

## Adults with cochlear implants can use prosody to determine the clausal structure of spoken sentences

Nicole M. Amichetti,<sup>1,a)</sup> Jonathan Neukam,<sup>2</sup> Alexander J. Kinney,<sup>1,b)</sup> Nicole Capach,<sup>2</sup> Samantha U. March,<sup>1,b)</sup> Mario A. Svirsky,<sup>2</sup> and Arthur Wingfield<sup>1,b)</sup>

<sup>1</sup>Department of Psychology, Brandeis University, Waltham, Massachusetts 02453, USA

<sup>2</sup>Department of Otolaryngology, New York University (NYU) Langone Medical Center, New York, New York 10016, USA

### ABSTRACT:

Speech prosody, including pitch contour, word stress, pauses, and vowel lengthening, can aid the detection of the clausal structure of a multi-clause sentence and this, in turn, can help listeners determine the meaning. However, for cochlear implant (CI) users, the reduced acoustic richness of the signal raises the question of whether CI users may have difficulty using sentence prosody to detect syntactic clause boundaries within sentences or whether this ability is rescued by the redundancy of the prosodic features that normally co-occur at clause boundaries. Twenty-two CI users, ranging in age from 19 to 77 years old, recalled three types of sentences: sentences in which the prosodic pattern was appropriate to the location of a clause boundary within the sentence (congruent prosody), sentences with reduced prosodic information, or sentences in which the location of the clause boundary and the prosodic marking of a clause boundary were placed in conflict. The results showed the presence of congruent prosody to be associated with superior sentence recall and a reduced processing effort as indexed by the pupil dilation. The individual differences in a standard test of word recognition (consonant-nucleus-consonant score) were related to the recall accuracy as well as the processing effort. The outcomes are discussed in terms of the redundancy of the prosodic features, which normally accompany a clause boundary and processing effort. © 2021 Acoustical Society of America.

<https://doi.org/10.1121/10.0008899>

(Received 24 May 2021; revised 4 November 2021; accepted 8 November 2021; published online 14 December 2021)

[Editor: Benjamin V Tucker]

Pages: 4315–4328

### I. INTRODUCTION

Cochlear implants (CI) have proven an effective option for younger and older adults whose severity of hearing loss is beyond the benefit from conventional hearing aids. In spite of factors, such as limited spectral information and tonotopic mismatch (Friesen *et al.*, 2001; Fu and Nogaki, 2005; Landsberger *et al.*, 2015; Svirsky *et al.*, 2015), post-lingually deafened CI users generally show good levels of speech intelligibility as measured by the accuracy in repeating phonemes, words, and short sentences in quiet (Budenz *et al.*, 2011; Friedland *et al.*, 2010; Roberts *et al.*, 2013).

Several issues emerge, however, when considering more complex speech material. One issue relates to the use of speech prosody as an aid in detecting the clausal structure of complex multi-clause sentences as can occur in everyday listening. *Prosody* is the generic term that includes the intonation pattern (pitch contour) of a sentence carried by the fundamental frequency ( $F_0$ ) of the voice, inter- and intra-word timing patterns, and lexical stress, which is itself a complex subjective variable encoded by changes in loudness

(amplitude), pitch, and syllabic duration (Beckman and Edwards, 1987; Lehiste, 1970; Nooteboom, 1997; Shattuck-Hufnagel and Turk, 1996).

Speech prosody is known to serve a number of functions in everyday speech communication. The pitch, vocal tension, and speaking rate can reveal the speaker emotion (Breitenstein *et al.*, 2010), the relative  $F_0$  can distinguish a male from a female speaker (Klatt and Klatt, 1990), word stress can signal the semantic focus of a sentence (Jackendoff, 1971; Morrett *et al.*, 2020), and a rising pitch across a sentence can indicate that an utterance is intended as a question rather than a statement (Lehiste, 1970). (A further discussion of the prosodic features and their functions can be found in Wagner and Watson, 2010.)

Although considerable acoustic information is delivered by the successful fitting of a CI, the reduced spectral information in the CI signal interferes notably with the perception of the pitch and pitch contour (cf. Holt and McDermott, 2013; Marx *et al.*, 2015; Oxenham and Kreft, 2014; Peng *et al.*, 2012). As a consequence, CI users can be less effective than normal hearing listeners in perceiving the speaker gender (Meister *et al.*, 2009) and speaker emotion (Everhardt *et al.*, 2020), both of which are heavily dependent on the vocal pitch. For the same reason, distinguishing between a question and a statement can present difficulties for CI users, albeit mitigated to some extent by the word stress and context (cf. Marx *et al.*, 2015; Meister *et al.*, 2009; Moore and Carlyon, 2005; Peng *et al.*, 2009).

<sup>a)</sup>Also at: Volen National Center for Complex Systems, Brandeis University, Waltham, MA 02453, USA. Electronic mail: [namich@brandeis.edu](mailto:namich@brandeis.edu)

<sup>b)</sup>Also at: Volen National Center for Complex Systems, Brandeis University, Waltham, MA 02453, USA.

Although all of the above aspects of speech prosody can contribute directly or indirectly to effective communication, our interest here is in the role of prosody as an aid to detect the clausal structure of a sentence as a step toward determining the meaning of the utterance (Engelhardt *et al.*, 2010; Hoyte *et al.*, 2009; Steinhauer *et al.*, 2010; Steinhauer *et al.*, 1999; Titone *et al.*, 2006). This ability takes on special importance when one goes beyond simple sentences to sentences with embedded or dependent clauses. For example, as one hears a sentence that begins, “Although the two friends pushed the car...” there is a temporary syntactic ambiguity as to whether there is a major clause boundary after the word *pushed* (e.g., “Although the two friends pushed, the car wouldn’t start”) or after the word *car* (e.g., “Although the two friends pushed the car, it wouldn’t start”).

In writing, such temporary syntactic ambiguities can be avoided by an inserted comma as illustrated in these examples. The present study focuses on spoken utterances in which the sentence-level prosody can signal the arrival of a major clause boundary by pitch contour, relative word stress, a brief pause at the clause boundary, and the lengthening of a clause-final word (Kjelgaard *et al.*, 1999; Shattuck-Hufnagel and Turk, 1996; Speer *et al.*, 1996). To the extent that these prosodic events can be detected, they can inform the listener about the syntactic structure of the sentence either to avoid (Nooteboom, 1997; Shattuck-Hufnagel and Turk, 1996; Steinhauer *et al.*, 1999) or quickly repair (Titone *et al.*, 2006) a syntactic ambiguity. In this regard, the reduced acoustic richness of the CI signal raises the question of whether CI users may have difficulty in using sentence prosody to detect syntactic clause boundaries within multi-clause sentences or whether the multiplicity of the prosodic features that mark a clause boundary may rescue this ability.

To answer this question, we used a cross-splicing technique (Garrett *et al.*, 1966; Wingfield *et al.*, 1992) to test the recall of conflicting-cue sentences. In these sentences, the prosodic marking of a clause boundary and the actual position of the clause boundary as defined by the lexical and syntactic (*lexico-syntactic*) content of the sentence occur at different points within the sentence. Prior research using this technique with normal hearing younger and older adults has shown a significant recall decrement for conflicting-cue sentences relative to sentences in which the prosodic marking of a clause boundary coincides with the lexico-syntactically defined clause boundary, as would normally be the case (Wingfield *et al.*, 1992).

If a CI user were unable to detect the prosody in such conflicting-cue sentences, one would expect recall to be similar to recall of sentences spoken with normal prosody. On the other hand, there is considerable redundancy in the prosodic cues that signal the presence of a clause boundary, including the co-occurrence of the previously cited pitch and stress cues, clause-final word-lengthening, and often a brief pause at the clause boundary. Although, as indicated, CI users have well-documented difficulties in perceiving

voice pitch, many CI users have a relatively normal ability to process temporal cues (Cosentino *et al.*, 2016; Hood *et al.*, 1987; Sagi *et al.*, 2009; Shannon, 1989; Tyler *et al.*, 1989). Thus, it is possible that CI users may show an influence of the speech prosody in determining the location of a clause boundary even with reduced accessibility of the full array of prosodic features. To the extent that this is the case, however, the attempt to resolve the conflict in conflicting-cue sentences may come at the cost of a greater effort than dealing with sentences in which the lexico-syntactically defined clause boundary and prosodic marking of the boundary position coincide.

To address this latter question, we made use of *pupillometry*: the measurement of task-evoked changes in dilation of the pupil of the eye. In addition to responding to changes in ambient light (Wang *et al.*, 2016) and emotional arousal (Kinner *et al.*, 2017; Kim *et al.*, 2000), there is now a deep literature reliably showing the dilation of the pupil of the eye as a person engages in a challenging perceptual or cognitive task. Pupillometry has resultantly received wide use as an objective, physiological index of the processing effort (see van der Wel and Steenbergen, 2018, and Zekveld *et al.*, 2018, for reviews).

An association between an increase in the pupil dilation and processing effort has been observed while individuals with normal hearing listen to degraded speech (Koelewijn *et al.*, 2012; Kuchinsky *et al.*, 2013; Kuchinsky *et al.*, 2014; Zekveld *et al.*, 2018; Zekveld *et al.*, 2011) or are asked to recall spoken sentences that increase in length and syntactic complexity (Just and Carpenter, 1993; Piquado *et al.*, 2010). Pupillometry has also been shown to be a sensitive index of the perceptual effort associated with speech processing by CI users (Winn and Moore, 2018). Pupillometry might, thus, address the question of whether a measure of the processing effort might supply converging evidence on the influence of sentence prosody on the boundary detection by CI users.

## A. The present experiment

In the experiment to be described, adult CI users were tested for the recall of sentences containing a major syntactic clause when the sentence had normal prosody or was computer-edited to place the prosodic marking of the clause boundary at a different position than the actual lexico-syntactically defined clause boundary. A third condition used sentences recorded with reduced prosody (as explained in Sec. II).

Two behavioral measures of the influence of the prosodic clause marking were obtained. The first measure was the accuracy of the sentence recall. For this measure, a significant influence of the prosody would result in poorer recall accuracy for the conflicting-cue sentences (those in which the lexico-syntactically defined clause boundary and prosodic cues, which ordinarily accompany a clause boundary, were in different positions within the sentence).

