochlear implant and their hearing aid during the practice runs. Both devices were set at the normal everyday setting.

Thresholds were obtained using a 2-down 1-up staircase procedure (Levitt, 1971). After an incorrect response, the height of the peak of the accented stimulus was increased by one-step. After two correct responses, the height was decreased by one-step. A test run proceeded until one of the following occurred: (1) ten reversals, (2) eight successive incorrect responses at the maximum stimulus difference, or (3) eight successive correct responses at the minimum stimulus difference. The F0 threshold was estimated from the mean of the F0 differences of the final six reversals.

## D. Experiment 2: Discrimination between questions and affirmations

Stimuli were six isolated bisyllabic or tri-syllabic words, all names of fruit, spoken with an affirmative intonation by a Dutch native speaker (a man aged 57 years) at an average F0 of 130 Hz. Interrogative versions of these words were created by imposing F0 patterns typical of questions in Dutch (and ‘exaggerated’ versions of them) on the last syllable of the utterance. None of the other markers of questions vs. affirmations were changed (e.g., duration, spectral pattern, overall F0 level).

### 1. Speech materials

To create the stimuli, three words were selected that carry stress on the first syllable: “aardbei” ([ˈaːrtbɛi];, strawberry), “paprika” ([ˈpaprikaː];, pepper) and “mango” ([ˈmɑŋɡoː];, mango), and three words were selected that carry stress on the final syllable: “citroen” ([siːˈtrun];, lemon), “banaan” ([baːˈnaːn];, banana) and “meloen” ([mɛːˈluːn];, melon). As described above, the pitch contours were manipulated using the PSOLA technique in PRAAT. All the original productions were provided with a linear falling pitch contour (“LL%”) of 6 semitones that is typical of Dutch utterances, which yielded the affirmative stimulus. While attempting to provide the affirmative stimulus with a rising intonation typical for questions (“LHH%”), it was found that the naturalness of the stimuli was best preserved when only the last part of the F0 contour was manipulated. Rules developed in the Text-To-Speech-system for Dutch, mentioned above, were adopted. They can be summarized as follows: F0 rise and fall

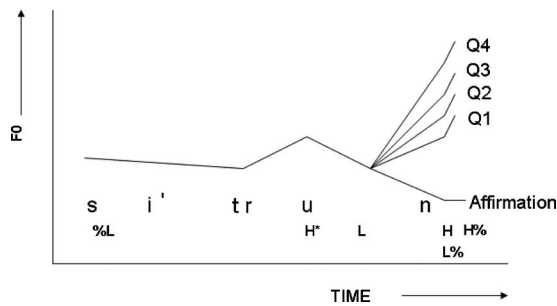


FIG. 2. Diagram of the final fall (L-L%) and rises (L-H-H%) of the F0 contour in the affirmation and in the four question variants (Q1, Q2, Q3 and Q4) of H-H%. From positions %L to L, the default Dutch F0 pattern was used, ending at H% or L%, with the lowest F0 in the affirmative variant. F0 at H% (male speaker) varied between 100 Hz (affirmation) and 183 Hz (Q4, see Table III). The stimulus was the word “citroen” ([si'trun]; lemon).

have a duration of 100 ms and the final falling movement (to L%) and rising movement (peak at H%) have a duration of 20 ms; and the starting locations were determined from the end of the words. In compliance with the rules of the Text-To-Speech-system, the start of the rise or fall for words ending with an [n] and for those with a low amplitude in the last 120 ms was moved to the beginning of the word. This is because it is known to be more difficult to perceive the LHH% pitch rise when the amplitude is low or during a nasal resonance.

The total duration of the stimuli varied from 440 ms to 580 ms. Four questions were created (Q1, Q2, Q3, and Q4), each with increasingly higher F0 rises at the end of the utterances. Figure 2 shows an example of the pitch contours of the Affirmation and the four question variants. Table III shows the increases in F0 as expressed in semitones and in percent. The step size between the four categories of questions was three semitones. In Dutch, LHH% rises that signal questions range from 6 to 9 semitones, which is comparable to Q1. The peak amplitude of each manipulated utterance was scaled to the same value (95% of the maximum amplitude of the sound buffer of the computer system).

## 2. Procedure

The stimuli were presented in blocks of 48 items containing 24 affirmations (6 tokens repeated 4 times) and 24 questions (6 tokens from each of the Q1, Q2, Q3 and Q4 categories). All six words were used equally often as an affirmation and as a question. Each word was also used equally

TABLE III. Final rise of the F0 contours in the four question categories, expressed in semitones and percentage difference from L (the start of the rising contour, see Fig. 2) to H and H% (the end of the rising contour). In Dutch, final rises that signal questions range from 6 to 9 semitones. Thus, the rising contour of Q1 is typical for Dutch questions.

End-pitch contour	L to H		L to H%	
	Semitones	%difference	Semitones	%difference
Q1	6	41	9	68
Q2	9	68	12	100
Q3	12	100	15	138
Q4	15	138	18	183

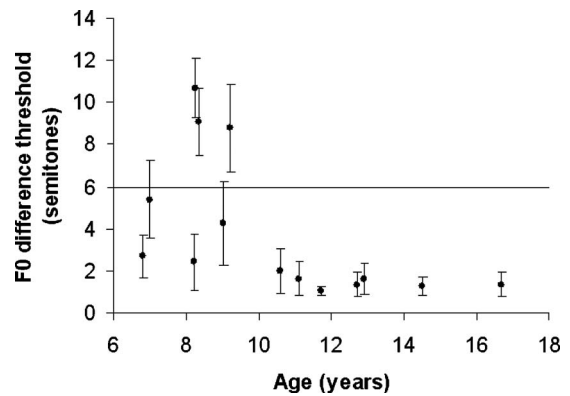


FIG. 3. Individual thresholds of F0 movements in the 14 children with normal hearing as a function of age at testing. Lower thresholds indicate better F0 discrimination. The horizontal line gives an indication of the F0 rise characteristic of the height of accents in Dutch. Data pooled over the male and female speakers. Error bars represent the measurement error of the adaptive threshold procedure (i.e., the mean SD of the final 6 reversals of the adaptive procedure).

often in each of the categories Q1-Q4. Each block of 48 stimuli was presented twice. The stimuli were randomly selected without replacement.

