More Than Words: the Relative Roles of Prosody and Semantics in the Perception of Emotions in Spoken Language by Postlingual Cochlear Implant Users

Riki Taitelbaum-Swead,[1,2] Michal Icht,[1] and Boaz M. Ben-David[3,4,5]

Objectives: The processing of emotional speech calls for the perception and integration of semantic and prosodic cues. Although cochlear implants allow for significant auditory improvements, they are limited in the transmission of spectro-temporal fine-structure information that may not support the processing of voice pitch cues. The goal of the current study is to compare the performance of postlingual cochlear implant (CI) users and a matched control group on perception, selective attention, and integration of emotional semantics and prosody.

Design: Fifteen CI users and 15 normal hearing (NH) peers (age range, 18–65 years) 1istened to spoken sentences composed of different combinations of four discrete emotions (anger, happiness, sadness, and neutrality) presented in prosodic and semantic channels—T-RES: Test for Rating Emotions in Speech. In three separate tasks, listeners were asked to attend to the sentence as a whole, thus integrating both speech channels (integration), or to focus on one channel only (rating of target emotion) and ignore the other (selective attention). Their task was to rate how much they agreed that the sentence conveyed each of the predefined emotions. In addition, all participants performed standard tests of speech perception.

Results: When asked to focus on one channel, semantics or prosody, both groups rated emotions similarly with comparable levels of selective attention. When the task was called for channel integration, group differences were found. CI users appeared to use semantic emotional information more than did their NH peers. CI users assigned higher ratings than did their NH peers to sentences that did not present the target emotion, indicating some degree of confusion. In addition, for CI users, individual differences in speech comprehension over the phone and identification of intonation were significantly related to emotional semantic and prosodic ratings, respectively.

Conclusions: CI users and NH controls did not differ in perception of prosodic and semantic emotions and in auditory selective attention. However, when the task called for integration of prosody and semantics, CI users overused the semantic information (as compared with NH). We suggest that as CI users adopt diverse cue weighting strategies with device experience, their weighting of prosody and semantics differs from those used by NH. Finally, CI users may benefit from rehabilitation strategies that strengthen perception of prosodic information to better understand emotional speech.

Key words: Cochlear implant, Emotion identification, Emotion perception, Postlingual deafness, Prosody.

(Ear & Hearing 2022;00;00–00)

INTRODUCTION

Cochlear implants (CIs) are now the standard of care for individuals with severe to profound hearing impairment, including prelingually and postlingually deaf children and adults (Wilson 2013, 2015; De Raeve et al. 2015). Of interest are postlingually deaf adults, born with intact auditory systems, who receive CIs after deafness occurs (Firszt et al. 2004; Gifford et al. 2008; Holden et al. 2013). Research efforts have focused on the success of CIs in restoring speech identification and on the perception of segmental features (Kelly et al. 2005; Lazard et al. 2012; Blamey et al. 2013). Several studies have also examined the challenges imposed by the CI on the perception of prosodic cues (Chatterjee & Peng 2008; Meister et al. 2009; Peng et al. 2012). Indeed, successful social communication relies on processing emotional prosody (rhythm, intonation, etc.) via suprasegmental features, as well as semantic content via segmental features (Bryant & Barrett 2008). Integration of these emotional channels—prosody and semantics—appears to be particularly challenging for people with hearing difficulties (Zupan et al. 2009; Ben-David et al. 2016b; Paulmann et al. 2008; Dupuis & Pichora-Fuller 2014; Ben-David et al. 2019).

The few studies that have assessed vocal emotion perception in postlingual CI users have found reduced prosody identification (Luo et al. 2007; Agrawal et al. 2013; Chatterjee et al. 2015; Gilbers et al. 2015). Yet, to date, no study has examined the integration of semantic and prosodic cues in the perception of emotional speech in this population. For instance, a voice message over the phone of “Do not waste my time!” (angry semantics) spoken with happy prosody may be interpreted differently by various groups of listeners. Research shows that young normal hearing (NH) listeners tend to base their interpretation on the prosodic content (Ben-David et al. 2016b, 2019, 2020; Oron et al. 2020, Leshem et al. 2020). Would CI users do the same, or would they base their interpretation on the semantic content, or on a combination of the two? The goal of the current study was to answer this question by comparing the performance of postlingual CI users and a matched control group of NH individuals on a spoken-emotion processing test.

Perception of Emotions in Speech

Spoken communication, and specifically the processing of spoken emotion, has an important role in daily social interactions (Loveland et al. 1997; Ben-David et al. 2013). Clearly, when a listener does not fully comprehend emotions conveyed by a speaker, miscommunication can ensue, negatively impacting quality of life and social well-being (Hudepohl et al. 2015; Heinrich et al. 2016). When visual information during communication is also absent (e.g., when talking over the phone), the ability to derive emotional meaning from spoken language

---

[1]Department of Communication Disorders, Ariel University, Israel; [2]Meuhedet Health Services, Tel Aviv, Israel; [3]Baruch Ivcher School of Psychology, Reichman University (IDC), Herzliya, Israel; [4]Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada; and [5]Toronto Rehabilitation Institute, University Health Networks (UHN), ON, Canada.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal’s Web site (www.ear-hearing.com).

<zdoi; 10.1097/AUD.0000000000001199>

relies on the integration of two auditory channels, semantics and prosody. However, CI users extracting spoken emotional meaning from degraded auditory information may experience altered perception (Taitelbaum-Swead & Fostick 2017).

The Test of Rating of Emotions in Speech (T-RES) was designed to assess the separate roles of emotional prosody and semantics in the processing of spoken emotions. In several T-RES studies involving over 180 young NH participants, Ben-David et al. (2016b, 2019, 2020); Oron et al. (2020); and Leshem et al. (2020) identified two main observed behaviors in typical spoken emotion processing: (1) *Failures of selective attention*, in which listeners fail to selectively attend to one channel while actively ignoring the other and (2) *Prosodic bias*, in which prosody is more strongly weighted on emotional ratings than is semantic information. These observed behaviors have yet to be examined in CI users. However, some relevant literature exists regarding the processing of spoken emotions and selective attention in CI users, as discussed next.

CI Users: Processing Spoken Emotion Through Identification of Semantics and Prosody

Semantics • In ideal listening conditions (e.g., quiet background, typical speech rate, familiar accent), spoken word recognition appears to be generally well-preserved in postlingual CI users, especially when semantics or syntactic context is provided. However, CI users find it especially challenging to recognize speech under degraded or adverse listening conditions (Dorman & Gifford 2017). For example, when the speaking rate increases, speech understanding worsens for CI users (Ji et al. 2013; Jaekel et al. 2017). Given that speaking rate varies in emotional speech and may change some acoustic features, CI users may have difficulty extracting semantic information (e.g., Breitenstein et al. 2001; Dupuis & Pichora-Fuller 2014). For example, when the speaking rate increases, “this is *bad*” may be confused with “this is *dad*”, altering the meaning of the emotional sentence (Ben-David et al. 2011b; Hadar et al. 2016; Nitsan et al. 2019). This confusion may be increased in CI users, who have difficulty discriminating between the consonants/b/ and /d/ (Mason & Kokkinakis 2014).

Prosody • Research on non-emotional prosody indicates that CI users do not differentiate as well as their NH peers do between statements and questions (Chatterjee & Peng 2008; Meister et al. 2009; Peng et al. 2012) and between stressed versus unstressed words (Kalathottukaren et al. 2015). This decreased performance hints at possible difficulties CI users might experience in identifying emotional prosodic cues as well.

The processing of emotional prosodic cues calls for correct analysis of supra-segmental features, such as changes in fundamental frequency (F0), pitch, intensity, and duration (Peng et al. 2009; Pichora-Fuller et al. 2016). CI systems have limitations in transmitting such spectro-temporal fine structure information, possibly accounting for impaired processing of emotional prosodic cues (Zeng 2002). Indeed, a recent meta-analysis found that CI users in general were less accurate than were NH listeners in identifying linguistic and emotional prosody because of inadequate transmission of F0 cues (Everhardt et al. 2020).

Only a few studies have tested emotional prosody perception in adult postlingual CI users (Agrawal et al. 2013; Chatterjee et al. 2015; Gilbers et al. 2015; Luo et al. 2007; Paquette et al. 2018). The studies used a variety of tools, emotions, and stimuli, thus limiting the ability to generalize the results. However, a general trend of reduced ability to identify (at least some) prosodic emotions was detected.