The second measure was an analysis of the participants’ reconstructions of the erroneously recalled conflicting-cue

sentences. In this case, the evidence for an influence of prosody in determining the location of a clause boundary would appear in the form of reconstructed sentences in which the prosodic marking of a clause boundary took precedence over the actual lexico-syntactically defined clause position.

Throughout each trial, a continuous recording of the pupil size was obtained, time-locked with the speech as it was being heard, while the participant was preparing to give his or her recall.

In conducting this experiment, attention was paid to the well-known variability among the CI users in their word recognition ability. This was estimated by the recognition accuracy for the consonant-nucleus-consonant (CNC) words (Lehiste and Peterson, 1959; Luxford, 2001). Of interest was the extent to which the individual differences in word recognition ability may affect the sentence recall and syntactic resolution across conditions.

II. METHODS

A. Participants

The participants were 22 CI users, 8 men and 14 women, ranging in age from 19 to 77 years old ( $M = 49.0$  years). Of the 22 participants, 17 had bilateral implants. Of the five participants who had unilateral implants, four had a profound hearing loss from 125 Hz to 8 kHz in their un-implanted ear as indexed by the pure tone thresholds, and one had a severe loss at 125–500 Hz, sloping to profound. The latter participants' low frequency thresholds were above the presentation level of the experimental stimuli, as will be described, such that a contribution to the  $F0$  prosodic cues would be unlikely.

The participants were tested with CNC-30 lists as a measure of the word recognition ability (Luxford, 2001). As is common among CI users, there was a wide range in the effectiveness of the word recognition with CNC-30 word recognition scores ranging from 20% to 78% correct ( $M = 50.23$ ). Because the CNC-30 recordings are of greater difficulty than the more commonly used CNC recordings (on average scores are 22% lower; Skinner *et al.*, 2006), they are less likely to result in ceiling effects, which is an advantage for the clinical research.

The participants' vocabulary knowledge was assessed with a 20-item version of the Shipley vocabulary test (Zachary, 1991). This is a written multiple-choice test in which the participant is asked to indicate which of six listed words mean the same or nearly the same as a given target word. The participants' scores ranged from 9 to 19 [ $M = 14.86$ ; standard deviation (SD) = 2.49].

A summary of the CI participants' demographic information is given in Table I, ordered by the participant age. The first two columns give the participants' age and sex, and the next three columns give their implant information. This is followed by the etiology of the hearing loss where known and the years of experience with their CIs. The final column gives their CNC-30 scores.

All of the participants reported themselves to be in good health with no known history of stroke, Parkinson's disease, or other neuropathology, which might interfere with their ability to perform the experimental task. Written informed consent was obtained from all of the participants, according to a protocol approved by the Institutional Review Boards of Brandeis University and the New York University School of Medicine.

TABLE I. The participant information.

Subject identification	Age (yr)	Sex	Manufacturer	Electrode L/R	Processor	Etiology	Experience L/R (yr)	CNC
1	19	F	Cochlear	CI24R CS/—	N7	Congenital	16.5/—	61%
2	21	M	MED-EL	C40+/-	Opus 2	Congenital	16.3/—	53%
3	22	M	Cochlear	CI512/CI24M	N7	Congenital	9.4/8.5	78%
4	25	F	Cochlear	CI24R/CI24RE	N7	Congenital	4.7/11	31%
5	26	F	Cochlear	CI24R CS/CI22M	N6	Unknown	15/24	66%
6	28	F	Cochlear	CI24/—	N6	Congenital	20.7/—	29%
7	29	M	Cochlear	N22/—	N6	Congenital	26.2/—	55%
8	30	F	Cochlear	CI24M/CI24RE	N6	Congenital	19.3/5.6	61%
9	35	M	AB	HR90K	Naida Q90	Unknown	13.1/6.7	42%
10	38	F	Cochlear	CI532/CI522	N6	Congenital	1.1/3.6	72%
11	60	F	Cochlear	CI24RE/CI512	N6	Unknown	12.9/8.9	48%
12	62	F	Cochlear	CI24RE	N6	Progressive	10.9/9.9	55%
13	62	F	Cochlear	CI532/L24 EAS	N7	Progressive	1/4.4	57%
14	63	M	MED-EL	Concerto	Sonnet	Progressive	4.2/3.3	20%
15	64	F	Cochlear	CI24R CS/CI24 CA	N7	Progressive	17.9/17.9	60%
16	65	F	Cochlear	—/CI24RE	N6	Unknown	—/5.8	46%
17	65	M	Cochlear	CI532	N7	Progressive	1.5/2.9	69%
18	69	F	Cochlear	CI532	Kanso	Meniere's	5.4/3.4	38%
19	70	F	AB	HR90K	Naida Q70	Unknown	21.2/10.2	55%
20	73	M	Cochlear	CI24RE	N7	Congenital	7.3/3.8	33%
21	75	M	AB	HR90K	Naida Q90	Unknown	10.4/12.8	33%
22	77	F	Cochlear	CI532	Kanso	Congenital	2.3/2.3	43%

**B. Stimuli**

Stimulus construction began with the creation of 36 pairs of 8- to 15-word sentences, each consisting of a dependent and an independent clause. The pairs were constructed such that they shared an identical word sequence but with the major clause boundary occurring at different points within the word sequence. This is illustrated by the first two examples shown in Table II. Here, it can be seen that the first two sentences share the same seven-word sequence, “romantic lighting the candle on the table,” which is highlighted in yellow in sentence (1) and green in sentence (2). However, because of the different sentence frames in which the shared word sequence is embedded (the unhighlighted segments that precede and follow the highlighted sequences), in sentence (1) the lexico-syntactically defined clause boundary (indicated by a comma) occurs between *lighting* and *the candle*, whereas in sentence (2), the clause boundary occurs between *romantic* and *lighting*.

When spoken with normal prosody, the clause boundary in each of the sentences was accompanied by their ordinary pitch, stress, and timing prosodic boundary cues [indicated by a double slash (/)]. The sentences were recorded onto computer sound files (16-bit; 44.1 kHz) by a male speaker of American English, using a prosodic pattern appropriate to the particular clausal structure. These original recordings represented the *congruent prosody* condition.

Once these original sentence pairs were recorded, speech editing was then used to create two new sentences by swapping the highlighted word sequence from sentences (1) into the sentence frame of sentence (2) to produce sentence (4), and the highlighted sequence of sentence (2) into the sentence frame of sentence (1) to produce sentence (3). This yielded two new sentences that carried the full inventory of the prosodic features, which normally mark a syntactic clause boundary occurring at a different point in the sentence than the lexico-syntactically defined clause boundary. These cross-spliced sentences [sentences (3) and (4) in this example] represent the *conflicting prosody* condition.

Splicing was conducted using computer speech editing via Sound Studio 4 (Felt Tip Inc., New York, NY). This was accomplished by selecting the points for cross-splicing on a visual display of the sentence waveforms and verified by auditory monitoring such that, after splicing, the sentences with conflicting prosody sounded naturally fluent with the points of splicing unnoticed in the playback.

A third condition consisted of the same speaker re-recording the original sentences with *reduced prosody*. This

was characterized by a reduced clause-related pitch variation and differential word stress and an absence of pauses and clause-final word-lengthening, which are especially salient markers of a major clause boundary (Hoyte *et al.*, 2009; Nooteboom, 1997; Shattuck-Hufnagel and Turk, 1996). This latter feature reduced the average duration of the reduced prosody sentences ( $M = 3.96$  s), relative to the durations of the congruent prosody ( $M = 4.44$  s) and conflicting prosody ( $M = 4.43$  s) sentences.

The recordings of the completed stimuli were equalized for the root mean square (RMS) intensity using MATLAB (MathWorks, Natick, MA). To avoid the potential effects of splicing on intelligibility differences across the conditions, the congruent and reduced prosody sentences were spliced at the positions that were spliced in the conflicting prosody sentences.

**C. Procedures**

The participants were told that they would hear a series of recorded sentences. Each sentence would be followed by a 4-s silent period after which a tone would be presented. At this point, they were to recall the just-heard sentence as accurately and completely as possible. The participants were encouraged to say what they thought they might have heard even if they were unsure. The recall responses were given aloud and audio-recorded for later transcription and scoring. The instructions made no mention of the splicing procedure or prosody differences among the sentences.

When the participant indicated that they had recalled as much as they believed they could, they were to press a key that initiated the next trial. Each sentence was preceded by 4 s of silence to supply a pre-sentence pupil size baseline for that sentence. Each trial, thus, consisted of a 4-s silent period for the baseline acquisition, a stimulus sentence, a 4-s silent period, and the sentence recall.

The participants were tested individually in a sound attenuating booth with the stimuli transmitted through a sound field loudspeaker. The loudspeaker was positioned at zero azimuth, 1 m away from the listener. The stimuli were presented at a mean sound level of 65 dB (C scale), which was reported to be a comfortable listening level by all of the participants. The testing was conducted with the participant using his or her everyday implant program settings.

Each participant heard 36 sentences, 12 sentences with congruent prosody, 12 with conflicting prosody, and 12 with reduced prosody. The sentences were presented inter-mixed with regard to the prosody condition. No participant heard

TABLE II. Examples of the prosody conditions. Note that the comma indicates the lexico-syntactically defined clause boundary; the *double slash* (//) indicates the position of the prosodic marking of a clause boundary. The quotations indicate a shared seven-word sequence, cross-spliced.

Sentence type	Example sentences
(1) Congruent prosody	Because she wanted “romantic lighting, // the candle on the table” was important.
(2) Congruent prosody	Because she was a “romantic, // lighting the candle on the table” became a ritual.
(3) Conflicting prosody	Because she wanted “romantic // lighting, the candle on the table” was important.
(4) Conflicting prosody	Because she was a “romantic, lighting // the candle on the table” became a ritual.

the same core sentence more than once; the particular prosody condition with which the sentences were heard was varied across the participants.

The main experiment was preceded by three practice trials to familiarize the participant with the experimental procedures. The practice sentences consisted of sentences in each of the three prosody conditions. None of these sentences was used in the main experiment.