In each trial, the subjects heard a single word and were required to identify it as either a “question” or an “affirmation” by pointing at a drawing of a person who was gesturing a question, or a person who was making an affirmative gesture. Visual feedback was provided. Each test started with a practice run of 12 different stimuli. All the implanted children used their cochlear implant and their hearing aid during the practice runs. Both devices were set at the normal everyday setting.

## III. RESULTS

### A. Experiment 1: Just noticeable difference in fundamental frequency (F0) of the non-word “baba”

#### 1. Subjects with normal hearing

Figure 3 shows the results of the NH subjects pooled over the male and female speaker. The mean F0 differences at threshold were 3.3 semitones (21% above baseline) with an SD of 2.7 semitones for the male speaker and 4.3 semitones (28% above baseline) with an SD of 4.2 semitones for the female speaker. A paired-samples t-test (SPSS version 13.0) did not show a significant difference in the responses to the male and female speakers ( $t_{13}=2.7$ ,  $p>0.12$ ;  $\eta^2_{\text{partial}}=0.172$ ;<sup>1</sup> observed power=0.332). The F0 difference at threshold, expressed in semitones, decreased with the logarithm of the age of the subject ( $F_{1,13}=11.7$ ;  $p<0.01$ ,  $R^2=0.493$ ). Figure 3 indicates that the performance of children older than 10 years was better than that of most of the younger children. This suggests that the test may have been too difficult for some of the younger children. The average F0 difference at threshold in the children aged 10 years and older was 1.5 semitones (SD=0.4).

#### 2. Subjects using a cochlear implant and a hearing aid

Figure 4 shows the mean F0 difference at threshold in the CI-only and CI+HA conditions compared to the NH

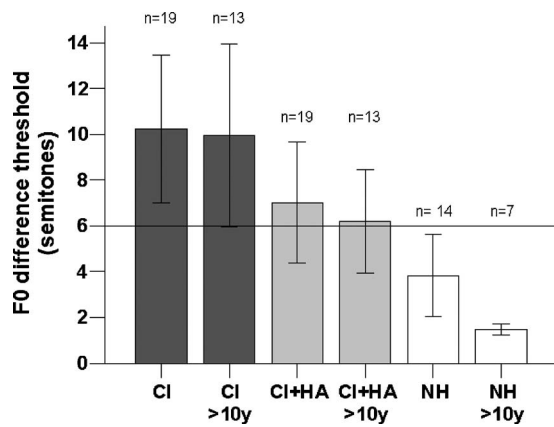


FIG. 4. Mean F0 difference thresholds and 95% confidence interval in the CI alone and the CI plus HA conditions compared to the group with normal hearing. Data from the whole group ( $n=19$  in the implanted group and  $n=14$  in the NH group) and from individuals older than 10 years ( $n=13$  in the implanted group and  $n=7$  in the NH group) are presented separately. Data pooled over the male and female speakers. Lower thresholds indicate better F0 discrimination. The horizontal line gives an indication of the F0 rise characteristic of word accents in Dutch.

group. In response to the male speaker, thresholds improved on average from 9.4 semitones (CI) to 6.9 semitones (CI+HA) (i.e., a decrease from 72% to 49% F0 difference). In response to the female speaker, thresholds improved from 11.3 (CI) to 7.5 (CI+HA) semitones (i.e., a decrease from 92% to 54%). Independent samples t-test showed a significant improvement in performance for the NH group compared to the implanted children in the CI-only condition ( $t_{31}=3.2$ ;  $p<0.01$ ;  $\eta^2_{\text{partial}}=0.246$ ). However, the scores of the NH group were not significantly better than the implanted children wearing their CI+HA ( $t_{31}=1.9$ ;  $p>0.05$ ;  $\eta^2_{\text{partial}}=0.101$ ; observed power=0.439). Figure 5 shows the individual F0 difference thresholds in the CI-only condition compared to the CI+HA condition. Most data points were on

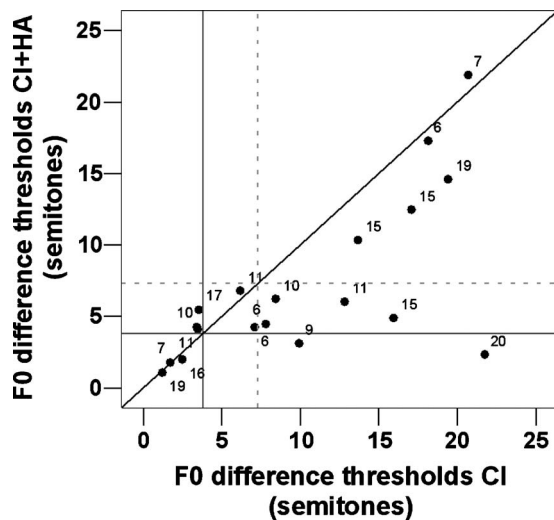


FIG. 5. Individual thresholds of the F0 movements in the CI alone compared to the CI plus HA conditions. Lower thresholds indicate better F0 discrimination. The black diagonal line shows complete agreement. Horizontal and vertical lines represent the mean thresholds (gray line) plus 1 SD (dotted line) of the group with normal hearing. Data points are labeled with the age (years) of the subject. Data pooled over the male and female speakers.