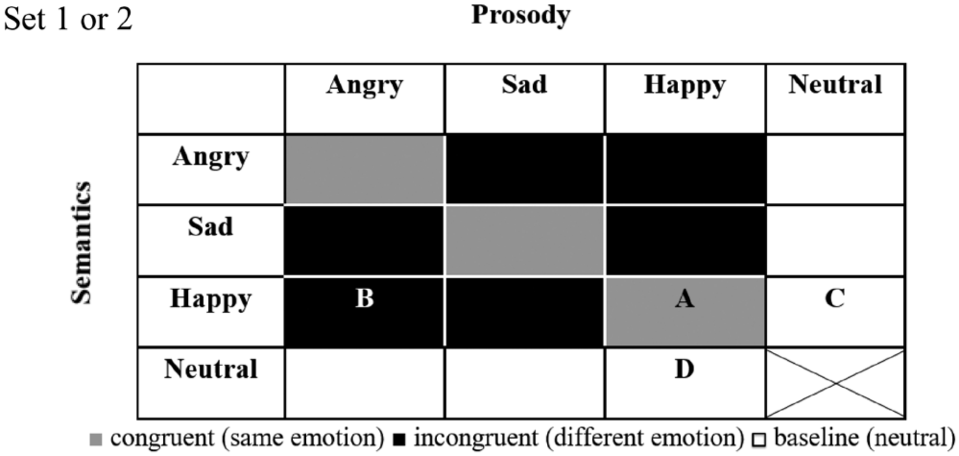
CI Users: Failures of Selective Attention • There is mixed evidence in the literature regarding postlingual CI users’ cognitive performance. Most studies have focused on the visual modality, testing selective attention and inhibition. These executive functions relate to the ability to focus on the information presented in one channel while inhibiting incongruous information presented in the other channel (Ben-David & Algom 2009; Melara & Algom 2003). The classic visual color-word Stroop test (e.g., responding “red” to the word BLUE printed in red) has been taken as the gold standard for assessing these abilities [but see Ben-David & Schneider (2009, 2010)]. While some studies have observed decreased performance in the color-word Stroop task in this group (Moberly et al. 2019), others have observed no significant differences in Stroop performance and some other attentional tasks (Moberly et al. 2016). Moreover, it is not clear whether performance on visual cognitive tasks can generalize to the auditory domain (Knight & Heinrich, 2018). Thus a study by Henkin and colleagues (2014) is noteworthy in its departure from focusing on the visual modality. Using an *auditory* Stroop task, they found postlingual adult CI users did not perform differently from their age-matched NH peers. However, the data related only to adults over age 60. It is not clear whether these findings can be generalized to younger CI users (Wingfield 2016).

The Current Study

In the present study, performance on a Hebrew-adapted version of the T-RES (Shakuf et al. 2016) was compared between a group of postlingual CI users (CI group) and a matched group of NH adults. In three separate tasks, participants were presented with spoken sentences in which the emotional semantic and prosodic content appeared in various combinations of congruence and incongruence from trial to trial. Participants were asked to rate the extent to which a predefined emotion was expressed by the prosody alone (prosodic rating task), the semantics alone (semantic rating task), or the sentence as a whole (general rating task). For a visual illustration of the test, see Figure 1. The current study, to the best of our knowledge, is the first to test how postlingual CI users *rate* the emotional semantic content of spoken sentences.

Predictions were made for the following research questions:

1. *Rating of Target Emotions (RTE): Do both groups rate semantic and prosodic emotions in a similar way?* The literature indicates that CI users may be able to successfully identify emotional semantics, but experience difficulties identifying prosodic emotions. Thus, we predicted no group differences when asked to rate semantics (RTE semantics), but reduced performance for CI users when asked to rate prosody (RTE prosody; see the white cells C and D in Fig. 1).
2. *Selective Attention: Is there a group difference in selective attention to the prosodic or semantic channels?* Evidence in the literature suggests no CI-related change in selective attention performance. Therefore, we predicted no group differences in failures of selective attention when asked to rate one channel (e.g., semantics), while ignoring the other (e.g., prosody; see the black cells in Fig. 1). This would be evident when the prosody and semantics present different emotions (incongruent trials).
3. *Integration of Channels and Channel Bias: Do both groups similarly assign weights to prosodic and semantic*



	Angry	Sad	Happy	Neutral
Angry	▨	■	■	□
Sad	■	▨	■	□
Happy	■ B	■	▨ A	□ C
Neutral	□	□	□ D	☒

▨ congruent (same emotion) ■ incongruent (different emotion) □ baseline (neutral)

Fig. 1. General design of the Test for Rating Emotions in Speech (T-RES): Stimuli. All combinations of prosody and semantics are presented in each emotional rating block (note: neutral semantics spoken with neutral prosody was deemed uninformative or confusing and therefore not presented).

channels? The literature suggests relatively preserved semantic processing in contrast to impaired prosodic processing for CI users. We predicted CI users would be less biased to the prosodic channel than their NH peers, when they are asked to judge the spoken sentence as a whole, with no direction to one of the two channels. Again, this would be evident in incongruent trials (see the black cells in Fig. 1).

4. *Individual Characteristics: Does CI users' T-RES performance relate to their speech perception functions and HA status?* The literature shows that postlingual CI users vary substantially in sensory abilities impacting speech processing. We predicted that reduced auditory perception function, as tested by the Speech Over the Phone questionnaire and the Linguistic Intonation test, would impact semantic and prosodic processing, respectively, across all tasks. It is also possible that bimodal users (CI and hearing aid) may be able to better utilize pitch cues compared to CI-only users.

MATERIALS AND METHODS

Participants

Fifteen postlingually deaf adults who use CI (CI group) were recruited via associations for hearing impaired individuals. In addition, 15 matched control participants with normal hearing for their age (NH group) were recruited via publications in social media. Inclusion criteria for all participants were as follows (as assessed by self-report questionnaires): (a) native Hebrew speaker; (b) possessing a high school or college diploma; (c) normal or corrected to normal vision; and (d) within the age range of 18 to 65 years old. The study was approved by the institutional ethics committee and all participants gave their informed consent before the experiment commenced. Participants received the equivalent of $25 (US) to compensate for the time they spent in the study.

The CI Group • Additional inclusion criteria implemented for the CI group were: (a) onset of hearing loss after language acquisition (postlingual), at the mean age of 16.8 years (SD = 14.6 years); (b) severe to profound hearing loss in both ears before implantation; and (c) use of at least one CI. Their mean age of implantation was 37.7 years (*SD* = 13.8 years). Individual demographic information and background data for each CI user are shown in Table 1.

The NH Group • Control participants were pair-matched to participants in the CI group in terms of age (within 18 months), gender, and level of education (see Table 2). The additional inclusion criterion for the NH group was a mean pure-tone average (PTA; 0.5, 1, 2, 4, and 8 kHz) of less than 20 dB HL in both ears, as verified by an audiology student.

Tools and Materials

Test of Rating of Emotions in Speech

T-RES stimuli • For this test, participants are presented with spoken sentences in which the emotional semantic and prosodic content appears in four different combinations from trial to trial. For example, consider Figure 1: cell "A", a congruent (matched) stimulus, represents a semantically happy sentence (e.g., "I got a raise in my salary") spoken with a congruent happy prosody, while cell "B", an incongruent (mismatched) stimulus, represents a semantically happy sentence (e.g., "I received an amazing gift") spoken with an incongruent angry prosody. Cell "C", a baseline for the semantic channel, represents a semantically happy sentence ("I won the lottery") spoken with a neutral (emotionless) prosody, whereas cell "D", a baseline for the prosodic channel, represents a semantically neutral sentence (e.g., "There are many hangers in the closet") spoken with happy prosody.

When the emotional semantic and prosodic cues are congruent (matched), the listener can identify the target emotion by either using both channels or focusing on only one. However, when the emotional semantic and prosodic cues are incongruent, integration across channels is necessary (Ben-David et al. 2011b). These instances of incongruent combinations enable us to pit prosodic and semantic cues against one another and assess their separate roles in spoken emotion processing.

We used the Hebrew version of the T-RES (Shakuf et al. 2016), with the following emotions: Anger, Happiness, Sadness, and Neutrality. Semantic sentences were comparable on main linguistic characteristics (e.g., frequency of usage, sentence length) across the four affective categories to avoid possible biases (Ben-David et al. 2011b). These sentences were recorded in the four different prosodies by a native Hebrew-speaking Israeli professional radio-drama actress. The final experimental

TABLE 1. Demographic information and background data for the CI group

Participant	Age (yrs)	Gender	Education	N. of CIs	Implanted Ear	HA Use Post-CI	HA Use Pre CI (yrs)	Age at First CI (yrs)	Duration of Implant Use (yrs)	Type of CI
1	57	F	Academic	1	L	N	20	40	17	C
2	65	F	Academic	1	R	Y	9	64	1	C
3	46	F	Academic	1	L	Y	14	43	3	M
4	42	M	Academic	1	L	N	10	32	10	M
5	59	M	Academic	1	L	Y	1	53	6	C
6	59	M	Academic	2	R, L	--	4	43	16	C
7	43	F	H School	2	R, L	--	13	38	5	M
8	52	F	H School	1	L	Y	7	46	6	C
9	21	F	H School	1	R	Y	14	19	2	AB
10	58	F	Academic	1	L	Y	6	57	1	C
11	26	F	Academic	2	R, L	--	22	25	1	AB
12	27	M	Academic	2	R, L	--	17	22	5	C
13	19	F	H School	1	R	N	3	18	1.5	M
14	52	F	Academic	1	R	N	15	34	18	C
15	33	F	H School	1	L	Y	15	32	1	AB

Gender: F = female, M = male; Education: H school = high school diploma, Academic = at least obtained a bachelor's degree at an accredited academic institute; Implanted ear: R=right, L=left; HA= Hearing Aid; Type of CI: C = Cochlear, AB=Advanced Bionics, M = Med-EL

stimuli were made of two different sets of 15 sentences each, in which each semantic category was represented once in each of the tested prosodies, generating a 4 (semantics) X 4 (prosody) matrix, as shown in Figure 1. It should be noted that the combination of neutral prosody and neutral semantics was considered uninformative and was not included (Ben-David et al. 2016b). All recorded spoken sentences were rated as distinctive and as exemplars of their respective prosodic and semantic categories by a group of trained raters, closely following the procedures of the English version (Ben-David et al. 2011b, 2013). Digital audio files were equated with respect to their root-mean-square amplitude. Spoken sentences were selected such that their duration did not differ significantly across emotions (see, Ben-David et al. 2013, 2019). An acoustic analysis of the Hebrew stimuli (Carl et al. 2022) using Praat software, version 6.1.07 (Boersma & Weenink, 2019) indicated that the mean F0 (fundamental frequency) and articulation rates varied across emotions, yet were naturally preserved. The full acoustical data are available in Table 3.