#### D. Pupillometry data acquisition

Throughout the course of the experiment, each participant's moment-to-moment pupil size was recorded via an EyeLink 1000 Plus eye-tracker (SR Research, Mississauga, Ontario, Canada), using a standard nine-point calibration procedure. The pupil size data were acquired at a rate of 1000 Hz and recorded using a program developed using the SR Research Experiment Builder, and processed via MATLAB software (MathWorks, Natick, MA). The participants placed their heads in a chin rest to reduce head movement and maintain a distance of 60 cm from the EyeLink camera. To further facilitate reliable pupil size measurements, the participants were instructed to keep their eyes on a centrally located fixation point continuously displayed on a computer screen, which was placed over the camera. The computer screen was filled with a medium gray color to avoid the ceiling or floor effects in the pupil size at the baseline (Winn *et al.*, 2018). The eye blinks were detected and removed using the algorithm described by Hershman *et al.* (2018) and filled by means of linear interpolation. The trials in which more than 85% of the data required interpolation were removed from the analyses.

The pupil sizes were baseline corrected to account for non-task changes in the base pupil size across the trials (Ayasse and Wingfield, 2020). This was accomplished by subtracting the mean pupil size measured over the last 1 s of the 4-s baseline silent period preceding each sentence from the task-related pupil size measures. (See Reilly *et al.*, 2019, for data and a discussion on linear versus proportional baseline scaling.)

An additional adjustment was made to account for the tendency for the older adults' base pupil size and dynamic range to be smaller than that for the younger adults (*senile miosis*; Bitsios *et al.*, 1996; Guillon *et al.*, 2016). Without an adjustment, one might underestimate the level of effort allocated to a task by the older adults relative to the younger adults. There is yet to be an agreed upon method for such an adjustment, whether using individual differences in the pupillary responses acquired in a reduced effort task as a reference point for the age differences in a cognitive task of interest (McLaughlin *et al.*, 2021) or light-range normalization based on the individual differences in the pupillary dynamic range (Piquado *et al.*, 2010). The two methods differ in the underlying assumptions and magnitude of the age differences in the apparent task-related effort, although both methods yield results in the same direction (McLaughlin *et al.*, 2021). In the present case, we employed light-range normalization, which involved representing the pupil sizes

as a percentage ratio of the individual's minimum pupil constriction and maximum pupil dilation.

The measures of the minimum pupil constriction and maximum pupil dilation were obtained by measuring the pupil size in response to the viewing light (199.8 cd/m<sup>2</sup>) and dark (0.4 cd/m<sup>2</sup>) screens presented for 60 s each prior to the main experiment. The percentage ratio was calculated as  $(d_M - d_{\min}) / (d_{\max} - d_{\min}) \times 100$ , where  $d_M$  was the participant's measured pupil size at a given time point;  $d_{\min}$  was the participant's minimum constriction, taken as the average pupil size over the last 30 s of viewing the light screen, and  $d_{\max}$  was the participant's maximum dilation measured as the pupil size averaged over the last 30 s of viewing the dark screen (e.g., Ayasse *et al.*, 2017; Ayasse and Wingfield, 2018; see Winn *et al.*, 2018, for a discussion of the procedures for adjusting for individual differences in the pupil size). The ambient light in the testing room was kept constant throughout the experiment.

### III. RESULTS

#### A. Recall accuracy

##### 1. Content words correct

The left panel of Fig. 1 shows the percentage of the content words reported correctly for the sentences heard with congruent prosody, reduced prosody, and conflicting prosody. The content words were defined as nouns, verbs, adjectives, and adverbs. The function words, such as determiners, prepositions, and conjunctions, were not scored.

The statistical tests conducted on the data in Fig. 1 and subsequent analyses were run using SPSS version 27.0. An alpha of 0.05 was used for all of the tests to determine whether the results were significantly different from chance.

A one-way repeated measures analysis of variance (ANOVA) confirmed the significant variance in the content words correctly recalled among the three prosody conditions [ $F(2,63) = 7.15$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.185$ ]. The pair-by-pair

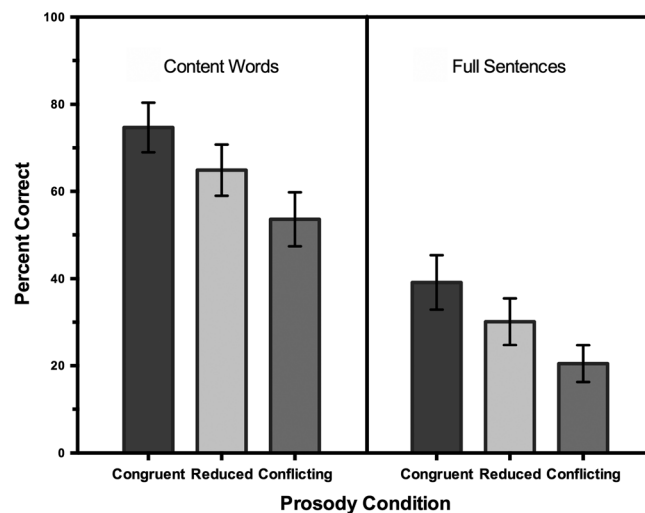


FIG. 1. The percentage of content words reported correctly (left) and the percentage of complete sentences correctly recalled (right) for sentences heard with congruent, reduced, and conflicting prosody. The error bars are one standard error.

comparisons confirmed that recall of the content words was superior when the sentence was supported by congruent prosody than for sentences with reduced prosody, [ $t(21) = 4.68, p < 0.001$ ], and words from sentences with reduced prosody were in turn better recalled than words from sentences with conflicting prosody [ $t(21) = 2.54, p = 0.019$ ].

**2. Complete sentences correct**

The right panel of Fig. 1 shows the data when the same CI users’ recall accuracy was assessed with the more demanding metric of the percentage of full sentences that were recalled accurately in each of the three prosody conditions. To be considered correct, the recall was required to be a complete sentence with all of the original words present, in the correct order, and with no extra words added. Minor changes, such as an article substitution (e.g., *the* for *a*), were accepted.

The pattern of prosody effects for this more stringent scoring was the same as for the more lenient scoring by content words. An ANOVA in this case also confirmed significant variance in the recall across the three prosody conditions [ $F(2,63) = 4.58, p = 0.014, \eta_p^2 = 0.127$ ]. The pair-by-pair comparisons confirmed that recall of the full sentences was superior when supported by congruent prosody than for sentences with reduced prosody [ $t(21) = 2.26, p = 0.035$ ] and the sentences with reduced prosody were better recalled than the sentences with conflicting prosody [ $t(21) = 2.25, p = 0.035$ ].

**B. Categories of recall errors**

It is well established in the memory literature that when meaningful sentences are incorrectly recalled, the erroneous recalls often retain coherence by the missing words being replaced in the recall by other words that plausibly fit the sentence context. The result in such cases is a sentence that departs from the original but nevertheless represents a grammatically correct and semantically reasonable response (e.g., Alba and Hasher, 1983; Little et al., 2006; Potter, 1993; Potter and Lombardi, 1990; Wingfield et al., 1995).

Of special interest in the present case were the recall errors, specifically in the conflicting prosody condition. Here, the converging evidence for the CI users’ ability to detect and to be influenced by the sentence prosody might be obtained by determining whether reconstructions in the imperfectly recalled sentences maintained a clause boundary at the lexico-syntactically defined position or whether the reconstructed “new” sentences had a clause boundary at the point suggested by the sentence prosody. This question was answered by examining a verbatim transcription of the erroneous responses written with no punctuation. The determination was then made as to whether the response was a complete sentence and, if so, where one would place the major clause boundary in the written text. This analysis showed the examples of reconstructions that maintained a major clause boundary at the lexico-syntactically defined

point as well as cases with a clause boundary shifted to follow the sentence prosody. Table III shows two representative examples of the latter case in which the participants’ erroneous recalls resolved the syntax-prosody conflict in favor of the sentence prosody.

For each example in Table III, the first entry shows the sentence with the shared word sequence from which the critical segment was spliced (*donor sentence*). Below that is the test sentence with a comma, indicating the lexico-syntactically defined clause boundary, and the double slash notation (*//*), indicating the prosodically marked boundary position. Below this is the participant’s recall response.

In the first example, the participant, in addition to having made a benign substitution of *that* for *the*, omitted the word *bets*. It can be seen that with this simple omission, the result was a fully grammatical sentence but one that had shifted the major clause boundary to the position marked by the sentence prosody. In the second example, a participant’s recall omitted the word *she*, shifting the syntactic clause boundary to follow *pictures*, as consistent with the prosodic marking. The boundary was then regularized by adding the word *her*. There were also minor changes in the form of the preposition *in* being changed to *with* and the word *painting* pluralized.

The relative influence of the prosody in the erroneous recalls of the conflicting prosody sentences is quantified in Fig. 2. The first two bars show the mean percentage of the incorrect responses, which formed a fully grammatical sentence with a major clause boundary placed at what had been the lexico-syntactically defined point in the original sentence or that had a major clause boundary that coincided with the prosodic marking. The next category represents the relatively small percentage of the cases in which the response was a grammatically coherent sentence but one that did not have a clause boundary at either point. These were largely recalls of an independent clause with the omission of its accompanying dependent clause. The remaining category of error responses consisted of sentence fragments that did not form a meaningful sentence.

TABLE III. Examples of donor sentences, test sentences, and prosody-influenced reproductions. Note that the comma indicates the lexico-syntactically defined clause boundary; the double slash (*//*) indicates the position of the prosodic marking of a clause boundary. The shared word sequences between the donor sentences and test sentences are underlined.

Donor sentence: Since the man at the casino was <u>winning</u> ,
<u>// bets</u> were placed on his success.
Test sentence: With too many people <u>winning</u> // <u>bets</u> ,
the casino had to fire many of their staff.
Response: With too many people winning,
(- - -) <b>that</b> casino had to fire many of their staff.
Donor sentence: Because of your <u>skill in painting</u> ,
<u>// the pictures</u> will be highly valued.
Test sentence: Because of her <u>skill in painting</u> // <u>the pictures</u> ,
she made a lot of money.
Response: Because of her skill with <b>paintings</b> ,
<b>her pictures</b> made a lot of money.