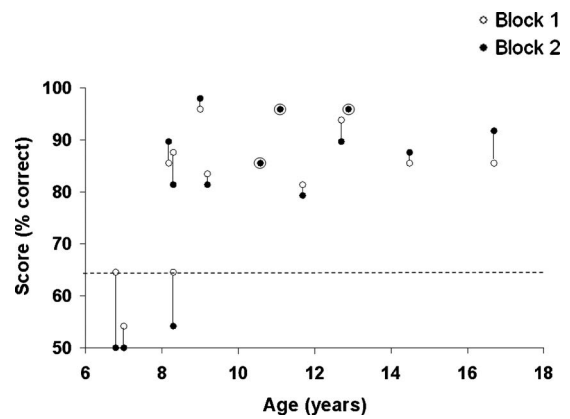


FIG. 6. Percentage correct scores on the question/affirmation identification task as a function of age in the 14 subjects with normal hearing. Open circles represent the first block scores, black points represent the second block scores. Individual scores are connected with an arrow. The arrow indicates the direction of change in performance between the first and the second block. Three subjects had identical scores on the first and second blocks (represented by black point in open circle). The horizontal dotted line represents chance level, based on binomial tests. Higher scores indicate better question/affirmation identification.

or under the diagonal line. Figure 5 shows that in the CI-only condition, the performance of 42% of the implanted subjects (8 out of the 19) was within 1 SD of the average of NH subjects. In the CI+HA condition, this percentage improved to 74% (14 out of 19) and all but three CI subjects performed within 2 SD from the average of NH subjects. A repeated measures analysis of variance with two within-subject fixed factors, namely Condition (CI-only and CI+HA) and Speaker (male and female), showed significantly better F0 discrimination in the CI+HA condition than in the CI condition ( $F_{1,18}=6.2$ ;  $p<0.02$ ;  $\eta^2_{\text{partial}}=0.26$ ; observed power = 0.65).<sup>2</sup> The effects of Speaker ( $F_{1,18}=1.4$ ;  $p>0.24$ ;  $\eta^2_{\text{partial}}=0.07$ ; observed power=0.21) and the interactions between Speaker and Condition ( $F_{1,18}=0.55$ ;  $p>0.46$ ;  $\eta^2_{\text{partial}}=0.03$ ; observed power=0.11) were not significant.

Linear correlations (Pearson), nonparametric correlations (Spearman & Kendall) and curvilinear correlations were calculated between the F0 difference thresholds and (1) age, (2) logarithmic function of age, (3) the pure tone average (PTA, i.e., mean threshold at 0.5, 1, 2 and 4 kHz), (4) average aided threshold with the hearing aid, (5) average aided threshold with the CI, (6) aided thresholds at 0.25 kHz, (7) duration of deafness, (8) duration of CI use, (9) duration of hearing aid use, (10) age at implantation, (11) age of deafness, (12) speech perception scores in quiet (i.e., phoneme or word scores) in HA-only, CI-only, and the bimodal condition. Of these, only one nonparametric correlation was found to be significant. F0 difference thresholds were marginally better for patients that were older at implantation (Kendall's tau:  $-0.283$ ,  $p=0.046$ ).

## B. Experiment 2: Discrimination between questions and affirmations

### 1. Subjects with normal hearing

All the normal hearing subjects participated in two blocks of 48 trials. The mean correct score was 81% (SD 15%). Their individual scores are shown in Fig. 6. A signifi-



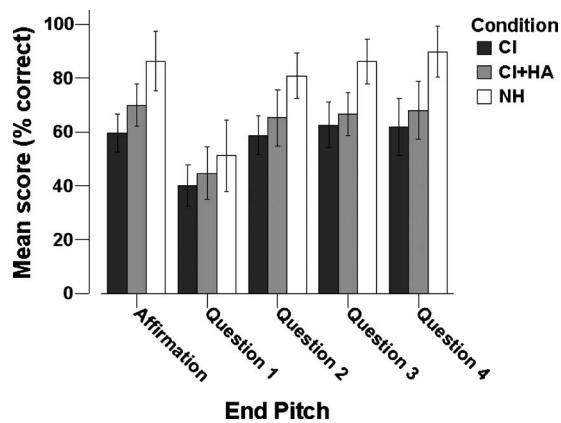


FIG. 7. Mean percentages of correctly classified utterances and 95% confidence interval of the Affirmation and the modified utterances Question 1, Question 2, Question 3 and Question 4 in the CI alone condition (dark gray bar) and the CI plus HA condition (light gray bar), compared to group with normal hearing (white bar). Higher scores indicate better question/affirmation identification. For details about the stimuli see Table III.

cant Pearson correlation coefficient was found between age and intonation discrimination ( $r_{14}=0.58$ ;  $p<0.05$ , one-tailed). However, it was clear that this correlation was mainly caused by the poor scores of the three youngest children ( $\leq 8.3$  years). Based on the binomial distribution, each child's performance was statistically compared to chance levels, with a criterion level of 64% correct ( $\alpha=0.05$ , one-sided). Performance was significantly better than chance for 11 subjects but not for the three youngest subjects (aged 6, 7 and 8 years).

Figure 7 shows the mean correct scores on the five categories of the manipulated utterances. As was expected, the higher the end-pitch rise, the higher the score. As the implanted group had missing values in experiment 2, a linear mixed model analysis was carried out with two within-subject fixed factors: End-pitch contour (affirmation, Q1, Q2, Q3 and Q4) and Block (first and second blocks). Utterances with higher end-pitch rises were more often correctly classified as questions ( $F_{4,117}=26.2$ ;  $p<0.001$ ). In contrast, the effect of Block ( $F_{1,117}=1.8$ ;  $p=0.158$ ) and the interaction between Block and End-pitch contour ( $F_{4,117}=0.39$ ;  $p>0.81$ ) were not significant. Pairwise comparisons of end-pitch contour using Bonferroni adjustment for multiple comparisons showed one significant contrast: Q1 yielded significantly fewer correct responses than all other categories ( $p<0.001$  in each case).