T-RES Design • Figure 2 presents the makeup of the T-RES paradigm. In each trial, listeners are asked to rate how much they agree that the speaker conveys a predefined emotion (Anger, Sadness, or Happiness, in three separate rating blocks) using a 6-point Likert scale. For example, "How much do you agree that the speaker conveys happiness? From 1—strongly disagree to 6—strongly agree." These three emotions were found to be universal (Zupan et al. 2009), easily recognized, and distinguished in both prosody and semantics (Laukka 2003; Scherer et al. 2001). While the original T-RES included the emotion "fear", it was found to be the least reliable in previous studies (Ben-David et al. 2016a, 2019; Pell et al. 2009). Therefore, it is not included in the current study, or in other recent studies (Oron et al. 2020; Ben-David et al. 2020; Leshem et al. 2020).

T-RES Tasks • As presented in Figure 2, the T-RES consists of three tasks: (a) general rating, in which listeners are asked to rate the overall emotion of the sentence as a whole, as if listening to a phone call; (b) prosodic rating, in which listeners are asked to rate the sentence based only on prosodic information

TABLE 2. Demographic, speech perception, and cognitive data for CI users and NH controls

	CI	NH	Test
n	15	15	
Age (mean, SD), yrs	43.9 (15.3)	44.2 (15)	$t(28) = 0.05$, $p = 0.96$
Gender (f, m)	11, 4	11, 4	
HAB-Q score (mean, SD)	61.3 (10.4)	----	
HAB-N score, SNR = 0 dB (mean, SD)	12.0 (9.2)	79 (6.3)	$t(28) = 23.2$, $p < 0.001$
Level of education			
Academic (≥Bachelor's degree)	10	10	
High school	5	5	
HeSPAC intonation score (mean, SD)	89.9 (8.6)	98.9 (4.4)	$t(28) = 3.6$, $p < 0.001$
100% accuracy	5/15*	14/15	
<100% accuracy	10/15†	1/15	$\chi^2(1) = 6.1$, $p = 0.01$
Comprehension over phone			
Some difficulties	7/15	—	
No difficulties	8/15	—	
Auditory forward digit span, number of digits recalled (mean, SD)	8.6 (1.45)	8.9 (1.4)	$t(28) = 0.5$, $p = 0.61$

Statistical analysis for differences between groups.
**Out of which, four used HA and CI.*
†Out of which, three used HA and CI.
HAB-N, HAB test in noise; HAB-Q, Hebrew version of the AB words test (HAB) in quiet.

TABLE 3. Means and standard deviations (in parentheses) for each acoustic measure, across emotional prosodies

	Anger	Happiness	Sadness	Neutrality
Mean F0 (Hz)	405.8 (32.7)	359.6 (82.3)	250.8 (11.5)	173.3 (20.6)
Range F0 (Hz)	350.9 (56)	265.1 (77.8)	315.2 (137.8)	171.1 (130.1)
Mean articulation rate (syllables/Sec)	4.95 (1.18)	4.14 (0.5)	3.77 (0.43)	4.41 (0.33)
Mean intensity (dB)	57.8 (3.4)	54.1 (3.5)	52 (4)	45.1 (5.3)
Range intensity (dB)	70.7 (4)	68.6 (2.4)	63.8 (1.8)	59.4 (5.7)
Mean duration (sec)	3.3 (.5)	3.1 (.3)	4.4 (.5)	3.7 (.8)

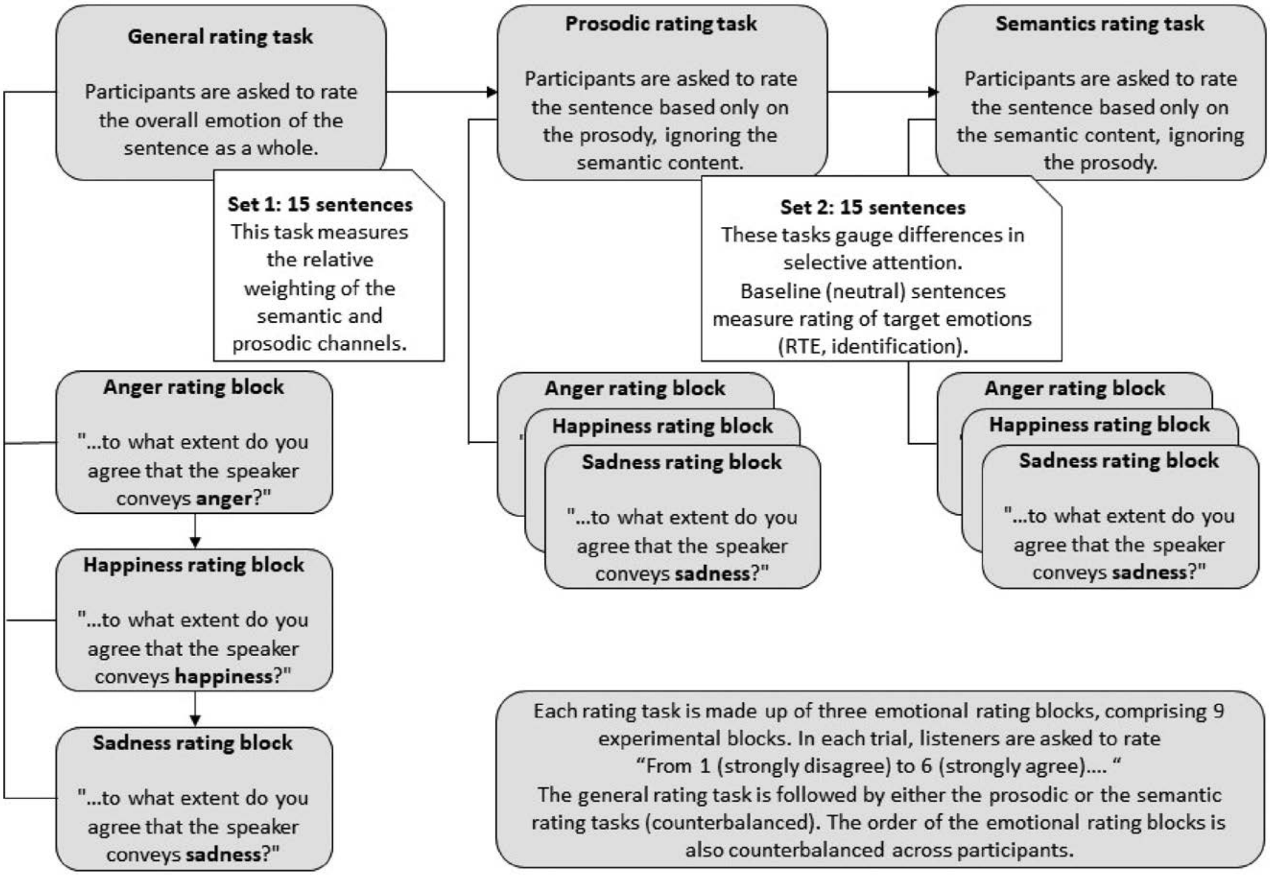
The data was adapted from Carl et al. (2022), by personal correspondence

(ignoring the semantics); and (c) semantic rating, in which listeners are asked to rate the sentence based only on semantic information (ignoring the prosody). Each task consisted of three emotional rating blocks, corresponding to the three target emotions (anger-rating, happiness-rating, and sadness-rating). The experimental session began with the general rating task for all participants. For a randomly chosen half of the participants in each group, this was followed by the semantic rating task and then the prosodic rating task. This order was reversed for the other half of the participants. The order of the three emotion rating blocks was counterbalanced by using a Latin square design, and the order of the trials in each block was fully randomized.

In sum, in the general rating task, each sentence from set 1 was presented once for each of the three emotion rating blocks (anger, sadness, and happiness), totaling 45 trials (15 sentences * 3 rating blocks). Similarly, in the semantic and prosodic rating tasks, each sentence from set 2 was presented once for each of the three emotion rating blocks (anger, sadness, and happiness), totaling 90 trials (15 sentences * 3 rating blocks * 2 tasks). This constituted 135 trials per session (less than 25 minutes). In each trial, each participant made a single rating judgment, thereby producing 135 responses in total.

Measures for Individual Characteristics

Speech Perception Test • The Hebrew version of the AB words test (HAB) was used in quiet conditions (HAB-Q) and with background white noise, at a signal-to-noise ratio of 0 dB (HAB-N) (Taitelbaum-Swead et al. 2005). The test is composed of lists of ten meaningful, one-syllable consonant-vowel-consonant phonetically balanced Hebrew words (i.e., in each list, every consonant appears once, and every vowel appears twice) presented in 70 dB SPL. For CI participants, four lists of ten words were presented, two in each condition (quiet and noise),



General rating task

Participants are asked to rate the overall emotion of the sentence as a whole.

Prosodic rating task

Participants are asked to rate the sentence based only on the prosody, ignoring the semantic content.

Semantics rating task

Participants are asked to rate the sentence based only on the semantic content, ignoring the prosody.

Set 1: 15 sentences
This task measures the relative weighting of the semantic and prosodic channels.

Set 2: 15 sentences
These tasks gauge differences in selective attention.
Baseline (neutral) sentences measure rating of target emotions (RTE, identification).