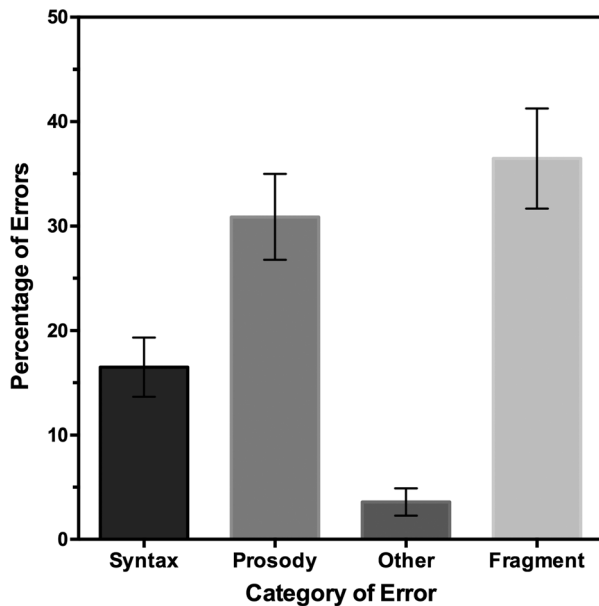


FIG. 2. The categories of incorrect recalls in the conflicting prosody condition. The first two bars show the percentage of responses that formed grammatically coherent sentences, which had a major clause boundary at the lexico-syntactically defined position in the original sentence (*syntax*), or at the point that had been suggested by the prosodic marking (*prosody*). The remaining two categories were grammatical responses that did not have a clause boundary at either point and sentence fragments that did not form a meaningful utterance. The error bars are one standard error.

Our major question in this analysis was the degree to which the prosodic marking of a clause boundary overrode the actual, lexico-syntactically defined, clause boundary. As seen in Fig. 2, when the participants produced grammatically coherent sentences, they showed a greater percentage of reconstructions that had a syntactic boundary at the prosodically marked position than at the lexico-syntactically defined position [ $t(21) = 3.23, p = 0.004$ ]. This result is consistent with the previously described effect of the prosody on recall accuracy in showing a significant influence of sentence prosody on the listeners' understanding of the clausal structure of what they were hearing.

### C. CNC score as a predictor

It is recognized in the CI literature that there are wide individual differences in speech recognition among CI recipients, which may be influenced by factors such as surgical placement within the cochlea (Aschendorff *et al.*, 2007; Finley *et al.*, 2008; Skinner *et al.*, 2007), the CI programming parameters (Firszt *et al.*, 2004; Zeng and Galvin, 1999), and others.

We conducted a series of hierarchical multiple regressions to examine the effects of the individuals' word recognition ability (assessed by CNC-30 scores), vocabulary knowledge (assessed by Shipley scores), and participant age at time of testing on the recall accuracy. This was performed for the two recall accuracy scoring methods (content words correct, full sentences correct) for each of the three prosody conditions. The predictor variables were entered into the

model in the following order: CNC score, Shipley vocabulary score, and participant age.

The results of the regression analyses are shown in Table IV. For each predictor variable in each condition, we show  $R^2$ , which represents the cumulative contribution of each variable along with the previously entered variables, and the change in  $R^2$ , which shows the contribution of each variable at each step. The next column shows the level of significance of each variable. It can be seen that the CNC score was a significant predictor of the percentage of content words recalled correctly for all three of the prosody conditions and percentage of full sentences recalled correctly in the congruent prosody condition. It can also be seen that neither the vocabulary score nor participants' age contributed significant variance to the recall accuracy once the CNC scores were taken into account.

### D. Pupillometry

Figure 3(A) shows the mean adjusted pupil sizes as the participants listened to sentences with congruent, reduced, and conflicting prosody. Also shown are the mean adjusted pupil sizes for the first 2 s of the 4-s silent period between the end of a sentence and the signal to begin the sentence recall. The vertical line at the zero-point on the abscissa is aligned with the ends of the sentences. The pupil sizes are shown relative to the pre-sentence baselines and scaled as a

TABLE IV. The summary of hierarchical regressions. Note that the  $p$ -values reflect the significance of the changes in  $R^2$  at each step of the model. The significant values are shown in bold.

Content words correct				
Prosody	Predictor	$R^2$	Change in $R^2$	$p$ value
Congruent	CNC score	0.433	0.433	<b>0.002</b>
	Vocabulary	0.473	0.041	0.268
	Age	0.49	0.017	0.476
Reduced	CNC score	0.453	0.453	<b>0.001</b>
	Vocabulary	0.486	0.033	0.314
	Age	0.546	0.06	0.166
Conflicting	CNC score	0.199	0.199	<b>0.049</b>
	Vocabulary	0.241	0.042	0.345
	Age	0.336	0.094	0.151
Complete sentences correct				
Condition	Predictor	$R^2$	Change in $R^2$	$p$ value
Congruent	CNC score	0.224	0.224	<b>0.035</b>
	Vocabulary	0.225	0.001	0.893
	Age	0.389	0.164	0.055
Reduced	CNC score	0.158	0.158	0.082
	Vocabulary	0.184	0.026	0.472
	Age	0.216	0.032	0.431
Conflicting	CNC score	0.032	0.032	0.449
	Vocabulary	0.117	0.085	0.217
	Age	0.184	0.067	0.27

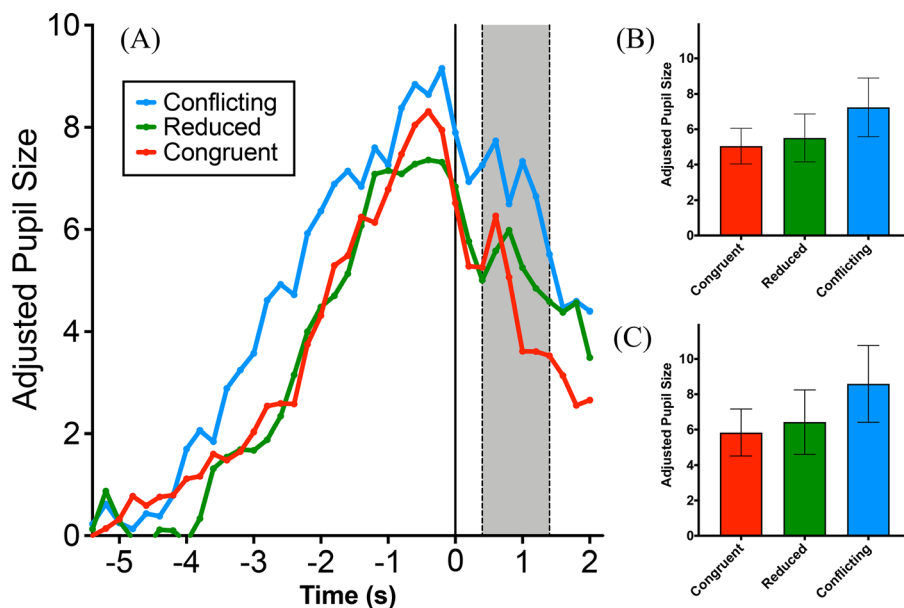


FIG. 3. (Color online) (A) shows the time course of adjusted pupil dilations while listening to sentences with conflicting, reduced, and congruent prosody, and for the first 2 s of the 4-s silent interval between the end of a sentence and the signal to begin the sentence recall. Pupil sizes are shown relative to the pre-sentence baselines, further scaled as a percentage of the individuals' pupillary dynamic range (see the text). The vertical line at the zero-point on the abscissa is aligned with the sentence endings. (B) shows the mean adjusted pupil sizes over a 1-s time window beginning 400 ms after a sentence ending for each of the three prosody conditions [shaded area in (A)]. (C) shows these data, excluding the participants with CNC scores below 45% correct. The error bars are one standard error.

percentage of the individuals' calculated dynamic range as previously described. These data are shown for 18 of the 22 participants. The pupillometry data were unavailable for one participant because of technical issues, and three participants were eliminated due to excessive blinking on all 12 trials in at least 1 prosody condition.

There are reports in the literature of an association between the rate of the participants' eye blinks and task difficulty, although with mixed findings as to the direction of this association (cf. Burg *et al.*, 2021; Holland and Tarlow, 1972; Rosenfield *et al.*, 2015; Wood and Hassett, 1983). In the present case, for these 18 participants included in Fig. 3, the percentage of samples lost due to blinking and filled by interpolation was similar across the three prosody conditions (13.0%, 12.8%, and 13.0% for the congruent, reduced, and conflicting prosody conditions, respectively). The present data did not reveal a direct association between the recall accuracy and pupil dilation for any of the three prosody conditions, either in terms of the content words or full sentences correct.

It can be seen in the time series data in Fig. 3(A) that the mean adjusted pupil sizes progressively increase as more and more of a sentence is being heard. It is especially notable that the differential effects of the prosody condition appear even after a sentence has ended. This latter point is quantified in Fig. 3(B), which shows the mean adjusted pupil sizes for the 18 participants, averaged over a 1-s time window, beginning at 400 ms after the sentences had ended. The starting point for this time window was selected by visual inspection as the point at which all three curves representing the three prosody conditions begins to diverge. The choice of a 1-s region of interest was to some extent arbitrary, but its selection captures the post-sentence time window, which includes the maximal separation of the three prosody condition curves. This time window is indicated by the shaded region shown in Fig. 3(A).

The data in Fig. 3(B) were submitted to a one-way repeated measures ANOVA, which confirmed a significant effect of the prosody condition on the mean pupil size [ $F(2,34) = 4.80$ ,  $p = 0.015$ ,  $\eta_p^2 = 0.22$ ]. The paired-comparisons showed that hearing sentences with conflicting prosody produced a larger mean pupil dilation in the post-sentence time window than for sentences with reduced prosody [ $t(17) = 2.53$ ,  $p = 0.022$ ] or congruent prosody [ $t(17) = 2.54$ ,  $p = 0.021$ ]. The comparison between the pupil dilations for the congruent prosody versus reduced prosody sentences was not significant [ $t(17) = 0.69$ ,  $p = 0.503$ ].

Although hearing sentences in which there was a conflict between the lexico-syntactically defined and prosodically marked clause boundary positions resulted in a significant increase in pupil dilation relative to the sentences with reduced and congruent prosody, the large error bars around the means in Fig. 3(B) are indicative of the wide individual differences in the pupil dilation within each of the prosody conditions. The finding that the CNC scores played a significant role in the content word recall, however, raised the possibility that individual differences in word recognition ability as measured by the CNC scores and attendant processing effort may have contributed to this variability in the pupil dilation. To explore this possibility within each of the three prosody conditions, we asked whether the pupil size varied with the individual participants' word recognition ability.