## 2. Subjects using a cochlear implant and a hearing aid

An experienced pediatric audiologist judged that 2 out of the 19 children (both 6 years of age) lacked concentration during the practice run. The results of these two subjects were therefore excluded from analysis. Due to lack of time or concentration, 9 of the 17 remaining subjects were unable to participate in the second block of the test. A bimodal benefit of 62% (CI+HA) against 56% (CI-only) was found for the 17 subjects who completed the first block, and 73% (CI+HA) against 62% (CI-only) for the repeated presentation (8 subjects). The higher scores in the second block sug-

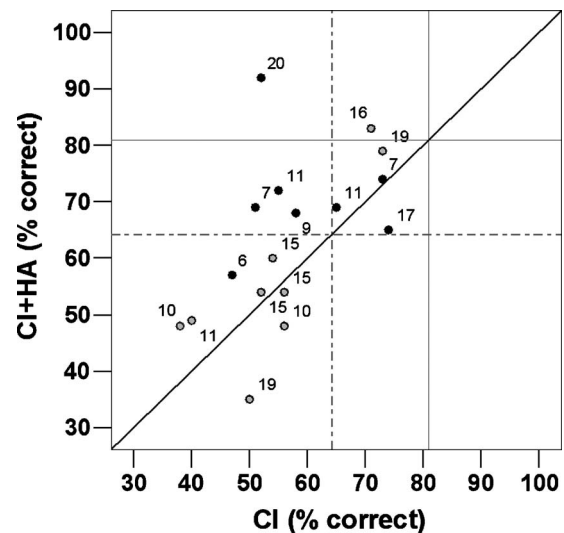


FIG. 8. Individual percentage correct question/affirmation identification scores in the CI alone condition compared to the CI plus HA condition. Higher scores indicate better question/affirmation identification. The diagonal line represents no effect of condition. Mean identification scores (gray horizontal/vertical lines) minus 1 SD (dotted horizontal/vertical lines) in the group with normal hearing. Gray points are the scores of the subjects who completed one block, black points are the scores of eight individuals who completed two blocks. All data points are labeled with the age of the subject (years).

gested a learning effect. In Fig. 8, the individual scores obtained in the CI-only condition are compared to the scores in the CI+HA condition. On the basis of the binomial distribution, performance was above chance for 5 of the 17 subjects (29%) in the CI-only condition and for 9 of the 17 subjects (53%) in the CI+HA condition (criterion level is 64% correct,  $\alpha=0.05$ , one-sided).

A linear mixed model analysis was carried out with three fixed within-subject factors: Condition (CI and CI+HA), Block (first and second blocks) and End-pitch contour (Affirmation, Q1, Q2, Q3 and Q4). Significantly better scores were found in the CI+HA condition compared to the CI-only condition ( $F_{1,213.33}=6.5$ ;  $p<0.02$ ). The effect of End-pitch contour were also significant ( $F_{4,213.33}=12.3$ ;  $p<0.001$ ). In contrast, the effects of Block and the interactions between Condition and Block ( $F_{1,213.33}=0.313$ ;  $p>0.57$ ), Condition and End-pitch ( $F_{4,213.33}=0.517$ ;  $p>0.72$ ), Block and End-pitch ( $F_{4,213.33}=0.838$ ;  $p>0.50$ ) and the three-way interaction between Condition, Block and End-pitch ( $F_{4,213.33}=1.425$ ;  $p>0.23$ ) were not significant.

When the 9 subjects who did not complete both blocks were excluded, approximately the same results were found. Again, significantly better scores were found in the CI+HA condition ( $F_{1,133}=7.4$ ;  $p<0.01$ ). The effect of End-pitch contour was also significant ( $F_{4,133}=11.8$ ;  $p<0.001$ ). As before, the effect of Block and the interactions between Condition and Block ( $F_{1,133}=0.004$ ;  $p>0.95$ ), Condition and End-pitch ( $F_{4,133}=0.434$ ;  $p>0.78$ ), Block and End-pitch ( $F_{4,133}=0.526$ ;  $p>0.72$ ) and the three-way interaction between Condition, Block and End-pitch ( $F_{4,133}=1.40$ ;  $p>0.24$ ) were not significant.

Additionally, an independent sample t-test comparing the 8 subjects who completed both blocks and the 9 subjects

who did not, showed no significant differences with respect to: (1) age, (2) age at implantation, (3) age at deafness, and (4) question-statement identification scores of the first block. As a result of the similar scores obtained in the group which completed both blocks and the results of the entire group, the results were pooled for further analysis.

Linear correlations (Pearson), nonparametric correlations (Spearman & Kendall) and curvilinear correlations were calculated for the whole group between intonation identification and (1) age, (2) logarithmic function of age, (3) the pure tone average (PTA, i.e., mean threshold at 0.5, 1, 2 and 4 kHz), (4) average aided threshold with the hearing aid, (5) average aided threshold with the CI, (6) aided thresholds at 0.25 kHz, (7) duration of deafness, (8) duration of CI use, (9) duration of hearing aid use, (10) age at implantation, (11) age of deafness, (12) speech perception scores in quiet (i.e., phoneme or word scores) in HA-only, CI-only, and the bimodal condition. No significant correlations were found.

Pairwise comparisons of the end-pitch contour on the basis of Bonferroni adjustment for multiple comparisons showed that significantly fewer Q1 utterances were correctly identified than Affirmations ( $t_{4,213.33}=5.8$ ;  $p<0.001$ ), Q2 ( $t_{4,213.33}=5.2$ ;  $p<0.001$ ), Q3 ( $t_{4,213.33}=5.5$ ;  $p<0.001$ ) and Q4 utterances ( $t_{4,213.33}=5.5$ ;  $p<0.001$ ). This effect was significant in both the CI-only and the CI+HA conditions (Fig. 7).