Anger rating block

"...to what extent do you agree that the speaker conveys **anger**?"

Happiness rating block

"...to what extent do you agree that the speaker conveys **happiness**?"

Sadness rating block

"...to what extent do you agree that the speaker conveys **sadness**?"

Each rating task is made up of three emotional rating blocks, comprising 9 experimental blocks. In each trial, listeners are asked to rate "From 1 (strongly disagree) to 6 (strongly agree)…. "
The general rating task is followed by either the prosodic or the semantic rating tasks (counterbalanced). The order of the emotional rating blocks is also counterbalanced across participants.

Fig. 2. General design of the Test for Rating Emotions in Speech (T-RES): experimental design. The three tasks, the nine rating blocks, and their related experimental sets.

and the participant was requested to repeat each word. NH participants were tested only in noise.

Identification of Linguistic Intonation • The intonation subtest of the Hebrew Speech Pattern Contrasts (HeSPAC; Kishon-Rabin et al. 2002) assesses the ability to discriminate between a statement and a yes/no question. Two lists are used, each containing 12 sentences comprised of familiar words that could have two possible intonation curves—a statement or a question. Half of the sentences are spoken with prosody expressing declarativity, and the other half with interrogative prosody. In each trial, participants are asked to indicate whether the spoken sentence is a statement or a question.

Auditory Forward Digit Span Test • Gauging working memory span (also related to short-term memory buffer, Harel-Arbeli et al. 2021; Nitsan et al. 2019), this test presents sets of random digits read aloud at a rate of one per second, with instructions to report them back verbatim in the order in which they are heard. The shortest list contains two digits and increases the number of digits progressively until the individual is no longer able to recall all of the digits accurately and in the correct order. Participants receive two lists of each length, and the individual's span is recorded as the maximum list length at which at least one of the two lists is accurately recalled (Wechsler 1997; Nitsan et al. 2019).

Hearing Experience Over the Phone • Participants were asked to report, using a 4-point Likert scale, their general experience of comprehension during phone conversations (from 1 = "much difficulty" to 4 = "no difficulties").

Design and Procedure

All participants were tested individually in a quiet room. Upon arrival, all participants received an explanation of the experimental tasks. Those wishing to participate provided signed informed consent. Next, demographic and background data were obtained (via questionnaire) and the following tests were performed: HAB-N, identification of linguistic intonation, and digit span. Members of the CI group were also asked to perform the HAB-Q test and answer the questionnaire regarding their hearing experience in phone conversations. All CI users performed all tasks using their personal hearing devices (CIs with or without hearing aids). For the NH group, pure tone audiometric thresholds were assessed using an Interacoustics-AD629 audiometer with headphones. The full experimental session lasted no more than one hour; breaks were offered upon participants' request. Following the session, participants were debriefed. None reported difficulties or expressed any other complaints regarding the study.

Statistical Analyses

Analyses used repeated measures mixed model ANOVAs (GLM) with average ratings as the dependent variable, group (X2: CI versus NH) as a between-participants variable, and target emotion (X3: anger, sadness, or happiness) and rated channel (X2: prosodic versus semantic rating) in all but general rating task analyses as within-participants variables. Each test included one or two other within-participants variable(s), as specified in Supplemental Appendix A, http://links.lww.com/EANDH/B2.

1. RTE was gauged as the difference between the average ratings of sentences that presented the target emotion in the attended channel versus sentences that did not. This is visually depicted in Figure 3, comparing "type 5" with "type 6" trials, respectively (also, see Equation 1 in Supplemental Appendix A, http://links.lww.com/EANDH/B2). The Prosodic and Semantic rating tasks were analyzed using baseline sentences (in which the to-be-ignored channel was neutral; see white cells in Fig. 1).
2. Selective attention was gauged as the difference between average ratings of sentences that presented the target emotion *only* in the to-be-ignored channel, with sentences that did not present the target emotion in either channel. This is visually depicted in Figure 3, comparing "type 7" with "type 8" trials, respectively (also, see Equation 2 in Supplemental Appendix A, http://links.lww.com/EANDH/B2). The prosodic and semantic rating tasks were analyzed using emotional sentences (see black and gray cells in Fig. 1).
3. Integration was gauged by comparing ratings for congruent trials (target emotion appears in both channels; visually depicted in Fig. 3 as type 1), prosody trials (target emotion appears only in the prosody; type 2 in Fig. 3), semantics trials (target emotion appears only in the semantics; type 3 in Fig. 3) and target-emotion-absent trials (i.e., the target emotion appears in neither the semantics nor the prosody; type 4 in Fig. 3). General rating task performance was analyzed.

As the order of the target emotions (3 X 2) and tasks (2) were fully counterbalanced across participants and groups, they were not included in the analyses. Supplemental Appendix B, http://links.lww.com/EANDH/B2 presents the data analyzed in each task, divided per emotion rating block for each group separately.

T-RES performance was tested against individual differences in speech perception measures for the CI group. For this purpose, we used the self-reported rating of hearing experience in phone conversations (converted to a binary scale: existence or absence of reported difficulties) and linguistic intonation (perfect performance, or not). As the HAB-Q and HAB-N were not found to generate any significant effect on any of the tests, they will not be further discussed. The effect of Hearing Aid (HA) use—bimodal (CI and HA) versus CI-only use (one or two implants)—was tested as well for T-RES performance. For the general rating task, we followed the ANOVA with a regression analysis focusing on incongruent stimuli, trying to directly gauge the different weights assigned to the prosodic and semantic channels by the two groups. Partial eta squared (η_p^2) was used as the measure for effect size in all statistically significant tests.

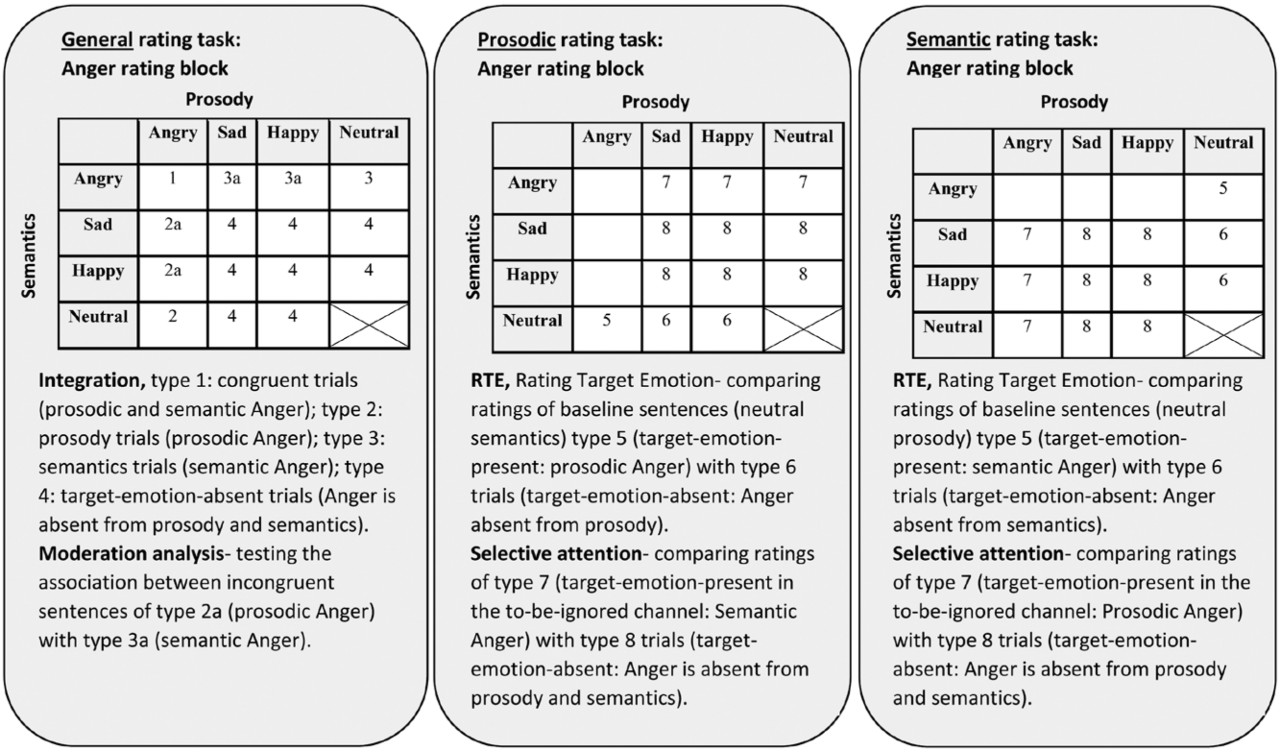
RESULTS

Individual Characteristics

Table 2 compares characteristics of participants in the CI and NH groups. It is noteworthy that although the two groups were matched on age, gender, level of education and cognitive skills (auditory forward digit span), they clearly differed in their performance on auditory measures (HAB-N score, HeSPAC intonation score).

The findings related to the study's research questions were as follows:

(1) Rating of target emotions (RTE): Do both groups rate semantic and prosodic emotions in a similar way?



General rating task: Anger rating block

Prosody / Semantics

	Angry	Sad	Happy	Neutral
Angry	1	3a	3a	3
Sad	2a	4	4	4
Happy	2a	4	4	4
Neutral	2	4	4	

Integration, type 1: congruent trials (prosodic and semantic Anger); type 2: prosody trials (prosodic Anger); type 3: semantics trials (semantic Anger); type 4: target-emotion-absent trials (Anger is absent from prosody and semantics).
Moderation analysis- testing the association between incongruent sentences of type 2a (prosodic Anger) with type 3a (semantic Anger).