Figure 4 shows for each prosody condition, the mean adjusted pupil size for each of the 18 participants for whom the pupillometry data were available as a function of their CNC scores. On visual inspection, it appears that for those participants with CNC scores of 45% or better (shaded portions in Fig. 4), the higher the CNC score, the smaller the pupil dilation, and this relationship appears in all three of the prosody conditions. Also shown in each panel is the formula for the plotted line of best fit along with the Pearson



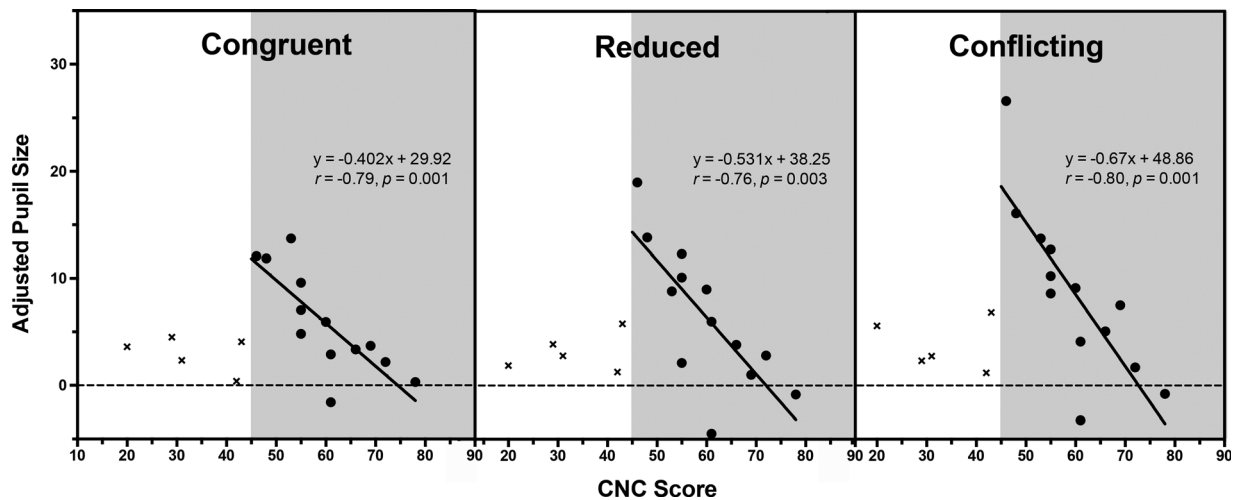


FIG. 4. The mean adjusted pupil size in the period prior to the recall signal in each of the three prosody conditions plotted as a function of the individual participants' CNC-30 score. The formulas for the best fit linear regressions and Pearson correlations are shown in each case for the participants with CNC-30 scores greater than 45% (shaded areas).

correlation, which can be seen to be significant in all three of the cases. It can also be seen, however, that for the five participants with CNC scores below 45%, there was no clear pattern relating the pupil size and CNC score in any of the three prosody conditions.

The mean adjusted pupil sizes in the same post-sentence time window as used in Fig. 3(B) are shown in Fig. 3(C) just for the participants with a CNC score above 45%. It can be seen that the pattern of the effects of prosody on the pupil size is similar to that in Fig. 3(B), where all of the participants were included. The data shown in Fig. 3(C) were submitted to a one-way repeated measures ANOVA, which confirmed the significance variance in the pupil sizes across the three prosody conditions [ $F(2,24) = 4.34, p = 0.025, n_p^2 = 0.266$ ]. The paired comparisons confirmed the larger average pupil sizes for the sentences heard with conflicting prosody than for those with reduced prosody [ $t(12) = 2.47, p = 0.003$ ] and larger pupil sizes for the sentences heard with conflicting prosody than for those in the congruent prosody [ $t(12) = 2.43, p = 0.032$ ]. A marginal difference appeared between the pupillary responses associated with the congruent versus reduced prosody conditions [ $t(12) = 0.064, p = 0.534$ ]. Neither the age nor vocabulary differences among the participants accounted for the variability in the adjusted pupil sizes.

The number of test stimuli heard by each participant (12 trials per condition) was predicated in part by the demanding nature of the decision task and in part by a desire to mitigate the likelihood of the participants developing artificial solution strategies specific to this task. This is, however, a small number of trials per condition for a pupillometry study, raising a concern that a few aberrant responses might have undue influence on the aggregate data. An examination of the participant-by-participant and trial-by-trial data suggests that this was not the case in the present experiment. At the same time, we align ourselves with the importance of having as large a sample size as is practicably possible to ensure reliability in the pupillometry studies (see

Winn *et al.*, 2018, pp. 7–8, for a more complete discussion of the desired number of trials and task demands).

#### IV. DISCUSSION

##### A. Effect of prosody on sentence recall

Numerous studies have shown that CI users often have difficulty in distinguishing linguistic contrasts that rely heavily on the intonation pattern (i.e., pitch contour) of an utterance, such as determining whether a word or sentence is spoken as a question or a statement (e.g., Holt and McDermott, 2013; Marx *et al.*, 2015; Peng *et al.*, 2009; Townshend *et al.*, 1987). Although the intonation pattern of an utterance is an important element in the full complex of the linguistic prosody, our present data imply that CI users are able to use other, potentially spared features of the sentence prosody to facilitate detection of the clausal structure of a multi-clause sentence. This was evidenced in two behavioral measures. The first was a significant recall advantage for sentences when their prosodic patterns and lexico-syntactically defined clause boundaries coincided and a significant disadvantage when the prosodic and lexico-syntactic structures were in conflict. This was so whether it was measured by the content words recalled or complete sentences recalled correctly.

It has been shown that the nature of the errors that individuals make in the presence of a degraded signal can offer valuable insight into the elements underlying the participants' perceptual operations, such as identifying specific phonological errors leading to misperceived words and their downstream consequences on the coherence of a perceptual report (Winn and Teece, 2021). In the present case, our focus was on the extent to which the nature of the participants' recall errors might offer converging evidence on the CI users' ability to detect and, hence, be influenced by the prosodic pattern of a sentence in determining its clausal structure.

Although elements of the speech signal may be processed online as it is arriving (Marslen-Wilson and Tyler, 1980), the interval between the end of a sentence and the instruction to begin the recall is not a period of passive retention. Rather, the literature on sentence memory has long shown that considerable processing continues after a sentence is presented with the degree, and presumed duration of active processing proportional to the syntactic complexity of the utterance (e.g., Savin and Perchonock, 1965).

This latter point was supported in the present study by the pupillometry data, which showed a differential effect of the prosody condition after the stimulus sentence itself had ended. In part, this post-sentence activity likely included, at different time scales, continued phonological processing, interactive matching of the input within the mental lexicon, and, in some cases, context-based repair of the initial misperceptions (cf. Tyler *et al.*, 2000; Winn and Moore, 2018; Winn and Teece, 2021). This post-sentence activity would presumably be a factor in all three of the prosody conditions. We suggest that in the present case, the differences between the post-sentence pupillary responses were reflective of the differences in the processing effort attendant to syntactic resolution of the clausal structure of the sentences when accompanied by supportive versus conflicting prosody. In this regard, the resolution of the clausal structure of a multi-clause sentence has special importance as the syntactic clauses represent major processing points in developing the semantic coherence of the utterance (Fallon *et al.*, 2004; Fodor *et al.*, 1974; Jarvella, 1971, 1979).

As was seen in the present data, in addition to a significant recall advantage for sentences spoken with congruent prosody, the influence of the sentence prosody also appeared in an examination of the nature of the errors in the subset of the incorrectly recalled conflicting-cue sentences in which the contextually based reconstructions resulted in syntactically coherent, meaningful sentences. The drive to give coherence in the recall of meaningful sentences is not a novel finding (e.g., Alba and Hasher, 1983; Little *et al.*, 2006; Potter, 1993; Potter and Lombardi, 1990; Wingfield *et al.*, 1992; Wingfield *et al.*, 1995). Indeed, such instances reflect the general principle that the memory for narratives and events is as much reconstructive as it is reproductive (Bartlett, 1932; Oldfield, 1954). Rather, the novel finding is that for CI users, many of the imperfectly recalled sentences in the conflicting-cue condition showed reconstructions that yielded a “new” sentence with a major clause boundary at the position suggested by the prosodic marking. Indeed, these instances occurred relatively more often than the reconstructions that honored the lexico-syntactically defined boundary position. Together with the influence of prosody on the recall accuracy, these data are inconsistent with a general prosody deficit in CI users, which extends to the use of the prosody as an aid to syntactic parsing.

## B. Predictors of recall accuracy

As is often found among CI users, the participants serving in the present study varied widely in word recognition

ability as measured by the CNC-30 scores (cf. Gifford *et al.*, 2008; Holden *et al.*, 2013). Determining the source of these individual differences exceeds the scope of this present study, although it is understood that factors, such as the age of onset of the loss and duration of the loss prior to receiving implants, with implications for the possible degeneration of spiral ganglion neurons (cf. Blamey *et al.*, 2013; Boisvert *et al.*, 2011; Cheng and Svirsky, 2021; Cohen and Svirsky, 2019; Leake *et al.*, 1999), surgical placement within the cochlea (Aschendorff *et al.*, 2007; Finley *et al.*, 2008; Skinner *et al.*, 2007; Yukawa *et al.*, 2004), and CI programming parameters (Firszt *et al.*, 2004; Zeng and Galvin, 1999) may all contribute to the variation in word recognition success.

It is not surprising that accurate recall of the content words from sentences in all three of the prosody conditions was predicted by the individuals’ CNC-30 scores. It is interesting that in spite of the range in age and vocabulary scores among our participants, the regression analysis failed to show that either factor predicted the recall accuracy once the CNC-30 scores were taken into account. It is also notable that the CNC-30 word recognition was a much weaker predictor of the recall accuracy for full sentences than for recall of content words from those same sentences, particularly in the reduced and conflicting prosody conditions. This finding would imply that in the absence of support from the congruent prosody, as was the case in the reduced and conflicting prosody conditions, resolving the syntactic structure drew on the linguistic operations, whose effects on the accuracy overrode or masked the effects due to the differences in word recognition ability.

## C. Pupil dilation and processing effort

Studies have shown that even when word recognition is successful, lexical activation with reduced spectral information can be effortful (Winn *et al.*, 2015). The link between the processing effort (which was indexed in the present study by pupil dilation) and recall accuracy followed the presumed order of the processing challenge with the mean pupil dilation during a post-sentence time window increasing from the sentences with congruent prosody to those with conflicting prosody. In spite of this general trend, the pupillary data were notable in the wide variability around these means.