In the CI-only condition, only 24% of the implanted children (4 of the 17) scored within 1 SD of the mean of the NH children. In the CI+HA condition, this percentage increased to 47% (8 out of the 17) and all but 3 implanted children performed within 2 SD of the mean of the NH children. Independent samples t-tests showed a significantly better performance by the NH group compared to the implanted children in the CI-only condition ( $t_{29}=5.4$ ,  $p<0.001$ ;  $\eta^2_{\text{partial}}=0.50$ ), and the CI+HA condition ( $t_{29}=3.5$ ,  $p<0.01$ ;  $\eta^2_{\text{partial}}=0.29$ ).

### C. Correlation between F0 perception and discrimination of questions and affirmations

A one-tailed Pearson correlation test showed that subjects with better F0 thresholds in Experiment 1 were also better in classifying questions and affirmations in Experiment 2 (condition CI-only:  $r_{17}=-0.45$ ;  $p<0.05$ , condition CI+HA:  $r_{17}=-0.44$ ;  $p<0.05$ ). There was also a significant correlation between the bimodal benefits (i.e., the difference between CI score and CI+HA score) in both experiments ( $r_{17}=0.53$ ;  $p<0.05$ ).

## IV. DISCUSSION

In the present investigation, bimodal fitting significantly supported the perception of intonation compared to using CI-only. Both the discrimination of F0 excursions signaling pitch accents and the ability to use F0 patterns to identify questions versus affirmations improved significantly. In both tasks, an F0 excursion of about 6 semitones was required to be detectable in the CI+HA condition, which was similar to the excursion needed by the poorer performing NH control subjects. The performance in the CI+HA condition was suf-

ficient to detect only the largest F0 excursions normally produced by Dutch speakers. It is reasonable to expect decreased performance in less than optimal listening situations when reverberation or (low frequency) noise may mask F0 (Lavandier and Culling, 2008).

### A. Just noticeable difference in fundamental frequency

The first experiment assessed the discrimination of fundamental frequency excursions used to signal a pitch accent. The addition of the hearing aid was of significant benefit to the implanted children. The JND improved from 9.4 (CI) to 6.9 semitones (CI+HA) in the test with the male speaker, versus 11.3 (CI) to 7.5 semitones (CI+HA) in the test with the female speaker. The effect of Speaker was not significant, but it was in the same direction as that reported by others (Green *et al.*, 2004; Laneau and Wouters, 2004; Chatterjee and Peng, 2008). Discrimination thresholds of the NH children (male speaker: 3.3 semitones, female speaker: 4.3 semitones) were almost half those of the CI subjects. In other words, the overall performance of the implanted children was poor.

Several studies on cochlear implant users reported JNDs for F0 changes that were considerably smaller than those found in the present study (Rogers *et al.*, 2006; Cleary *et al.*, 2005; Geurts and Wouters, 2001; Laneau *et al.*, 2004), but considerable differences may well explain the discrepancies. In the first place, there were differences in subject population (i.e., adults vs. children). Psychophysical studies have shown that frequency discrimination thresholds in young children are higher than in adults. Halliday *et al.* (2008) measured frequency discrimination thresholds in adults and children. The performance of the children was clearly poorer than that of the adults (Halliday *et al.*, 2008). In the present study better F0 difference thresholds were found for subjects that were older at implantation. This seems to be an unexpected finding. However, age at implantation was highly correlated with age at testing, possibly indicating that older subjects were better at this test, as expected from the results for frequency discrimination. Furthermore, this correlation was weak and non-significant after a Bonferroni correction. Additionally, differences in speech materials may also explain the discrepancies in the results. For instance, Geurts and Wouters (2001) measured JNDs of 0.7 to 2.2 semitones in four adult cochlear implant users with an F0 reference stimulus of 150 Hz in steady-state synthetic vowels (for a reference stimulus of 250 Hz, results were considerably poorer). Instead of a steady state F0, we used stimuli with rapidly changing F0 contours. The JND for F0 differences between steady state complex harmonic stimuli can be less than 1% in normal hearing subjects (Klatt, 1973). If rapidly changing F0 contours are presented to normal hearing subjects, a JND of 19% or 2.3 semitones can be found ('t Hart, 1981), which is similar to the JND of 1.6 semitones for normal hearing adults reported by van de Sandt (2008) who used the same materials and procedure as in the present study.

The results of our NH subjects suggest that children younger than 10 years could not reliably perform the F0

discrimination task. Surprisingly, this was not observed in the implanted group, which may have been due to their greater familiarity with auditory testing in general. Additionally, it must be taken into account that this effect might be absent due to the variability in the implanted group in duration of deafness, duration of CI use, age at implantation, age of deafness, amount of residual hearing, or a combination of these variables. However, no correlations were found between these possible sources of variation and test performance.

## B. Discrimination between questions and affirmations

In the second experiment, the average results of the implanted children show a modest but significant benefit of bimodal fitting for discriminating between questions and affirmations. The percentage correct improved significantly from 56% (CI) to 62% (CI+HA) in the first presentation and from 62% (CI) to 73% (CI+HA) in the repeated presentation.

Our results differ somewhat from those reported by Green *et al.* (2005) who reported mean percentage correct scores of 69% and 81% in the CI-only condition in response to a male speaker and a female speaker, respectively. In Green *et al.*'s experiment, the subjects were all postlingually deaf adults instead of the prelingually deaf children that we tested. A second explanation for the difference in results is that in the study by Green *et al.*, the pitch rises were slower (at least 200 ms) than our study (120 ms). This difference relates to the discrimination of steady state versus rapidly changing F0 contours. Indeed, White and Plack (White and Plack, 1998) found that F0 discrimination improved with increasing duration of the pitch rise.

Another important difference is that Green *et al.* used naturally spoken affirmations and questions; whereas, our study manipulated the last part of the F0 contour. Therefore, other cues may have been available to their subjects besides F0, such as differences in formant spectrum, intensity, and duration. We found significant correlations in the range of 0.4 and 0.5 between F0 discrimination and question/statement identification in the CI-only and the CI+HA condition. In support of our hypothesis, the results suggest that subjects used F0 information rather than other cues in the question versus affirmation task.