Prosodic rating task: Anger rating block

Prosody / Semantics

	Angry	Sad	Happy	Neutral
Angry		7	7	7
Sad		8	8	8
Happy		8	8	8
Neutral	5	6	6	

RTE, Rating Target Emotion- comparing ratings of baseline sentences (neutral semantics) type 5 (target-emotion-present: prosodic Anger) with type 6 trials (target-emotion-absent: Anger absent from prosody).
Selective attention- comparing ratings of type 7 (target-emotion-present in the to-be-ignored channel: Semantic Anger) with type 8 trials (target-emotion-absent: Anger is absent from prosody and semantics).

Semantic rating task: Anger rating block

Prosody / Semantics

	Angry	Sad	Happy	Neutral
Angry				5
Sad	7	8	8	6
Happy	7	8	8	6
Neutral	7	8	8	

RTE, Rating Target Emotion- comparing ratings of baseline sentences (neutral prosody) type 5 (target-emotion-present: semantic Anger) with type 6 trials (target-emotion-absent: Anger absent from semantics).
Selective attention- comparing ratings of type 7 (target-emotion-present in the to-be-ignored channel: Prosodic Anger) with type 8 trials (target-emotion-absent: Anger is absent from prosody and semantics).

Fig. 3. General design of the Test for Rating Emotions in Speech (T-RES): Tested effects. Visual depiction of the trials compared in the separate statistical analyses, with examples from the Anger rating blocks in the three tasks.

Results related to this research question are depicted in the top rows of Table 4. A 2 X 2 X 3 X 2 mixed model repeated measures ANOVA was conducted, with RTE (X2: target emotion present vs. absent), rated channel (X2: prosodic versus semantic rating) and target emotion (X3: anger, happiness, or sadness) as within-participants variables, and group membership (X2: CI versus NH) as a between-participants variable. The complete analysis is provided in Supplemental Appendix C, http://links.lww.com/EANDH/B2.

Mainly, the analysis found a significant main effect for RTE, $F(1, 28) = 399.0$, $p < 0.001$, $\eta_p^2 = 0.93$. This effect indicates that, in general, listeners were able to easily identify the presented emotion. In answer to the first research question, RTE did not interact with group membership, $F(1, 28) = 1.93$, $p = 0.18$. This indicates that the CI and NH groups did not differ significantly in their ability to rate emotions in general (i.e., provide higher rates when the target emotion was present).

(2) Selective attention: is there a group difference in selective attention to the prosodic or semantic channels?

Results related to this research question are depicted in the middle rows of Table 4. A 2 X 2 X 3 X 2 mixed model repeated measures ANOVA was conducted with selective attention (X2: target-emotion-present or target-emotion-absent in the to-be-ignored channel), rated channel (X2: prosodic vs. semantic rating), and target emotion (X3: anger, happiness, or sadness) as within-participants variables, and group membership (X2: CI or NH) as a between-participants variable. The complete analysis is provided in Supplemental Appendix C, http://links.lww.com/EANDH/B2.

Results showed a significant main effect for selective attention, indicating failures of selective attention for both groups across all emotions, $F(1, 28) = 8.3$, $p = 0.007$, $\eta_p^2 = 0.23$. However, in answer to the second research question, no significant interaction between (failures of) selective attention and group was found, $F(1, 28) = 0.35$, $p = 0.56$, suggesting that the CI and NH groups did not differ significantly in their ability to selectively attend to emotional auditory channels.

We note that a significant interaction of selective attention with the target channel factor was noted, $F(1, 28) = 27.1$, $p < 0.001$, $\eta_p^2 = 0.49$. This indicates that significant failures were evident only when listeners were asked to inhibit the prosody, $F(1, 28) = 23.1$, $p < 0.001$, $\eta_p^2 = 0.45$ (*Mean difference* = 0.52), but not when they were asked to inhibit the semantic, $F(1, 28) = 0.4$, $p = 0.52$ (*Mean difference* = −0.05). However, pertinent to the second research question, no significant interactions of selective attention and group membership were found when listeners were asked to inhibit the prosody, $F(1, 28) = 0.09$, $p = 0.77$, or the semantics, $F(1, 28) = 2.2$, $p = 0.15$. To conclude, the two groups did not differ in the degree of failures to selectively inhibit the prosody while attending to the semantics.

(3) Integration of channels and channel bias: Is there a difference in the weights assigned to prosodic and semantic channels between CI user and NH control participants?

Results related to this research question are depicted in the bottom rows of Table 4, and graphically depicted in Figure 4. A 1 X 3 X 2 mixed model repeated measures ANOVA was conducted in which the linear trend (X1: congruent > prosody > semantics) was tested with the target emotion (X3: anger, happiness, or sadness) as a within-participants variable, and group membership (X2: CI versus NH) as a between-participants variable. We also included planned comparisons of possible

TABLE 4. Summary of results (means and standard errors), averaged across target emotions

	CI Users		NH Controls		Group Interaction
Rating of target emotions, RTE (prosodic rating and semantic rating of baseline sentences)					
	Prosody	Semantics	Prosody	Semantics	
Target-emotion-present	4.9 (.19)	5.5 (.14)	5.0 (.19)	5.8 (.14)	
Target-emotion-absent	2.6 (.22)	1.7 (.17)	1.9 (.22)	1.9 (.17)	
RTE:					*F*(1, 28) = 1.9, *p* = 0.18
Selective attention (prosodic rating and semantic rating)					
	Prosody	Semantics	Prosody	Semantics	
Target emotion in the to-be-ignored channel	2.7 (.18)	2.4 (.19)	1.6 (.18)	2.0 (.19)	
Target-emotion-absent	2.6 (.14)	1.9 (.12)	1.8 (.14)	1.5 (.12)	
Selective attention					*F*(1, 28) = 0.35, *p* = 0.56
Integration (general rating)					
Congruent sentence	5.2 (.15)		5.6 (.15)		*F*(1, 28) = 2.93, *p* = 0.1
Prosodic sentences	4.2 (.21)		4.5 (.20)		*F*(1, 28) = 1.79, *p* = 0.2
Semantics sentences	3.4 (.26)		2.7 (.25)		*F*(1, 28) = 5.33, *p* = 0.029, η_p^2 = 0.16
Target-emotion-absent	2.2 (.14)		1.7 (.14)		*F*(1, 28) = 7.36, *p* = 0.011, η_p^2 = 0.21

group differences for each trial type (congruent, prosody, and semantics).

Mainly, analyses indicate a significant linear trend across groups (post-hoc analyses show that the linear trend was significant in each separate group, CI: $F(1, 14) = 33.76$, $p < 0.001$, $\eta_p^2 = 0.71$, and NH: $F(1, 14) = 109.03$, $p < 0.001$, $\eta_p^2 = 0.89$), $F(1, 28) = 177.89$, $p < 0.001$, $\eta_p^2 = 0.82$, that interacted significantly with group, $F(1, 28) = 7.19$, $p = 0.012$, $\eta_p^2 = 0.20$, with no main effect for group, $F(1, 28) = 0.05$, $p = 0.83$. Planned comparisons revealed the source of the interaction: in semantics trials, the CI group provided higher ratings than did the NH group, $F(1, 28) = 5.33$, $p = 0.029$, $\eta_p^2 = 0.16$. Group differences were not significant for congruent trials, $F(1, 28) = 2.93$, $p = 0.10$, nor for prosodic trials, $F(1, 28) = 1.79$, $p = 0.19$, $\eta_p^2 = 0.49$.

To answer the third research question, both groups exhibited the expected linear trend: congruent trials received the highest emotional ratings, followed by prosody trials, and then semantics trials (see Figure 4). However, the CI group provided higher weights for the semantics than did the NH group.

A separate analysis for target-emotion-absent trials indicated significantly higher ratings for the CI compared with the NH group, $F(1, 28) = 7.36$, $p = 0.011$, $\eta_p^2 = 0.21$. In other words, the CI group provided ratings indicating the presence of an emotion when, in fact, it was absent from both prosodic and semantic channels. This suggests some degree of confusion between emotions for the CI group.

Note, in one of the analyses an interaction with target-emotion was found. Yet, it did not affect the general pattern (see details in Supplemental Appendix D, http://links.lww.com/EANDH/B2).

In the next stage, we focused on incongruent sentences in the general rating task, comparing the ratio of prosodic to semantic ratings in the two groups. This, in order to gauge the different weights assigned to the prosody and semantics (visually depicted in Figure 3, as type “2a” and “3a” trials, respectively). The data show a higher ratio for the NH group than the CI group (ratios of 2.11, $SE = 0.37$, and 1.32, $SE = 0.13$, respectively), $t(28) = 2.02$, $p = 0.05$. This indicates a larger bias for prosody over semantics for the NH than the CI group. In fact, the correlation between prosodic and semantic ratings of incongruent sentences was also higher for the NH than the CI group, $Z = 1.724$, $p = 0.04$. In sum, the analyses suggest that NH participants rate incongruent emotions mainly based on one channel—prosody—whereas CI participants showed a lower bias for prosody.

(4) Individual Characteristics: Does CI users’ T-RES performance relate to their speech perception functions and HA status?

