The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation

Matthew B. Winn,[1,2] Jan R. Edwards,[1,3] and Ruth Y. Litovsky[1–3]

Objectives: This study measured the impact of auditory spectral resolution on listening effort. Systematic degradation in spectral resolution was hypothesized to elicit corresponding systematic increases in pupil dilation, consistent with the notion of pupil dilation as a marker of cognitive load.

Design: Spectral resolution of sentences was varied with two different vocoders: (1) a noise-channel vocoder with a variable number of spectral channels; and (2) a vocoder designed to simulate front-end processing of a cochlear implant, including peak-picking channel selection with variable synthesis filter slopes to simulate spread of neural excitation. Pupil dilation was measured after subject-specific luminance adjustment and trial-specific baseline measures. Mixed-effects growth curve analysis was used to model pupillary responses over time.

Results: For both types of vocoder, pupil dilation grew with each successive degradation in spectral resolution. Within each condition, pupillary responses were not related to intelligibility scores, and the effect of spectral resolution on pupil dilation persisted even when only analyzing trials in which responses were 100% correct.

Conclusions: Intelligibility scores alone were not sufficient to quantify the effort required to understand speech with poor resolution. Degraded spectral resolution results in increased effort required to understand speech, even when intelligibility is at 100%. Pupillary responses were a sensitive and highly granular measurement to reveal changes in listening effort. Pupillary responses might potentially reveal the benefits of aural prostheses that are not captured by speech intelligibility performance alone as well as the disadvantages that are overcome by increased listening effort.

Key words: Cochlear implant, Listening effort, Pupil dilation, Pupillometry, Spectral degradation, Spectral resolution, Vocoder.

(Ear & Hearing 2015;36;e153–e165)

INTRODUCTION

Listening effort is a component of auditory perception involving cognitive processing or cognitive load. People with hearing impairment (HI) routinely report that they experience elevated listening effort and demonstrate increased effort associated with speech perception (Kramer et al. 1997). Numerous studies suggest a connection between elevated effort and negative psychosocial consequences (Edwards 2007), including greater need for recovery after work (Nachtegaal et al. 2009), increased incidence of stress-related sick leave (Kramer et al. 2006), increased unemployment among young adults (Parving & Christensen 1993; Järvelin et al. 1997), and early retirement (Danermark & Gellerstedt 2004). Additionally, people with HI experience higher degrees of social isolation (Demorest & Erdman 1987; Grimby & Ringdahl 2000). While increased listening effort alone cannot account for all of these problems, these findings have helped to contribute to an increased awareness of this problem. There is a pressing need to quantify aspects of hearing loss that are not reflected by speech audiometry alone.

Spectral resolution is the ability of a listener to perceptually resolve sounds of different frequencies. This ability underlies the capacity to distinguish acoustically similar consonant pairs such as /b/-/d/ and /t/-/k/, which are notoriously difficult for people with HI and people with cochlear implants (CIs) (Dubno et al. 1982; Munson et al. 2003). Spectral resolution is known to be particularly poor in CIs (Nelson et al. 1995; Boëx et al. 2003, Henry et al. 2005; Won et al. 2007; Jones et al. 2013). Toward the goal of better understanding the challenges of cochlear implantation, the present work was designed to explore the relationship between spectral resolution and listening effort.

Virtually all of the published works on listening effort and HI have been conducted with people who do not use CIs. Unimplanted individuals with HI are likely to have poorer spectral resolution than their normal-hearing (NH) counterparts (Glasberg & Moore 1986) but are still likely to have better spectral resolution than CI users (c.f., Henry et al. 2005). It is therefore expected that spectral resolution–related difficulties encountered by people with HI could be magnified for people who use CIs.

Numerous studies in recent years have applied dual-task paradigms to study the listening effort in the context of communication disorders. For example, a primary task measuring speech perception is assumed to occupy a certain amount of cognitive load, while a secondary task is completed using the remaining capacity (Kahneman 1973). Increases in the cognitive load demanded by the primary task are observed as decrements in performance in the secondary task. There are numerous interpretations of how cognitive resources are allocated (i.e., whether both tasks draw from the same available load or work somewhat in parallel). Nonetheless, there is generally sufficient literature to support the notion that dual-task experiments are a reliable metric for cognitive load.