In the current study, question/affirmation identification was assessed using five different utterances (*viz.* one affirmation and four different question categories: Q1, Q2, Q3, Q4). The increasing final rise in pitch across the question categories Q1-Q4 resulted in increasingly improved identification, but the largest difference in performance was between Q1 and the other stimuli. This effect was present in the NH and implanted children. It appears that the rising F0 contour that spanned 6–9 semitones in Q1 was too small to be perceived as a question by most of the subjects and was therefore frequently identified as an affirmation. This is in agreement with the JND of 6.9 semitones in the implanted subjects in experiment 1. As a rising contour of approximately 6 semitones is typical of Dutch questions, this result suggests that even when implanted children use a contralat-

eral hearing aid, they require exaggerated F0 excursions to identify a question, particularly when other cues, such as duration, are lacking. In contrast, NH subjects should have been able to hear the rising contour of Q1 as they had a mean JND of 3.3 semitones. It is not clear why Q1 was more often perceived as an affirmation, but it may have been due to either a lack of duration cues or the complexity of the task.

Chatterjee *et al.* (Chatterjee and Peng, 2008) also reported that implanted adults required exaggerated F0 excursions to identify a question, and even with the most exaggerated F0 excursion, not all utterances were identified as questions by their subjects. The overall performance of their group was better than in our study, however the subjects in Chatterjee *et al.*'s study were all adults and the majority of their subjects (8 out of the 10) were postlingually deaf.

## C. The effect of age on the perception of intonation

In the implanted children, no significant correlations were found between age at testing and performance in the two experiments. In contrast, these correlations were present in the group of NH children. They showed particularly large spread in F0 discrimination thresholds at the age of 9 years and younger. This spread ranged from nearly as good as the older children to much poorer. On the question/affirmation identification task, the three youngest children in the NH group ( $\leq 8.3$  years) performed below chance level. This may be related to not understanding the task, lack of concentration, or a small attention span. A possible explanation for the lack of correlation between age and performance in the CI group is their greater familiarity with auditory testing. Alternatively, perhaps not all implanted children, even the older ones, could master the task; any bimodal benefit, if present, would only show in older children after they have developed sufficient skills for the task. Furthermore, comparison of Figs. 3 and 4 shows more variability in the results of implanted children compared to those of older NH children. This may have obscured any effect of age.

The results for implanted children in the first experiment indicate no significant correlations between F0 JND on the one hand, and the aided thresholds, duration of deafness, duration of CI use, and duration of HA use on the other. In addition, no significant correlations were found between F0 JND and age at implantation or age of deafness. These findings suggest that possibly a combination of these factors, or factors currently not investigated are responsible for the variability in performance [*e.g.*, language ability or neural survival, placement of the electrodes, educational or learning variables (Cleary *et al.*, 2005; Peng *et al.*, 2004)]. Future research should consider these issues.

## D. Benefit of the hearing aid for the perception of intonation

In contrast with previous studies that assessed prosody perception in implanted children who depended solely on their CI, the present study investigated the effect of bimodal fitting. Compared to the CI-only condition, performance in the CI+HA condition improved significantly in experiments 1 and 2. As discussed in the introduction, acoustic stimula-

tion provides access to several F0 cues that are not available to the same extent when using a CI-only. These cues include additional low-frequency periodicity information, resolved lower harmonics of F0, and amplitude modulations due to the beating of unresolved higher harmonics within channels. A prerequisite for the use of these cues is audibility. The average aided thresholds were about 50 dB SPL up to 1000 Hz, 55 dB SPL at 2000 Hz, and 75 dB SPL at 4000 Hz. The range between subjects suggests that after amplification, the louder speech segments were audible for most subjects, however, for many subjects frequencies of 1000 Hz and higher may have been relatively soft in the acoustically stimulated ear. This suggests that the severely hearing-impaired subjects may have relied on F0 periodicity information, the presence of resolved lower harmonics, or both periodicity information and lower harmonics. However, we cannot exclude the beating of higher harmonics of up to 1000 Hz as an additional cue to F0. The fact that we did not find a correlation between aided thresholds and bimodal benefit provides some support to the notion that low frequency cues, rather than beating within higher frequency channels, dominate performance.

Part or all of the superior F0 discrimination and question/affirmation identification in the bimodal condition may have been simply due to the performance of the ear with the hearing aid. Due to time constraints, we did not include a hearing aid only condition and we are therefore unable to assign the bimodal benefit to either the true integration of information from both ears, or the subject's use of the ear best equipped for this task. However, in a study with eight postlingually deaf subjects, F0 discrimination was tested with the same task as in Experiment 1. Thresholds improved from 16 (CI-only) to 12 semitones both in the CI+HA and in the HA-only condition, indicating that performance in the CI+HA condition solely depended on the ear provided with the hearing aid (van de Sandt, 2008). A "better-ear" effect was also observed for melody recognition with EAS (Kong *et al.*, 2005).

It is possible that part of the bimodal benefit found in our study is the result of decreased performance by subjects in the CI-only condition, because this is not the natural listening condition for these subjects. However, we assume that the latter effect is important in experiments where subjects are required to recognize words or sentences, but it is less important in our study where subjects were required to complete psychoacoustic tests with little or no semantic or syntactic load. One way to shed light on this issue would be to test a representative control group of unilateral CI users who are deaf in the opposite ear and do not use a hearing aid. However, a large number of subjects would be required to achieve sufficient power given the interindividual variability observed in this study.

## E. F0 perception and speech perception in quiet

Although recent studies showed that F0 may provide benefit in speech understanding, especially in noise (Dorman *et al.*, 2008; Holt *et al.*, 2005), we found no correlation between the speech perception task and F0 perception. F0 also contributes to speech perception in quiet, suggesting that F0

might be an important cue for several linguistic features, such as consonant voicing (Holt *et al.*, 2001), lexical boundaries (Spitzer *et al.*, 2009; Spitzer *et al.*, 2007), and contextual emphasis (Fry 1955). However, the stimuli we used were monosyllabic words. Thus, features such as lexical boundaries or contextual emphasis were not available to the subjects. If we had used multisyllabic words, a correlation between F0 perception and speech perception may have been possible due to the top-down processes of word preselection based on the F0 contour (Marslen-Wilson, 1987).