The variability in the processing effort was, as that for the recall accuracy, also related to the individuals’ CNC-30 word recognition scores. For those participants with a CNC score better than 45%, there was a clear inverse relationship between the word recognition ability as measured by the CNC-30 score and the processing effort as indexed by the pupil dilation. This systematic relationship disappeared, however, for the participants with CNC-30 scores below 45% correct. It would seem likely that when the word recognition ability is relatively poor, factors other than a simple recognition difficulty-expenditure of effort relationship may come into play.

This pattern of the word recognition/pupil dilation relationship breaking down when the former drops below 45% is consistent with reports by [Wendt \*et al.\* \(2018\)](#) and [Ohlenforst \*et al.\* \(2017\)](#). Testing the word recognition as the signal-to-noise ratio was decreased, these authors found that the pupil dilation would peak when the accuracy level was about 40%–590%. When the signal-to-noise ratio resulted in an accuracy level below 40%, a drop in the pupil sizes was observed. As in the present case, it is reasonable to suggest that when the word recognition difficulty exceeds this level of difficulty, the listeners essentially give up with a concomitant withdrawal of effort ([Ohlenforst \*et al.\*, 2017](#); [Wendt \*et al.\*, 2018](#); see also [Ayasse and Wingfield, 2018](#); [Kuchinsky \*et al.\*, 2014](#); [Zekveld and Kramer, 2014](#)).

As characterized in *Framework for Understanding Effortful Listening* (FUEL) by [Pichora-Fuller \*et al.\* \(2016\)](#), exerting effort and investing mental resources are seen as intimately related if not synonymous. It is further recognized that how much effort is allocated to a task while depending on the task demands may also be moderated by motivational factors. That is, individuals will engage effort and, hence, an allocation of processing resources, only to the extent that they believe that additional effort will bring success ([Richter, 2016](#)). In terms of behavioral economics, there may be an interdependence between the commitment of the effort and likely payoff, and that effort may not be expended if the individual believes there is an unlikely return on their investment ([Eckert \*et al.\*, 2016](#)). We may speculate that in the present experiment, those CI users with poorer word recognition ability may have varied among themselves in their willingness to invest the effort with an uncertain belief in likely success. This may be contrasted with those with better word recognition scores, where effort, as indexed by the pupil dilation, was inversely related to their ease of word recognition in a systematic way. This apparent dissociation, we suggest, warrants future investigation, especially as it may open a window to motivational factors in the outcome success.

#### D. Why is prosody spared?

It is the case that clause boundaries can be identified in the absence of prosodic marking ([Garrett \*et al.\*, 1966](#)). When available, however, prosody can be a significant aid to clause boundary detection as a step toward comprehension of a sentence's meaning, a finding not limited to English (cf. [Butt \*et al.\*, 1980](#); [Kang and Speer, 2005](#); [Nootheboom, 1997](#)). Nevertheless, when present and especially when the listener is faced with a relatively complex multi-clause sentence, it is clear that the presence of linguistically tied sentence prosody can aid in the determination of the clausal structure ([Kjelgaard and Speer, 1999](#); [Kjelgaard \*et al.\*, 1999](#); [Nootheboom, 1997](#); [Shattuck-Hufnagel and Turk, 1996](#); [Titone \*et al.\*, 2006](#)). The question remains as to why CI users can show a significant influence of the prosodic pattern of a sentence in spite of the loss of the spectral

richness that appears in, for example, the reduced ability to detect the important prosodic feature of the pitch contour.

We attribute the maintained influence of the prosody by the CI users to the presence of multiple acoustic features, which typically co-occur at major clause boundaries that include inter- and intra-word timing cues and lexical stress as well as pitch contour. In the information theoretic sense ([Attneave, 1959](#); [Shannon and Weaver, 1949](#); [van Rooij and Plomp, 1991](#)), the prosody can be said to add *redundancy* to the speech signal, reducing the potential uncertainty as to the clausal structure of a sentence at or near a clause boundary. This redundancy offers the CI listener an array of prosodic features rather than necessarily requiring reliance on a single feature, which may be poorly conveyed by the CI signal. That is, to the extent that the limited spectral information available from a CI may limit the successful use of the pitch contour (cf. [Chatterjee and Peng, 2008](#); [Holt and McDermott, 2013](#); [Marx \*et al.\*, 2015](#); [Peng \*et al.\*, 2012](#)), the availability of other features, such as the lexical stress, which is marked among other features by an increase in the relative amplitude and pauses and word-lengthening at a clause boundary, may be more than adequate.

It should be noted that although the above features may co-occur at or near a clause boundary, not all of these features may be present from utterance to utterance, even by the same speaker. For example, on some occasions, there may be no pause at a clause boundary or a relative lengthening of the clause-final word. Instead, the location of the clause may be suggested by the pattern of the lexical stress in the vicinity of the clause boundary. It is possible to target a single feature, such as artificially introducing silent periods of various durations or varying clause-final word durations, to determine the minimum durations necessary to elicit the perception of a clause boundary (e.g., [Engelhardt \*et al.\*, 2010](#); [Lehiste \*et al.\*, 1976](#); [Nootheboom, 1997](#)). However, although informative, this will only give a limited picture of what is normally a complex interaction of the pitch, amplitude, and timing. That is, as has been argued elsewhere, the syntax-prosody interaction operates at an utterance-level prosodic representation rather than on specific local features such as the absolute pause duration, stress-indicating amplitude, or a specific pitch height ([Carlson \*et al.\*, 2001](#); [Kjelgaard and Speer, 1999](#); [Kjelgaard \*et al.\*, 1999](#)).

We cannot say from the present data which prosodic features, alone or in combination, contributed to the CI users' clause boundary detection. For normal hearing adults, timing variations in the form of pauses and clause-final word-lengthening are weighted especially heavily in detecting a syntactic clause boundary ([Hoyte \*et al.\*, 2009](#); [Nootheboom, 1997](#); [Sharpe \*et al.\*, 2017](#); [Shattuck-Hufnagel and Turk, 1996](#)), and this may also be the case for CI users. This speculation is supported by the fact that many CI users have a relatively normal ability to process temporal cues (e.g., [Sagi \*et al.\*, 2009](#); [Shannon, 1993](#)).

CI users, like all listeners, expect what they hear in everyday discourse to have semantic coherence as they

strive for the meaning of what is being heard. This can be seen in the extra effort that listeners spontaneously commit to in determining the meaning of an ambiguous homophone that makes the most sense within a particular sentence context (Kadem *et al.*, 2020). This assumption of rationality can also be seen in the way that listeners with normal hearing (Lash *et al.*, 2013), mild-to-moderate hearing loss (Benichov *et al.*, 2012), and CI users (Amichetti *et al.*, 2018) readily employ the linguistic context to infer the identity of indistinct or otherwise degraded words, a manifestation of the implicit belief that the syntactic and semantic content of the utterance will follow reasonable expectations. Although it is ordinarily adaptive, this expectation of coherence can result in cases of *false hearing*, which is when an especially strong context can cause a listener to “hear” a highly predictable word instead of a phonologically similar word that was actually presented (Rogers *et al.*, 2012; Rogers and Wingfield, 2015).

Winn and Teece (2021) have shown that the expectation of semantic coherence reliably appears in CI users as is evidenced by an increased pupillary response, which is indicative of the engagement of extra effort when the stimuli were encountered that violated this expectation. In the present case, the CI users’ pupillary responses revealed a significant increase in the processing effort when the syntactic coherence of a sentence was disturbed by a prosodic pattern in conflict with the position of a major clause boundary relative to when the prosodic pattern supported the correct boundary detection.

Taken together, the present results demonstrate that CI users are able to make use of the sentence prosody to facilitate the detection of the clausal structure of a multi-clause sentence, an important skill when parsing real-world speech input, and which may not be adequately assessed by tests that measure the identification of single words or individual simple sentences.

## ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIH) Grant No. R01 DC016834 from the National Institute of Deafness and Other Communication Disorders. We also gratefully acknowledge support from the Stephen J. Cloobek Research Fund.

Alba, J. W., and Hasher, L. (1983). “Is memory schematic?,” *Psychol. Bull.* **93**, 203–231.

Amichetti, N. M., Atagi, E., Kong, Y.-Y., and Wingfield, A. (2018). “Linguistic context versus semantic competition in word recognition by younger and older adults with cochlear implants,” *Ear. Hear.* **39**, 101–109.

Aschendorff, A., Kromeier, J., Klenzner, T., and Laszig, R. (2007). “Quality control after insertion of the nucleus contour and contour advance electrode in adult,” *Ear. Hear.* **28**(Suppl 2), 75S–79S.

Attneave, F. (1959). *Applications of Information Theory to Psychology* (Holt, Rinehart, and Winston, New York, NY).

Ayasse, N. D., Lash, A., and Wingfield, A. (2017). “Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment,” *Front. Aging Neurosci.* **8**, 329.

Ayasse, N. D., and Wingfield, A. (2018). “A tipping point in listening effort: Effects of linguistic complexity and age-related hearing loss on sentence comprehension,” *Trends Hear.* **22**, 2331216518790907.

Ayasse, N. D., and Wingfield, A. (2020). “Anticipatory baseline pupil diameter is sensitive to differences in hearing thresholds,” *Front. Psychol.* **10**, 2947.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology* (Cambridge University Press, Cambridge, UK).

Beckman, M., and Edwards, J. (1987). “On lengthenings and shortenings and the nature of prosodic constituency,” in *Papers in Laboratory Phonology*, edited by M. Beckman and J. Kingston (Cambridge University Press, Cambridge, UK), pp. 117–143.

Benichov, J., Cox, L. C., Tun, P. A., and Wingfield, A. (2012). “Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function,” *Ear. Hear.* **33**, 250–256.

Bitsios, P., Prettyman, R., and Szabadi, E. (1996). “Changes in autonomic function with age, a study of pupillary kinetics in healthy young and old people,” *Age Ageing* **25**, 432–438.

Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R., Fraysse, B., Gallégo, S., Govaerts, P. J., Green, K., Huber, A. M., Kleine-Punte, A., Maat, B., Marx, M., Mawman, D., Mosnier, I., O’Connor, A. F., O’Leary, S., Rousset, A., Schauwers, K., Skarzynski, H., Skarzynski, P. H., Sterkers, O., Terranti, A., Truy, E., Van de Heyning, P., Venail, F., Vincent, C., and Lazard, D. S. (2013). “Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients,” *Audiol. Neurootol.* **18**, 36–47.