With regard to speech perception, dual-task experiments have shown that, among other things, binaural listening is easier than monaural listening (Feuerstein 1992), hearing loss interferes with visual memory (Rakerd et al. 1996), and the presence of masking noise demands extra effort that is not alleviated by the availability of visual cues (Picou et al. 2011). Sarampalis et al. (2009) suggested that listening effort can be dissociated from speech intelligibility performance, as is sometimes anecdotally observed in the audiology clinic. For example, it is often the case that patients exhibit stress and fatigue when engaged in word recognition testing, yet their responses might still be correct. Sarampalis et al. showed that for listeners with NH in poor signal-to-noise ratio conditions, digital noise reduction resulted in no measured benefit in terms of speech perception

[1]Waisman Center, University of Wisconsin-Madison, Madison, Wisconsin, USA; [2]Department of Surgery, University of Wisconsin-Madison, Madison, Wisconsin, USA; and [3]Department of Communication Sciences and Disorders, University of Wisconsin-Madison, Madison, Wisconsin, USA.
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).

performance but did result in better performance on word-memory tasks and also resulted in shorter reaction times for processing of visual stimuli presented concurrently with the speech. Such results suggest that benefits of changes in signal processing might emerge in ways other than intelligibility scores.

Spectral degradation is an ideal aspect of speech signals to explore with measures of listening effort because the signal quality can be reduced without necessarily resulting in poor intelligibility (Shannon et al. 1995). Pals et al. (2013) investigated the impact of spectral resolution on listening effort by using NH listeners who heard vocoded speech with a variable number of spectral channels. Pals et al. characterized the vocoded speech as "CI simulations," determined by the historical match in performance between NH listeners presented with 8-channel vocoded speech and successful CI listeners (Friesen et al. 2001). To measure effort, they used a dual-task paradigm where listeners heard speech while completing a visual mental-rotation task or a rhyme-judgment task. Response time (RT) measurements for both tasks were used as measures of the effort exerted by subjects when listening to degraded speech. A systematic decline (improvement) in RT was observed as the number of channels was increased from 2 to 8 channels in 2-channel steps, suggesting that signals with better resolution required less effort to be understood. In that study, a plateau in intelligibility was observed at 6 channels, implying that the reduction in effort from 6 to 8 channels was captured by RT but not by intelligibility scores. No RT decrease was observed for signals with more than 8 channels (consistent with the commonly observed pattern described by Friesen et al. 2001). These results suggest that spectral resolution has an impact on listening effort but leave open the possibility that effort continues to shrink as resolution is improved beyond 8 channels or beyond the plateau in intelligibility. Those possibilities are explored in the present study.

Another reliable index of cognitive load is pupillometry or the measurement of pupil dilation. This technique has been used in a variety of listening tasks to gauge how listening effort is affected by sentence intelligibility (Zekveld et al. 2010), different masker types (Koelewijn et al. 2012), HI (Kramer et al. 1997; Zekveld et al. 2011), and cognitive function (Zekveld et al. 2011). More generally, the pupillary response has also been associated with measures of memory load (Kahneman & Beatty 1966), selective attention (Hillyard et al. 1973), motivation (Kahneman et al. 1968), and linguistic coherence of stimuli (Schluroff 1983). In each of these domains, the pupillary response corresponds to an intuitive understanding of changing cognitive demands (e.g., increased pupil dilation for a longer string of digits to be memorized, compared to a shorter string). Pupil dilation occurs in response to changes in attended stimuli and can be absent when changes occur in unattended stimuli (Hillyard et al. 1973). Pupil dilation is also increased when subjects exert greater effort to solve arithmetic problems (Polt 1970). Finally, pupillometry is a sensitive tool that can be used to measure differences across and within individuals (Beatty 1982), implying that changes in signal quality for individuals could potentially result in differences in pupil dilation within the same experiment.

The pupillary response is affected by numerous influences apart from cognitive load, including the pupillary light reflex, and other responses related to various states of arousal (consistent with sympathetic nervous system innervation). To exploit pupillometry as a window into listening effort, these ongoing influences on pupil diameter need to be rigorously controlled, and stimulus-time-locked averages are used, just as for evoked-potential methods such as EEG. As Beatty (1982) remarked, "A task-evoked pupillary response bears the same relation to the pupillary record from which it is derived as does an event-related brain potential to spontaneous electroencephalographic activity."