## V. CONCLUSION

The present study evaluated the benefits of using a cochlear implant and a hearing aid in the non-implanted ear in children. For the implanted group, performance significantly improved when the hearing aid was added. This benefit was found in the discrimination of pitch movements in bisyllabic non-words and in the ability to distinguish between questions and affirmations. However, even with a hearing aid, the implanted group required exaggerated F0 excursions to perceive a pitch accent and to identify a question. These exaggerated excursions are close to the maximum excursions typically used by Dutch speakers. Nevertheless, the results of this study showed that compared to the implant only condition, bimodal fitting improved the perception of intonation

<sup>1</sup>Partial  $\eta^2 = SS_{\text{factor}} / (SS_{\text{factor}} + SS_{\text{error}})$ , with  $0 \leq \eta^2 \leq 1$ .

<sup>2</sup>Huyn-feldt corrected df and p-values are reported when appropriate.

- Assmann, P. F., and Summerfield, Q. (1990). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **88**, 680–697.
- Berenstein, C. K., Mens, L. H. M., Mulder, J. J. S., and Vanpoucke, F. J. (2008). "Current steering and current focusing in cochlear implants: Comparison of monopolar, tripolar, and virtual channel electrode configurations," *Ear Hear.* **29**, 250–260.
- Boersma, P., and Weenink, D. (2006). "Praat: Doing phonetics by computer (version 4.4.33)," <http://www.praat.org/> (Last viewed 10/12/2006).
- Bosman, A. J. (1989). "Speech perception by the hearing impaired," Ph.D. thesis, University Utrecht, The Netherlands.
- Brox, J. P. L., and Nooteboom, S. G. (1982). "Intonation and the perception separation of simultaneous voices," *J. Phonetics* **10**, 23–26.
- Brown, C. A., and Bacon, S. P. (2009). "Low-frequency speech cues and simulated electric-acoustic hearing," *J. Acoust. Soc. Am.* **125**, 1658–1665.
- Chatterjee, M., and Peng, S. C. (2008). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," *Hear. Res.* **235**, 143–156.
- Ching, T. Y., Incerti, P., and Hill, M. (2004). "Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears," *Ear Hear.* **25**, 9–21.
- Ciocca, V., Francis, A. L., Aisha, R., and Wong, L. (2002). "The perception of Cantonese lexical tones by early-deafened cochlear implantees," *J. Acoust. Soc. Am.* **111**, 2250–2256.
- Cleary, M., Pisoni, D. B., and Kirk, K. I. (2005). "Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants," *J. Speech Lang. Hear. Res.* **48**, 204–223.
- Davidson, L. S., Skinner, M. W., Holstad, B. A., Fears, B. T., Richter, M. K., Matusofsky, M., Brenner, C., Holden, T., Birath, A., Kettel, J. L., and Scollie, S. (2009). "The effect of instantaneous input dynamic range setting on the speech perception of children with the nucleus 24 implant," *Ear Hear.* **30**, 340–349.
- Dorman, M. F., Gifford, R. H., Spahr, A. J., and McKarns, S. A. (2008). "The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies," *Audiol. Neuro-Otol.* **13**, 105–112.

- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Fu, Q. J., Chinchilla, S., Nogaki, G., and Galvin, J. J., III (2005). "Voice gender identification by cochlear implant users: the role of spectral and temporal resolution," *J. Acoust. Soc. Am.* **118**, 1711–1718.
- Geurts, L., and 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., and Wouters, J. (2004). "Better place-coding of the fundamental frequency in cochlear implants," *J. Acoust. Soc. Am.* **115**, 844–852.
- Green, T., Faulkner, A., and 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**, 2155–2164.
- Green, T., Faulkner, A., and 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., and Macherey, O. (2005). "Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification," *J. Acoust. Soc. Am.* **118**, 375–385.
- Gussenhoven, C., Rietveld, A. C. M., and Terken, J. (1999). "ToDI: Transcription of Dutch intonation," <https://lands.let.kun.nl/todi> (Last viewed 10/20/2006).
- Gussenhoven, C., and Rietveld, T. (2000). "The behavior of H\* and L\* under variations in pitch range in Dutch rising contours," *Lang Speech* **43**, 183–203.
- Halliday, L. F., Taylor, J. L., Edmondson-Jones, A. M., and Moore, D. R. (2008). "Frequency discrimination learning in children," *J. Acoust. Soc. Am.* **123**, 4393–4402.
- Hamilton, N., Green, T., and Faulkner, A. (2007). "Use of a single channel dedicated to conveying enhanced temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification," *Int. J. Audiol.* **46**, 244–253.
- Hart, J., Collier, R., and Cohen, A. (1990). *A Perceptual Study of Intonation* (Cambridge University Press, Cambridge), pp. 10–37.
- Henry, B. A., and Turner, C. W. (2003). "The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners," *J. Acoust. Soc. Am.* **113**, 2861–2873.
- Holt, L. L., Lotto, A. J., and Kluender, K. R. (2001). "Influence of fundamental frequency on stop-consonant voicing perception: A case of learned covariation or auditory enhancement?," *J. Acoust. Soc. Am.* **109**, 764–774.
- Holt, R. F., Kirk, K. I., Eisenberg, L. S., Martinez, A. S., and Campbell, W. (2005). "Spoken word recognition development in children with residual hearing using cochlear implants and hearing AIDS in opposite ears," *Ear Hear.* **26**, 82S–91S.
- Kerckhoff, J. J. P., and Rietveld, A. C. M. (1995). "The generation of prosody in the Nijmegen rule oriented speech synthesis system," in ESCA 4th European Conference on Speech Communication and Technology, GTH, UPM, Madrid, Spain, edited by J. M. J. Pardo, E. O. Enriquez, J. Ferreira, J. Macias, and F. J. Valverde, pp. 1831–1834.
- Klatt, D. H. (1973). "Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception," *J. Acoust. Soc. Am.* **53**, 8–16.
- Kong, Y. Y., and 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., Deeks, J. M., Axon, P. R., and Carlyon, R. P. (2009). "Limits of temporal pitch in cochlear implants," *J. Acoust. Soc. Am.* **125**, 1649–1657.
- Kong, Y. Y., Stickney, G. S., and Zeng, F. G. (2005). "Speech and melody recognition in binaurally combined acoustic and electric hearing," *J. Acoust. Soc. Am.* **117**, 1351–1361.
- Laneau, J., and Wouters, J. (2004). "Multichannel place pitch sensitivity in cochlear implant recipients," *J. Assoc. Res. Otolaryngol.* **5**, 285–294.
- Laneau, J., Wouters, J., and Moonen, M. (2004). "Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear implantees," *J. Acoust. Soc. Am.* **116**, 3606–3619.
- Laneau, J., Wouters, J., and Moonen, M. (2006). "Improved music perception with explicit pitch coding in cochlear implants," *Audiol. Neuro-Otol.* **11**, 38–52.
- Lavandier, M., and Culling, J. F. (2008). "Speech segregation in rooms: monaural, binaural, and interacting effects of reverberation on target and interferer 1," *J. Acoust. Soc. Am.* **123**, 2237–2248.
- Lee, K. Y., van Hasselt, C. A., Chiu, S. N., and Cheung, D. M. (2002). "Cantonese tone perception ability of cochlear implant children in comparison with normal-hearing children," *Int. J. Pediatr. Otorhinolaryngol.* **63**, 137–147.
- Lee, S. H., Lee, K. Y., Huh, M. J., and Jang, H. S. (2008). "Effect of bimodal hearing in Korean children with profound hearing loss," *Acta Oto-Laryngol.* **128**, 1–6.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Luo, X., and Fu, Q. J. (2006). "Contribution of low-frequency acoustic information to Chinese speech recognition in cochlear implant simulations," *J. Acoust. Soc. Am.* **120**, 2260–2266.
- Marslen-Wilson, W. D. (1987). "Functional parallelism in spoken word-recognition," *Cognition* **25**, 71–102.
- Moore, B. C. (2003). "Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants," *Otol. Neurotol.* **24**, 243–254.
- Moore, B. C., and Carlyon, R. P. (2005). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," *Pitch: Neural Coding and Perception* (Springer, New York), pp. 234–277.
- Moore, B. C., and Peters, R. W. (1992). "Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity," *J. Acoust. Soc. Am.* **91**, 2881–2893.
- Moulines, E., and Charpentier, F. (1990). "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones," *Speech Commun.* **9**, 453–467.
- O'Halpin, R., Faulkner, A., Rosen, S., and Viani, L. (2005). "The perception of stress and intonation by children with cochlear implants," Conference on Implantable Auditory Prostheses, Asilomar, CA.
- Oxenham, A. J. (2008). "Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants," *Trends Amplif.* **12**, 316–331.
- Peng, S. C., Tomblin, J. B., Cheung, H., Lin, Y. S., and Wang, L. S. (2004). "Perception and production of mandarin tones in prelingually deaf children with cochlear implants," *Ear Hear.* **25**, 251–264.
- Rogers, C. F., Healy, E. W., and Montgomery, A. A. (2006). "Sensitivity to isolated and concurrent intensity and fundamental frequency increments by cochlear implant users under natural listening conditions," *J. Acoust. Soc. Am.* **119**, 2276–2287.
- Shannon, R. V. (1983a). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," *Hear. Res.* **11**, 157–189.
- Shannon, R. V. (1983b). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," *Hear. Res.* **12**, 1–16.
- Spitzer, S., Liss, J., Spahr, T., Dorman, M., and Lansford, K. (2009). "The use of fundamental frequency for lexical segmentation in listeners with cochlear implants," *J. Acoust. Soc. Am.* **125**, EL236–EL241.
- Spitzer, S. M., Liss, J. M., and Mattys, S. L. (2007). "Acoustic cues to lexical segmentation: A study of resynthesized speech," *J. Acoust. Soc. Am.* **122**, 3678–3687.
- 't Hart, J. (1981). "Differential sensitivity to pitch distance, particularly in speech," *J. Acoust. Soc. Am.* **69**, 811–821.
- Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., and Henry, B. A. (2004). "Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing," *J. Acoust. Soc. Am.* **115**, 1729–1735.
- Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W., and McDermott, H. J. (2005). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," *J. Acoust. Soc. Am.* **117**, 3126–3138.
- von de Sandt, E. (2008). "Bimodal hearing in adults: Additional benefit of wearing a unilateral cochlear implant and a contralateral hearing aid ENT/Audiology," Master Thesis, Radboud University Nijmegen Medical Center, Nijmegen, pp. 1–81.
- von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Sturzebecher, E., and Klinke, R. (1999). "Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss," *ORL J. Otorhinolaryngol. Relat. Spec.* **61**, 334–340.
- White, L. J., and Plack, C. J. (1998). "Temporal processing of the pitch of complex tones," *J. Acoust. Soc. Am.* **103**, 2051–2063.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). "Better speech recognition with cochlear implants," *Nature (London)* **352**, 236–238.
- Wilson, B. S., Schatzer, R., Lopez-Poveda, E. A., Sun, X., Lawson, D. T., and Wolford, R. D. (2005). "Two new directions in speech processor design for cochlear implants," *Ear Hear.* **26**, 73S–81S.
- Zeng, F. G. (2004). "Trends in cochlear implants," *Trends Amplif.* **8**, 1–34.