Boisvert, I., McMahon, C. M., Tremblay, G., and Lyxell, B. (2011). “Relative importance of monaural sound deprivation and bilateral significant hearing loss in predicting cochlear implantation outcomes,” *Ear. Hear.* **32**, 758–766.

Breitenstein, C., Van Lancker, D., and Daum, I. (2010). “The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample,” *Cognit. Emot.* **15**, 57–79.

Budenz, C. L., Cosetti, M. K., Coelho, D. H., Birenbaum, B., Babb, J., Waltzman, S. B., and Roehm, P. C. (2011). “The effects of cochlear implants on speech perception in older adults,” *J. Am. Geriatr. Soc.* **59**, 446–453.

Burg, E. A., Thakkar, T., Fields, T., Misurelli, S. M., Kuchinsky, S. E., Roche, J., Lee, D. J., and Litovsky, R. Y. (2021). “Systematic comparison of trial exclusion criteria for pupillometry data analysis in individuals with single-sided deafness and normal hearing,” *Trends Hear.* **25**, 23312165211013256.

Butt, J., Wingfield, A., and Sandoval, A. W. (1980). “Effets de la prosodie sur la résolution syntaxique de la parole comprimée” (“Effects of prosody on the syntactic resolution of compressed speech”), *Annee Psychol.* **80**, 33–50.

Carlson, K., Clifton, C., and Frazier, L. (2001). “Prosodic boundaries in adjunct attachment,” *J. Mem. Lang.* **45**, 58–81.

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

Cheng, Y.-S., and Svirsky, M. A. (2021). “Meta-analysis—Correlation between spiral ganglion cell counts and speech perception with a cochlear implant,” *Audiol. Res.* **11**, 220–226.

Cohen, S. M., and Svirsky, M. A. (2019). “Duration of unilateral auditory deprivation is associated with reduced speech perception after cochlear implantation: A single-sided deafness study,” *Cochlear Implants Int.* **20**, 51–56.

Cosentino, S., Carlyon, R. P., Deeks, J. M., Parkinson, W., and Bierer, J. A. (2016). “Rate discrimination, gap detection and ranking of temporal pitch in cochlear implant users,” *J. Assoc. Res. Otolaryngol.* **17**, 371–382.

Eckert, M. A., Teuber-Rhodes, S., and Vaden, K. I. (2016). “Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions,” *Ear. Hear.* **37**(Suppl 1), 101S–110S.

Engelhardt, P. E., Ferreira, F., and Patsenko, E. G. (2010). “Pupillometry reveals processing load during spoken language comprehension,” *Q. J. Exp. Psychol.* **63**, 639–645.

Everhardt, M. K., Sarampalis, A., Coler, M., Baskent, D., and 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.

- Fallon, M., Kuchinsky, S., and Wingfield, A. (2004). "The salience of linguistic clauses in young and older adults' running memory for speech," *Exp. Aging Res.* **30**, 359–371.
- Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., Hullar, T. E., and Skinner, M. W. (2008). "Role of electrode placement as a contributor to variability in cochlear implant outcomes," *Otol. Neurotol.* **29**, 920–928.
- Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., Runge-Samuels, C. L., and 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.
- Fodor, J. A., Bever, T. G., and Garrett, M. F. (1974). *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar* (McGraw-Hill, New York, NY).
- Friedland, D. R., Runge-Samuels, C., Baig, H., and Jensen, J. (2010). "Case-control analysis of cochlear implant performance in elderly patients," *Arch. Otolaryngol. Head Neck Surg.* **136**, 432–438.
- 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., and Nogaki, G. (2005). "Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing," *J. Assoc. Res. Otolaryngol.* **6**, 19–27.
- Garrett, M., Bever, T., and Fodor, J. (1966). "The active use of grammar in speech perception," *Percept. Psychophys.* **1**, 30–32.
- Gifford, R. H., Shallop, J. K., and Peterson, A. M. (2008). "Speech recognition materials and ceiling effects: Considerations for cochlear implant programs," *Audiol. Neurootol.* **13**, 193–205.
- Guillon, M., Dumbleton, K., Theodoratos, P., Gobbe, M., Wooley, C. B., and Moody, K. (2016). "The effects of age, refractive status, and luminance on pupil size," *Optom. Vis. Sci.* **93**, 1093–1100.
- Hershman, R., Henik, A., and Cohen, N. (2018). "A novel blink detection method based on pupillometry noise," *Behav. Res. Methods* **50**, 107–114.
- 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., and Skinner, M. W. (2013). "Factors affecting open-set word recognition in adults with cochlear implants," *Ear Hear.* **34**, 342–360.
- Holland, M. K., and Tarlow, G. (1972). "Blinking and mental load," *Psychol. Rep.* **31**, 119–127.
- Holt, C. M., and McDermott, H. J. (2013). "Discrimination of intonation contours by adolescents with cochlear implants," *Int. J. Audiol.* **52**, 808–815.
- Hood, L. J., Svirsky, M. A., and Cullen, J. K., Jr. (1987). "Discrimination of complex speech-related signals with a multichannel electronic cochlear implant as measured by adaptive procedures," *Ann. Otol. Rhinol. Laryngol.* **96**, 38–40.
- Hoyte, K. J., Brownell, H., and Wingfield, A. (2009). "Components of speech prosody and their use in detection of syntactic structure by older adults," *Exp. Aging Res.* **35**, 129–151.
- Jackendoff, R. J. (1971). *Semantic Interpretation in Generative Grammar* (MIT Press, Cambridge, MA).
- Jarvella, R. J. (1971). "Syntactic processing of connected speech," *J. Verb. Learn. Verb. Behav.* **10**, 409–416.
- Jarvella, R. J. (1979). "Immediate memory and discourse processing," in *The Psychology of Learning and Motivation* (Academic, New York, NY), Vol. 13.
- Just, M. A., and Carpenter, P. A. (1993). "The intensity dimension of thought: Pupillometric indices of sentence processing," *Can. J. Exp. Psychol.* **47**, 310–339.
- Kadem, M., Herrmann, B., Rodd, J. M., and Johnsrude, I. S. (2020). "Pupil dilation is sensitive to semantic ambiguity and acoustic degradation," *Trends Hear.* **24**, 2331216520964068.
- Kang, S., and Speer, S. R. (2005). "Effects of prosodic boundaries on syntactic disambiguation," *Stud. Linguist.* **59**, 244–258.
- Kim, M., Beversdorf, D. Q., and Heilman, K. M. (2000). "Arousal response with aging: Pupillographic study," *J. Int. Neuropsychol. Soc.* **6**, 348–350.
- Kinner, V. L., Kuchinke, L., Dierolf, A. M., Merz, C. J., Otto, T., and Wolf, O. T. (2017). "What our eyes tell us about feelings: Tracking pupillary responses during emotion regulation processes," *Psychophysiology* **54**, 508–518.
- Kjelgaard, M. M., and Speer, S. R. (1999). "Prosodic facilitation and interference in the resolution of temporary syntactic ambiguity," *J. Mem. Lang.* **40**, 153–194.
- Kjelgaard, M. M., Titone, D. A., and Wingfield, A. (1999). "The influence of prosodic structure on the interpretation of temporary syntactic ambiguity by young and elderly listeners," *Exp. Aging Res.* **25**, 187–207.
- Klatt, D. H., and Klatt, L. C. (1990). "Analysis, synthesis, and perception of voice quality variations among male and female talkers," *J. Acoust. Soc. Am.* **87**, 820–857.
- Koelwijn, T., Zekveld, A. A., Festen, J. M., and Kramer, S. E. (2012). "Pupil dilation uncovers extra listening effort in the presence of a single-talker masker," *Ear Hear.* **33**, 291–300.
- Kuchinsky, S. E., Ahlstrom, J. B., Cute, S. L., Humes, L. E., Dubno, J. R., and Eckert, M. A. (2014). "Speech-perception training for older adults with hearing loss impacts word recognition and effort," *Psychophysiology* **51**, 1046–1057.
- Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cute, S. L., Humes, L. E., Dubno, J. R., and Eckert, M. A. (2013). "Pupil size varies with word listening and response selection difficulty in older adults with hearing loss," *Psychophysiology* **50**, 23–34.
- Landsberger, D. M., Syrakic, M., Roland, J. T., Jr., and Svirsky, M. (2015). "The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants," *Ear Hear.* **36**, e207–e213.
- Lash, A., Rogers, C. S., Zoller, A., and Wingfield, A. (2013). "Expectation and entropy in spoken word recognition: Effects of age and hearing acuity," *Exp. Aging Res.* **39**, 235–253.
- Leake, P. A., Hradek, G. T., and Snyder, R. L. (1999). "Chronic electrical stimulation by a cochlear implant promotes survival of spiral ganglion neurons after neonatal deafness," *J. Comp. Neurol.* **412**, 543–562.
- Lehiste, I. (1970). *Suprasegmentals* (MIT Press, Cambridge, MA).
- Lehiste, I., Olive, J. P., and Streeter, L. A. (1976). "Role of duration in disambiguating syntactically ambiguous sentences," *J. Acoust. Soc. Am.* **60**, 1199–1202.
- Lehiste, I., and Peterson, G. E. (1959). "Linguistic considerations in the study of speech intelligibility," *J. Acoust. Soc. Am.* **31**, 280–286.
- Little, D. M., McGrath, L. M., Prentice, K. J., and Wingfield, A. (2006). "Semantic encoding of spoken sentences: Adult aging and the preservation of conceptual short-term memory," *Appl. Psycholinguist.* **27**, 487–511.
- Luxford, W. (2001). "Minimum speech test battery for postlingually deafened adult cochlear implant patient," *Otolaryngol. Head Neck Surg.* **124**, 125–126.
- Marslen-Wilson, W., and Tyler, L. K. (1980). "The temporal structure of spoken language understanding," *Cognition* **8**, 1–71.
- Marx, M., James, C., Foxton, J., Capber, A., Fraysse, B., Barone, P., and Deguine, O. (2015). "Speech prosody perception in cochlear implant users with and without residual hearing," *Ear Hear.* **36**, 239–248.
- McLaughlin, D. J., Zink, M. E., Gaunt, L., Spehar, B., Van Engen, K. J., Sommers, M. S., and Peelle, J. E. (2021). "Pupillometry reveals cognitive demands of lexical competition during spoken word recognition in young and older adults," *Psychon. Bull. Rev.* (published online).
- Meister, H., Landwehr, M., Pyschny, V., Walger, M., and 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.
- Moore, B. C. J., and Carlyon, R. P. (2005). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, and R. R. Fay (Springer, New York, NY), pp. 234–277.
- Morrett, L. M., Roche, J. M., Fraundorf, S. H., and Partland, J. C. Mc. (2020). "Contrast is in the eye of the beholder: Infelicitous beat gesture increases cognitive load during online spoken discourse comprehension," *Cog. Sci.* **44**, e12912.
- Nooteboom, S. (1997). "The prosody of speech: Melody and rhythm," in *The Handbook of Phonetics Sciences*, edited by W. J. Hardcastle and J. Laver (Blackwell, Oxford, UK), pp. 640–673.
- Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., and Kramer, S. E. (2017). "Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation," *Hear. Res.* **351**, 68–79.
- Oldfield, R. C. (1954). "Memory mechanisms and the theory of schemata," *Br. J. Psychol.* **45**, 14–23.