RTE • Results are visually depicted in Figure 5. An interaction of semantic RTE and self-reported hearing experience over the phone was observed, $F(1, 13) = 4.5$, $p = 0.05$, $\eta_p^2 = 0.26$. This interaction was not significant for prosodic ratings, $F(1, 13) = 0.04$, $p = 0.56$. When testing linguistic intonation scores, an interaction was found with prosodic RTE, $F(1, 13) = 6.4$, $p = 0.02$, $\eta_p^2 = 0.33$, and not with semantic RTE, $F(1, 13) = 0.46$, $p = 0.5$. None of the other factors led to significant interactions. HA use did not lead to a significant interaction, $F(1,13) = 0.29$, $p = 0.6$.

To answer the fourth research question with respect to RTE, participants in the CI group who reported having difficulties understanding speech over the phone provided less discriminable semantic ratings than their CI peers, but similar prosodic ratings. The reverse pattern was observed for participants in the CI group with impaired identification of linguistic intonation; they provided less discriminable prosodic ratings than did their CI peers, but similar semantic ratings.

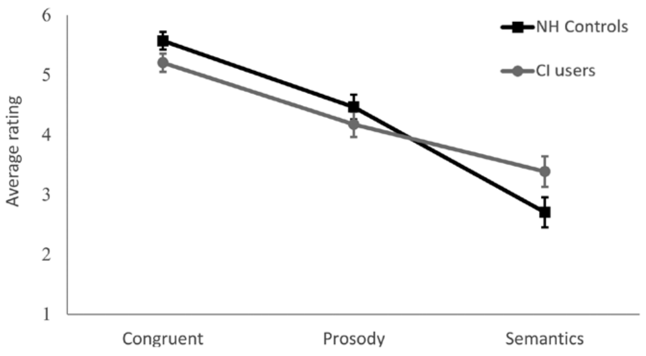


Fig. 4. Graphic depiction of ratings in the general rating task for cochlear implant (CI) users (gray line) and normal hearing (NH) controls (black line). Ratings are averaged for congruent trials (target emotion appears in both channels), prosody trials (target emotion appears only in prosody); semantics trials (target emotion appears only in semantics). These trials are visually depicted in Figure 1C sas “types 1”, “type 2”, and “type 3”, respectively. Error bars represent standard errors.

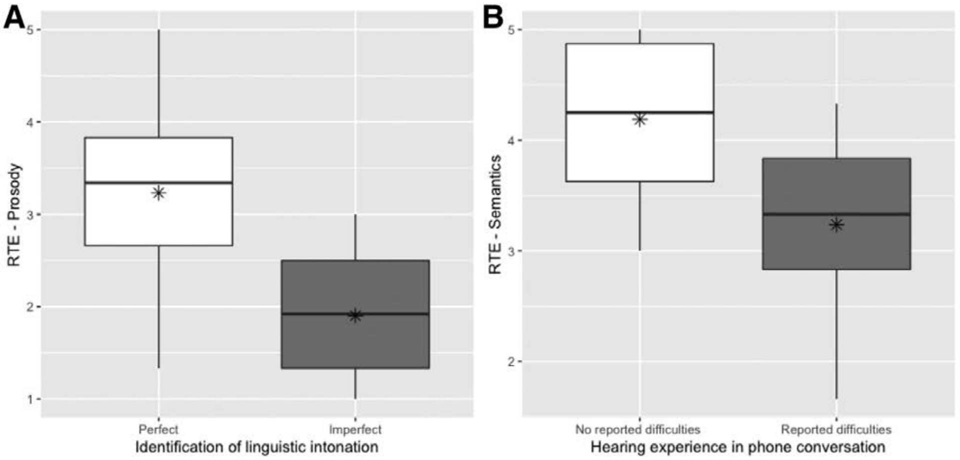


Fig. 5. Box plots depicting the effects of individual characteristics in the cochlear implant group on Rating of Target Emotion (RTE). A: Average RTE prosody with identification of linguistic information; B: RTE semantics with reported hearing experience in phone conversation.

Selective attention • To answer the fourth research question with respect to selective attention, performance for the CI group was not affected by individual characteristics (self-reported experience regarding hearing over the phone and identification of linguistic intonation), $F < 2.0$, $p > 0.18$, nor by HA use, $F(1,13) = 0.357$, $p = 0.561$.

Integration • A significant interaction of ratings of target-emotion-absent trials and linguistic intonation was found, $F(1, 13) = 5.71$, $p = 0.03$, $\eta_p^2 = 0.31$, indicating lower ratings provided by CI users with better linguistic intonation. In fact, their ratings were not significantly different from those given by the NH group, $F(1,18) = 0.02$, $p = 0.89$. No other significant interactions of linguistic intonation were found for the congruent- prosodic- and semantic trials, for the analysis of other individual characteristics, $F < 2.0$, $p > 0.2$, or for the effect of HA use, $F(1,13) < 1.1$, $p > 0.3$.

To answer the fourth research question, the integration of prosodic and semantic channels was not affected by individual characteristics. Though, the confusion in emotional rating in CI users may be related to difficulties in identifying prosodic cues.

DISCUSSION

In the current study, a group of postlingual CI users and a matched group of NH control participants were asked to rate the type and extent of emotion presented in the semantic and prosodic channels of spoken sentences. The study's findings highlight the following main trends:

1. *Rating of target emotions (RTE)*—both groups similarly rated emotions separately presented in the prosodic and semantic channels when no competitive information was presented in the other channel.
2. *Failure of selective attention*—no differences between CI users and matched NH participants were found in selective attention to the prosodic or to the semantic channel.
3. *Integration of channels and channel bias*—when two different emotions were presented via the semantic and the prosodic channels (incongruent trials), NH participants based their ratings mainly on the prosody. However, CI users were not biased in their ratings to either of the two channels, proving higher relative weights to semantics than NH participants did. In the CI group, some measure of confusion was indicated by responses suggesting the presence of emotions that were absent.
4. *Individual characteristics for CI users*—self-reports of difficulties in speech comprehension over the phone were negatively related to semantic ratings on the T-RES, while Linguistic Intonation test scores were positively related to rating of prosodic emotions on the T-RES and to the level of confusion in rating emotions on the general rating task.

RTE: Do Both Groups Rate Semantic and Prosodic Emotions in a Similar Way?

The current study's data suggest no group differences in rating of spoken emotions when there was no need to inhibit or integrate across channels. It should be noted that the current study's cohort of postlingual CI users developed typical language skills before deafness occurred via intact peripheral and central auditory systems. Moreover, gaining experience with the implant, they may have adopted efficient acoustic cue-weighting strategies with the available information transmitted by the CI. Thus, they achieved successful identification of prosodic cues, especially in relatively simple tasks.

As mentioned in the introduction, the small number of studies testing the identification of emotional prosody in postlingual CI users have suggested reduced identification (Agrawal et al. 2013; Chatterjee et al. 2015; Gilbers et al. 2015; Luo et al. 2007). These identification difficulties were related to limited pitch information delivered by the CI device, and the necessity of relying on different acoustic cues (Everhardt et al. 2020). The seemingly inconsistent findings of the current study may be understood in light of the different methodologies used. For example, previous studies used a forced choice paradigm, in which listeners selected the matching emotion from a set list of emotions. This restricted mode of response simplifies analysis, but provides a limited vantage point, as no information can be obtained about the options rejected by the listener. For example, with a forced choice response set, a listener may choose "sad" for an utterance spoken with angry prosody. But, on a rating scale, s/he may rate the same sentence as 4/6 on "anger" and 4/6 on "sadness." This more detailed information is lost using the forced choice methodology. Accordingly, the current study findings suggest that when using a rating scale in a laboratory-based study, CI users can indeed rate prosodic emotions in quiet just as well as their matched control NH peers. Future studies should further compare performance on a rating scale with performance on a forced-choice response set, as group-differences in response accuracy may ensue, even when both groups provide similar rating.

Selective Attention: Is There a Group Difference in Selective Attention to the Prosodic or Semantic Channels?

In the current study, a similar degree of failure of selective attention in the auditory modality was found for both groups. To date, the literature on selective attention in postlingual CI users has been inconsistent and mostly focused on the visual modality. Focusing on auditory selective attention, Henkin et al. (2014) failed to find a difference between postlingual CI users, aged 60-77 years, and age-matched NH controls. The current findings generalize these results to younger age groups.

The finding of preserved abilities, at least in some measures of auditory selective attention, is of potential clinical importance given previous evidence of a general decrease in executive functions in CI users (Moberly et al. 2018). This finding has added importance and external validity due to the study's methodology, which utilized contextually engaging stimuli – meaningful sentences, rather than single words (Ben-David et al. 2018; Cohen-Zimerman & Hassin 2018). Indeed, rehabilitation efforts for postlingual CI users may choose to use this preserved cognitive-auditory ability to train speech perception performance in adverse listening conditions. This may be of special importance, as selective attention performance was found to correlate with speech recognition in noise for post-lingual CI users (Moberly et al., 2016).

Integration of Channels and Channel Bias: Do Both Groups Similarly Assign Weights to Prosodic and Semantic Channels?

When asked to listen to the sentence as a whole, CI users appeared to over-use semantics, whereas NH participants rated these sentences mainly based on the prosody. It is interesting to compare the current results with those observed in other populations characterized by auditory difficulties (i.e., older adults and people with tinnitus) using the same T-RES paradigm (Ben-David et al. 2019; Oron et al. 2020). Across the three studies, when participants were asked to rate prosody alone (prosodic rating task), performance did not differ between the target and the control groups. However, when asked to integrate both channels (general rating task), the target groups gave higher relative weights to semantics. Taken together, it may be that the processing of prosodic cues calls for more cognitive resources than does the processing of semantic cues. This is particularly true when auditory input is distorted even if the source and the type of distortion is very different (Schneider and Pichora-Fuller 2000; Ben-David et al. 2011a, 2014). This imbalance in resource demands may explain why CI users over-emphasize the semantic channel.