- Oxenham, A. J., and Kreft, H. (2014). "Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing," *Trends Hear.* **18**, 2331216514553783.
- Peng, S. C., Chatterjee, M., and Lu, N. (2012). "Acoustic cue integration in speech intonation recognition with cochlear implants," *Trends Amplif.* **16**, 67–82.
- Peng, S. C., Lu, N., and Chatterjee, M. (2009). "Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners," *Audiol. Neurotol.* **14**, 327–337.
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., and Wingfield, A. (2016). "Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL)," *Ear Hear* **37**(Suppl.1), 5s–27s.
- Piquado, T., Isaacowitz, D., and Wingfield, A. (2010). "Pupillometry as a measure of cognitive effort in younger and older adults," *Psychophysiology* **47**, 560–569.
- Potter, M. C. (1993). "Very short-term conceptual memory," *Mem. Cognit.* **21**, 156–161.
- Potter, M. C., and Lombardi, L. (1990). "Regeneration in the short-term recall of sentences," *J. Mem. Lang.* **29**, 633–654.
- Reilly, J., Kelly, A., Kim, S. H., Jett, S., and Zuckerman, B. (2019). "The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry," *Behav. Res. Methods* **51**, 865–878.
- Richter, M. (2016). "The moderating effect of success importance on the relationship between listening demand and listening effort," *Ear Hear.* **37**(Suppl 1), 111S–117S.
- Roberts, D. S., Lin, H. W., Herrmann, B. S., and Lee, D. J. (2013). "Differential cochlear implant outcomes in older adults," *Laryngoscope* **123**, 1952–1956.
- Rogers, C. S., Jacoby, L. L., and Sommers, M. S. (2012). "Frequent false hearing by older adults: The role of age differences in metacognition," *Psychol. Aging* **27**, 33–45.
- Rogers, C. S., and Wingfield, A. (2015). "Stimulus-independent semantic bias misdirects word recognition in older adults," *J. Acoust. Soc. Am.* **138**, EL26–EL30.
- Rosenfield, M., Jahan, S., Nunez, K., and Chan, K. (2015). "Cognitive demand, digital screens and blink rate," *Comput. Hum. Behav.* **51**, 403–406.
- Sagi, E., Kaiser, A. R., Meyer, T. A., and Svirsky, M. A. (2009). "The effect of temporal gap identification on speech perception by users of cochlear implants," *J. Speech Lang. Hear. Res.* **52**, 385–395.
- Savin, H. B., and Perchonock, E. (1965). "Grammatical structure and the immediate recall of English sentences," *J. Verb. Learn. Verb. Behav.* **4**, 348–353.
- Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication* (The University of Illinois Press, Urbana, IL).
- Shannon, R. V. (1989). "Detection of gaps in sinusoids and pulse trains by patients with cochlear implants," *J. Acoust. Soc. Am.* **85**, 2587–2592.
- Shannon, R. V. (1993). "Psychophysics," in *Cochlear Implants: Audiological Foundations*, edited by R. S. Tyler (Singular, San Diego, CA), pp. 357–388.
- Sharpe, V., Fogerty, D., and den Ouden, D.-P. (2017). "The role of fundamental frequency and temporal envelope in processing sentences with temporary syntactic ambiguities," *Lang. Speech* **60**, 399–426.
- Shattuck-Hufnagel, S., and Turk, A. E. (1996). "A prosody tutorial for investigators of auditory sentence processing," *J. Psycholinguist. Res.* **25**, 193–247.
- Skinner, M. W., Holden, L. K., Fourakis, M. S., Hawks, J. W., Holden, T., Arcaroli, J., and Hyde, J. (2006). "Evaluation of equivalency in two recordings of monosyllabic words," *J. Am. Acad. Audiol.* **17**, 350–366.
- Skinner, M. W., Holden, T. A., Whiting, B. R., Voie, A. H., Brunsdon, B., Neely, J. G., Saxon, E. A., Hullar, T. E., and Finley, C. C. (2007). "In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea," *Ann. Otol. Rhinol. Laryngol.* **116**, 2–24.
- Speer, S. R., Kyelgaard, M. M., and Dobroth, K. M. (1996). "The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities," *J. Psycholinguist. Res.* **25**, 247–268.
- Steinhauer, K., Abada, S. H., Pauker, E., Itzhak, I., and Baum, S. R. (2010). "Prosody syntax interactions in aging: Event-related potentials reveal dissociations between on-line and off-line measures," *Neurosci. Lett.* **472**, 133–138.
- Steinhauer, K., Alter, K., and Friederici, A. D. (1999). "Brain potentials indicate immediate use of prosodic cues in natural speech processing," *Nat. Neurosci.* **2**, 191–196.
- Svirsky, M. A., Fitzgerald, M. B., Sagi, E., and Glassman, E. K. (2015). "Bilateral cochlear implants with large asymmetries in electrode insertion depth: Implications for the study of auditory plasticity," *Acta Otolaryngol.* **135**, 354–363.
- Titone, D. A., Koh, C. K., Kjelgaard, M. M., Bruce, S., Speer, S. R., and Wingfield, A. (2006). "Age-related impairments in the revision of syntactic misanalyses: Effects of prosody," *Lang. Speech* **49**, 75–99.
- Townshend, B., Cotter, N., van Compernelle, D., and White, R. L. (1987). "Pitch perception by cochlear implant subjects," *J. Acoust. Soc. Am.* **82**, 106–115.
- Tyler, L. K., Voice, J. K., and Moss, H. E. (2000). "The interaction of meaning and sound in spoken word recognition," *Psychon. Bull. Rev.* **7**, 320–326.
- Tyler, R. S., Moore, B. C., and Kuk, F. K. (1989). "Performance of some of the better cochlear-implant patients," *J. Speech Lang. Hear. Res.* **32**, 887–911.
- van der Wel, P., and Steenbergen, H. (2018). "Pupil dilation as an index of effort in cognitive control tasks: A review," *Psychonom. Bull. Rev.* **25**, 2005–2015.
- van Rooij, J. C. G. M., and Plomp, R. (1991). "The effect of linguistic entropy on speech perception in noise in young and elderly listeners," *J. Acoust. Soc. Am.* **90**, 2985–2991.
- Wagner, M., and Watson, D. G. (2010). "Experimental and theoretical advances in prosody: A review," *Lang. Cognit. Process.* **25**, 905–945.
- Wang, Y., Zekveld, A. A., Naylor, G., Ohlenforst, B., Jansma, E. P., Lorens, A., Lunner, T., and Kramer, S. E. (2016). "Parasympathetic nervous system dysfunction, as identified by pupil light reflex, and its possible connection to hearing impairment," *PLoS One* **11**, 1–26.
- Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., and Lunner, T. (2018). "Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio in the pupillary response while performing a speech-in-noise test," *Hear. Res.* **369**, 67–78.
- Wingfield, A., Tun, P. A., and Rosen, M. J. (1995). "Age differences in veridical and reconstructive recall of syntactically and randomly segmented speech," *J. Gerontol. Psychol. Sci.* **50B**, P257–P266.
- Wingfield, A., Wayland, S. C., and Stine, E. A. L. (1992). "Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentence," *J. Gerontol. Psychol. Sci.* **47**, P350–P356.
- Winn, M. B., Edwards, J. R., and Litovsky, R. Y. (2015). "The impact of auditory spectral resolution on listening effort revealed by pupil dilation," *Ear Hear.* **36**, 153–165.
- Winn, M. B., and Moore, A. N. (2018). "Pupillometry reveals that context benefit in speech perception can be disrupted by later-occurring sounds, especially in listeners with cochlear implants," *Trends Hear.* **22**, 1–22.
- Winn, M. B., and Teece, K. H. (2021). "Listening effort is not the same as speech intelligibility score," *Trends Hear.* **25**, 233121652110276.
- Winn, M. B., Wendt, D., Koelewijn, T., and Kuchinsky, S. E. (2018). "Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started," *Trends Hear.* **22**, 1–32.
- Wood, J. G., and Hassett, J. (1983). "Eyeblinking during problem solving: The effect of problem difficulty and internally vs externally directed attention," *Psychophysiology* **20**, 18–20.
- Yukawa, K., Cohen, L., Bulamey, P., Pyman, B., Tungvachirakul, V., and O'Leary, S. (2004). "Effects of insertion depth of cochlear implant electrodes upon speech perception," *Audiol. Neurotol.* **9**, 163–172.
- Zachary, R. A. (1991). *Shipley Institute of Living Scale: Revised Manual* (Western Psychological Services, Los Angeles, CA).
- Zekveld, A. A., Koelewijn, T., and Kramer, S. E. (2018). "The pupil dilation response to auditory stimuli: Current state of knowledge," *Trends Hear.* **22**, 1–25.
- Zekveld, A. A., and Kramer, S. E. (2014). "Cognitive processing load across a wide range of listening conditions: Insights from pupillometry," *Psychophysiology* **51**, 277–284.
- Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2011). "Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response," *Ear Hear.* **32**, 498–510.
- Zeng, F.-G., and Galvin, J. J. (1999). "Amplitude mapping and phoneme recognition in cochlear implant listeners," *Ear Hear.* **20**, 60–74.