Indeed, the literature suggests that the processing of prosodic emotions for CI users is especially challenging. After deafness, and following CI implantation, some acoustic cues (especially spectral cues) are degraded, and less available via the device. With time, postlingual CI users adapt their acoustic cue-weighting strategies to the available acoustic information (Winn et al. 2012). It may be that before deafness, such individuals could rely on F0 information, while after deafness and implantation, these cues became less available (Everhardt et al. 2020). For example, Peng et al. (2012) found that postlingual CI users relied more heavily on intensity cues (better represented by the CI) to identify prosody, as compared to NH adults who relied on F0 contours. Of particular interest is the study by O'Neill et al. (2019), which found that postlingual CI users make more use of semantic context in speech processing than do NH listeners, possibly due to life experience and reduced acoustic information. In sum, the current study findings suggest that postlingual CI users may compensate for the lack of acoustic information by overusing emotional semantics

Individual Characteristics: Does CI Users' T-RES Performance Relate to Their Speech Perception Functions and HA Status?

The effects of individual characteristics on T-RES performance in CI users support the external validity of the T-RES as a gauge for emotional speech processing. Indeed, T-RES prosodic and semantic ratings were related to performance on respective standard tests (linguistic prosody, comprehension of phone conversations). The analysis of individual characteristics also hints that the source for confusion between emotions (in general ratings) in CI users relates to identification of intonation. The latter analysis may explain why CI users overuse the semantic cues, suggesting a compensatory mechanism.

Comparisons of bimodal (CI and HA) and CI-only users show somewhat surprising results. The literature suggests that bimodal users may receive better temporal fine structure (TFS) cues from residual acoustic hearing in the low frequencies and compared with CI-only users (Shpak et al. 2014). This might suggest improved performance in prosodic identification. However, our data did not reveal any significant difference in T-RES performance between these two sub-groups of CI users. Note, this finding is not new: Cullington and Zeng (2011) did not find a difference in affective prosody discrimination between bimodal and CI-only users. It appears that CI users exploit the full spectrum of prosodic cues in speech processing, including pitch, rate, intensity and stress, among others.

Limitations and Future Studies

The current study has a few possible limitations. The group of CI users is varied on several background characteristics that may influence the processing of spoken emotions. In the current study, we tried to minimize variance by choosing only native Hebrew speakers, postlingual CI users, with education levels equivalent to that of their NH peers. We also acknowledge that CI users were varied with respect to HA use, yet it appears that no significant effect was found for the latter (bimodal versus CI-only). Another possible limitation is the relatively small number of participants. However, it is not smaller than that found in the pertinent literature (e.g., N = 8, Luo et al.2007). Indeed, in a recent meta-analysis, Everhardt et al. (2020) commented that "CI research is typically limited in number of participants" (p. 1092 there). Future studies may wish to tackle the abovementioned issues directly, with a larger group of CI (both bimodal and CI-only) users, controlling for background variability.

A few limitations also relate to the T-RES paradigm itself. First, the sentences were recorded by a professional female actress, rather than several unprofessional speakers. This may harm the generalizability of the test, but it minimizes confounding factors. Second, the T-RES includes basic and concrete emotions; differences may be further inflated in more abstract or complex emotions such as envy (Icht et al. 2021). Third, the T-RES is a laboratory measure. As such, it uses material designed for this purpose—for example, the discrepancy between the emotional category of prosody and semantics presented in incongruent sentences. Fourth, in order to directly gauge the separate weights assigned to each channel in incongruent prosody-semantics sentences, the T-RES does not include a scale corresponding to humor, irony, or sarcasm (i.e., a discrepancy between an expected and presented emotion). Fifth, one may also wish to assess performance in more realistic conditions, for example by embedding speech in noise (Mama et al. 2018). We hope that future studies with larger samples will be able to address these factors.

Future studies should also consider comparing performance of postlingual and prelingual CI users in order to assess the specific role of acquiring acoustic cue weighting strategies from intact hearing before hearing loss.

Summary and Recommendations

The current study found preserved skills for rating of prosodic and semantic emotional information in postlingual CI users, as well as intact selective attention abilities. However, when asked to integrate prosody and semantics, CI users overrelied on semantic information relative to NH peers. This group difference may relate to the differentially larger degradation of prosodic (rather than semantic) acoustic cues by CIs. The results of the current study may assist in developing new strategies for rehabilitation of postlingual CI users, improving their social interaction with the hearing population at large. Future development of CIs should aim to better provide valuable prosodic information.

ACKNOWLEDGMENTS

We would like to thank Ms. Maya Mentzel, Ms. Efrat Ben Tora, and Ms. Efrat Galinsky for their assistance on this project.

The authors have no conflict of interest to disclose.

Address for correspondence: Boaz M. Ben-David, Communication Aging and Neuropsychology Lab (CANlab), Baruch Ivcher School of Psychology, Reichman University (IDC), 8 University st., Herzliya, Israel 4610101. E-mail: boaz.ben.david@idc.ac.il

Received May 16, 2020; accepted November 27, 2021

REFERENCES

Agrawal, D., Thorne, J. D., Viola, F. C., Timm, L., Debener, S., Büchner, A., Dengler, R., Wittfoth, M. (2013). Electrophysiological responses to emotional prosody perception in cochlear implant users. *Neuroimage Clin, 2*, 229–238.

Ben-David, B. M., & Algom, D. (2009). Species of redundancy in visual target detection. *J Exp Psychol Hum Percept Perform, 35*, 958–976.

Ben-David, B. M., Avivi-Reich, M., Schneider, B. A. (2016a). Does the degree of linguistic experience (native versus nonnative) modulate the degree to which listeners can benefit from a delay between the onset of the maskers and the onset of the target speech? *Hear Res, 341*, 9–18.

Ben-David, B. M., Ben-Itzchak, E., Zukerman, G., Yahav, G., Icht, M. (2020). The perception of emotions in spoken language in undergraduates with high functioning autism spectrum disorder: A preserved social skill. *J Autism Dev Disord, 50*, 741–756.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *J Speech Lang Hear Res, 54*, 243–262.

Ben-David, B. M., Eidels, A., Donkin, C. (2014). Effects of aging and distractors on detection of redundant visual targets and capacity: Do older adults integrate visual targets differently than younger adults? *PLoS One, 9*, e113551.

Ben-David, B. M., Gal-Rosenblum, S., van Lieshout, P. H. H. M., Shakuf, V. (2019). Age-related differences in the perception of emotion in spoken language: The relative roles of prosody and semantics. *J Speech Lang Hear Res, 62*, 1188–1202.

Ben-David, B. M., Malkin, G., Erel, H. (2018). Ageism and neuropsychological tests. In: L. Ayalon, & C. Tesch-Römer (Eds.), *Contemporary Perspectives on Ageism. International Perspectives on Aging* (pp. 279–297). Springer.

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., van Lieshout, P. H. (2016b). Prosody and semantics are separate but not separable channels in the perception of emotional speech: Test for rating of emotions in speech. *J Speech Lang Hear Res, 59*, 72–89.

Ben-David, B. M., & Schneider, B. A. (2009). A sensory origin for aging effects in the color-word Stroop task: An analysis of studies. *Aging Neuropsychol Cogn, 16*, 505–534.

Ben-David, B. M., & Schneider, B. A. (2010). A sensory origin for aging effects in the color-word Stroop task: Simulating age-related changes in color-vision mimic age-related changes in Stroop. *Aging Neuropsychol Cogn, 17*, 730–746.

Ben-David, B. M., Thayapararajah, A., van Lieshout, P. H. (2013). A resource of validated digital audio recordings to assess identification of emotion in spoken language after a brain injury. *Brain Inj, 27*, 248–250.

Ben-David, B. M., van Lieshout, P. H., Leszcz, T. (2011). A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Inj, 25*, 206–220.

Blamey, P., Artieres, F., Başkent, D., et al. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiol Neurootol, 18*, 36–47.

Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer (ver 6.1.07) [Computer program].

Breitenstein, C., Lancker, D. V., Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cogn Emot, 15*, 57–79.

Bryant, G., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *J Cogn Cult, 8*, 135–148.

Carl, M., Icht, M., Ben-David, B. M. (2022). A cross-linguistic validation of the Test for Rating Emotions in Speech (T-RES): Acoustic analyses of emotional sentences in English, German, and Hebrew. *J Speech Lang Hear Res*. [Epub ahead of print].

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

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hear Res, 322*, 151–162.

Cohen-Zimerman, S., & Hassin, R. R. (2018). Implicit motivation improves executive functions of older adults. *Conscious Cogn, 63*, 267–279.

Cullington, H. E., & Zeng, F. G. (2011). Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. *Ear Hear, 32*, 16–30.

De Raeve, L., Vermeulen, A., Snik, A. (2015). Verbal cognition in deaf children using cochlear implants: Effect of unilateral and bilateral stimulation. *Audiol Neurootol, 20*, 261–266.

Dorman, M. F., & Gifford, R. H. (2017). Speech understanding in complex listening environments by listeners fit with Cochlear Implants. *J Speech Lang Hear Res, 60*, 3019–3026.

Dupuis, K., & Pichora-Fuller, M. K. (2014). Intelligibility of emotional speech in younger and older adults. *Ear Hear, 35*, 695–707.

Everhardt, M. K., Sarampalis, A., Coler, M., Başkent, D., Lowie, W. (2020). Meta-analysis on the identification of linguistic and emotional prosody in cochlear implant users and vocoder simulations. *Ear Hear, 41*, 1092–1102.

Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., Runge-Samuelson, C. L., Wackym, P. A. (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear Hear, 25*, 375–387.

Gifford, R. H., Shallop, J. K., Peterson, A. M. (2008). Speech recognition materials and ceiling effects: Considerations for cochlear implant programs. *Audiol Neurootol, 13*, 193–205.

Gilbers, S., Fuller, C., Gilbers, D., Broersma, M., Goudbeek, M., Free, R., & Başkent, D. (2015). Normal-hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech. *I-perception, 6*, 1–19.

Hadar, B., Skrzypek, J. E., Wingfield, A., Ben-David, B. M. (2016). Working memory load affects processing time in spoken word recognition: Evidence from eye-movements. *Front Neurosci, 10*, 221.

Harel-Arbeli, T., Wingfield, A., Palgi, Y., Ben-David, B. M. (2021). Age-related differences in the online processing of spoken semantic context and the effect of semantic competition: Evidence from eye gaze. *J Speech Lang Hear Res, 64*, 315–327.

Heinrich, A., Gagné, J. P., Viljanen, A., Levy, D. A., Ben-David, B. M., Schneider, B. A. (2016). Effective communication as a fundamental aspect of active aging and well-being: Paying attention to the challenges older adults face in noisy environments. *Soc. Inq. into Well-Being, 2*, 51–68.

Henkin, Y., Yaar-Soffer, Y., Steinberg, M., Muchnik, C. (2014). Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. *Audiol Neurootol, 19*(Suppl 1), 21–26.

Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., Gotter, B. D., Vanderhoof, S. S., Mispagel, K., Heydebrand, G., Skinner, M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear, 34*, 342–360.

Hudepohl, M. B., Robins, D. L., King, T. Z., Henrich, C. C. (2015). The role of emotion perception in adaptive functioning of people with autism spectrum disorders. *Autism, 19*, 107–112.

Icht, M., Zukerman, G., Ben-Itzchak, E., Ben-David, B. M. (2021). Keep it simple: Identification of basic versus complex emotions in spoken language in individuals with autism spectrum disorder without intellectual disability: A meta-analysis study. *Autism Res, 14*, 1948–1964.

Jaekel, B. N., Newman, R. S., Goupell, M. J. (2017). Speech rate normalization and phonemic boundary perception in cochlear-implant users. *J Speech Lang Hear Res, 60*, 1398–1416.

Ji, C., Galvin, J. J. 3rd, Xu, A., Fu, Q. J. (2013). Effect of speaking rate on recognition of synthetic and natural speech by normal-hearing and cochlear implant listeners. *Ear Hear, 34*, 313–323.

Kalathottukaren, R. T., Purdy, S. C., Ballard, E. (2015). Prosody perception and musical pitch discrimination in adults using cochlear implants. *Int J Audiol, 54*, 444–452.

Kelly, A. S., Purdy, S. C., Thorne, P. R. (2005). Electrophysiological and speech perception measures of auditory processing in experienced adult cochlear implant users. *Clin Neurophysiol, 116*, 1235–1246.

Kishon-Rabin, L., Taitelbaum, R., Muchnik, C., Gehtler, I., Kronenberg, J., Hildesheimer, M. (2002). Development of speech perception and production in children with cochlear implants. *Ann Otol Rhinol Laryngol Suppl, 189*, 85–90.

Knight, S., & Heinrich, A. (2018). Visual Inhibition Measures Predict Speech-in-Noise Perception Only in People With Low Levels of Education. *Front Psychol, 9*, 2779.

Laukka, P. (2003). Categorical perception of emotion in vocal expression. *Ann N Y Acad Sci, 1000*, 283–287.

Lazard, D. S., Vincent, C., Venail, F., et al. (2012). Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: A new conceptual model over time. *PLoS One, 7*, e48739.

Leshem, R., Icht, M., Bentzur, R., Ben-David, B. M. (2020). Processing of emotions in speech in forensic patients with schizophrenia: Impairments in identification, selective attention, and integration of speech channels. *Front Psychiatry, 11*, 601763.

Loveland, K. A., Tunali-Kotoski, B., Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A., Gibbs, M. C. (1997). Emotion recognition in autism: Verbal and nonverbal information. *Dev Psychopathol, 9*, 579–593.

Xin Luo, Fu, Q. J., Galvin, J. J. 3rd. (2007). Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends Amplif, 11*, 301–315.

Mama, Y., Fostick, L., Icht, M. (2018). The impact of different background noises on the production effect: Evidence for costs and benefits in free recall. *Acta Psychol, 185*, 235–242.

Mason, M., & Kokkinakis, K. (2014). Perception of consonants in reverberation and noise by adults fitted with bimodal devices. *J Speech Lang Hear Res, 57*, 1512–1520.

Meister, H., Landwehr, M., Pyschny, V., Walger, M., von Wedel, H. (2009). The perception of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients. *Int J Audiol, 48*, 38–48.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychol Rev, 110*, 422–471.

Moberly, A. C., Houston, D. M., Castellanos, I. (2016). Non-auditory neurocognitive skills contribute to speech recognition in adults with cochlear implants. *Laryngoscope Investig Otolaryngol, 1*, 154–162.

Moberly, A. C., Mattingly, J. K., Castellanos, I. (2019). How does nonverbal reasoning affect sentence recognition in adults with cochlear implants and normal-hearing peers? *Audiol Neurootol, 24*, 127–138.

Moberly, A. C., Patel, T. R., Castellanos, I. (2018). Relations between self-reported executive functioning and speech perception skills in adult cochlear implant users. *Otol Neurotol, 39*, 250–257.

Nitsan, G., Wingfield, A., Lavie, L., Ben-David, B. M. (2019). Differences in working memory capacity affect online spoken word recognition: Evidence from eye-movements. *Trends Hear, 23*, 1–12.

O'Neill, E. R., Kreft, H. A., Oxenham, A. J. (2019). Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *J Acoust Soc Am, 146*, 195.

Oron, Y., Levy, O., Avivi-Reich, M., Goldfarb, A., Handzel, O., Shakuf, V., Ben-David, B. M. (2020). Tinnitus affects the relative roles of semantics and prosody in the perception of emotions in spoken language. *Int J Audiol, 59*, 195–207.

Paquette, S., Ahmed, G. D., Goffi-Gomez, M. V., Hoshino, A. C. H., Peretz, I., Lehmann, A. (2018). Musical and vocal emotion perception for cochlear implants users. *Hear Res, 370*, 272–282.

Paulmann, S., Pell, M. D., Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain Lang, 104*, 262–269.

Pell, M. D., Monetta, L., Paulmann, S., Kotz, S. A. (2009). Recognizing emotions in a foreign language. *J of Nonverbal Behav, 33*, 107–120.

Peng, S. C., Chatterjee, M., Lu, N. (2012). Acoustic cue integration in speech intonation recognition with cochlear implants. *Trends Amplif, 16*, 67–82.

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.

Pichora-Fuller, M. K., Dupuis, K., Van Lieshout, P. (2016). Importance of F0 for predicting vocal emotion categorization. *J Acoust Soc Am, 140*, 3401.

Scherer, K. R., Banse, R., Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *J of Cross-Cultural Psychol, 32*, 76–92.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of Cognitive Aging II* (pp. 155–219). Erlbaum.

Shakuf, V., Gal-Rosenblum, S., van Lieshout, P. H. H. M., Ben-David, B. M. (2016). The psychophysics of aging: In emotional speech, older adults attend to semantic, while younger adults to prosody. In Proceedings of the 32nd Annual Meeting of the International Society for Psychophysics. *Proceedings of Fechner Day, 32*.

Shpak, T., Most, T., Luntz, M. (2014). Fundamental frequency information for speech recognition via bimodal stimulation: Cochlear implant in one ear and hearing aid in the other. *Ear Hear, 35*, 97–109.

Taitelbaum-Swead, R., Kishon-Rabin, L., Kaplan-Neeman, R., Muchnik, C., Kronenberg, J., Hildesheimer, M. (2005). Speech perception of children using Nucleus, Clarion or Med-El cochlear implants. *Int J Pediatr Otorhinolaryngol, 69*, 1675–1683.

Taitelbaum-Swead, R., & Fostick, L. (2017). Audio-visual speech perception in noise: Implanted children and young adults versus normal hearing peers. *Int J Pediatr Otorhinolaryngol, 92*, 146–150.

Wechsler, D. (1997). *WAIS-III, Wechsler adult intelligence scale: Administration and scoring manual*. Psychological Corporation.

Wilson, B. S. (2013). Toward better representations of sound with cochlear implants. *Nat Med, 19*, 1245–1248.

Wilson, B. S. (2015). Getting a decent (but sparse) signal to the brain for users of cochlear implants. *Hear Res, 322*, 24–38.

Wingfield, A. (2016). Evolution of models of working memory and cognitive resources. *Ear Hear, 37*(Suppl 1), 35S–43S.

Winn, M. B., Chatterjee, M., Idsardi, W. J. (2012). The use of acoustic cues for phonetic identification: Effects of spectral degradation and electric hearing. *J Acoust Soc Am, 131*, 1465–1479.

Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hear Res, 174*, 101–106.

Zupan, B., Neumann, D., Babbage, D. R., Willer, B. (2009). The importance of vocal affect to bimodal processing of emotion: Implications for individuals with traumatic brain injury. *J Commun Disord, 42*, 1–17.