This motivation of the present study was similar to that of the study by Pals et al. (2013), but with a number of important differences. First, listening effort in the present study was measured via pupil dilation rather than by a secondary task; in this study, there was no task except for speech perception/sentence repetition. Second, time-series (growth curve) analyses were used to model pupil dilation over time, which is distinctly different from the analysis of mean pupil dilation in most pupillometry studies. Finally, while a conventional channel vocoder is used to vary spectral resolution according to the number of carrier channels, a second experiment used a separate vocoder that more closely replicates the front-end processing of modern CIs, where resolution is controlled not by the number of channels, but instead via channel carrier bandwidth, to explicitly investigate the role of current spread as a major contributor to spectral resolution in CI listeners.

EXPERIMENT 1

In the first experiment, pupillary responses were measured in response to speech processed using the conventional channel vocoder, where spectral resolution was altered by changing the number of spectral channels. The experiment was conducted with approval from the Institutional Review Board at the University of Wisconsin-Madison.

MATERIALS AND METHODS

Participants

Twenty young adults (age range, 18–33 years) with NH participated. A majority of participants were recruited from undergraduate courses in the Communication Sciences and Disorders department at the University of Wisconsin-Madison; most participants were students who were eligible for undergraduate course credit for participation. All listeners were native speakers of American English, and listeners with fluency in any other language were excluded from the analysis. Nearly all the listeners were naive to vocoded speech.

Stimuli and Processing

Stimuli were taken from the Institute of Electrical and Electronics Engineers (IEEE)/Harvard sentence corpus (IEEE 1969), which consists of lists of sentences with five keywords each. The first 40 (out of 72) 10-sentence lists were used. All were digitized original recordings from the IEEE, spoken by a male talker. Sentence duration ranged from roughly 2 to 4 sec, with an average duration of 2.78 sec. These sentences are generally regarded as relatively easy, with most containing clear semantic coherence (e.g., "Steam hissed from the broken valve").

Sentences were vocoded using a conventional method (c.f., Shannon et al. 1995; Xu et al. 2005), whereby the spectrum was divided into a variable number of frequency bands that occupied equivalent cochlea space (determined according to the function from Greenwood 1990) between 150 and 8000 Hz.

The amplitude envelope of each band was extracted using half-wave rectification and a 300-Hz low-pass filter. The envelope was used to modulate a band of noise whose bandwidth was equivalent to the corresponding analysis filter. Essentially, the fine spectral structure is replaced by noise, and the temporal envelope is mostly preserved. Various studies (e.g., Friesen et al. 2001; Xu et al. 2005) have shown that speech intelligibility improves with increasing number of vocoder channels. Stimuli were either unprocessed/normal or vocoded with 32, 16, 8, or 4 channels using the AngelSim software (Fu 2013). These parameters have been shown to elicit a moderate range of intelligibility (c.f., Shannon et al. 2004) and comprise a reasonably wide range of perceived difficulty.

Procedure

Participants were seated in a quiet, dimly-lit room with low reverberation, inside a large research suite, at a distance of 60 cm from a Tobii T60 XL distal eye-tracker whose monitor filled most of the visual field. Consistent with Chapman et al. (1999) and Zekveld et al. (2010), the luminance of the visual field was controlled to avoid the floor and ceiling of the range of pupil size, which is affected relatively more strongly by the light reflex than by cognitive load. For each participant, the color of the computer screen was adjusted from black to white in successive shades of gray to elicit the range of pupil sizes attributable to the light reflex. The luminance required to elicit an intermediate pupil size (midway between the minimum and maximum measured sizes) was calculated, and the corresponding shade of gray used as screen background color for the rest of the experiment. Typically, screen color was adjusted to a level between {75, 75, 75} and {130, 130, 130} in the red/green/blue color mode.

Audio signals were presented using the Tobii T60 XL free-field loudspeakers. Calibration to 65 dB was performed using a 1-kHz tone matched to the intensity of the speech stimuli. Trials consisted of an auditory stimulus (a sentence), during which a red cross appeared on the screen. The color of the cross changed from red to green (controlled for equal luminance) 1.5 s after the conclusion of the sentence to signal that the participant should repeat back the sentence. Although there was no explicit test for colorblindness, all participants demonstrated the ability to reliably respond to the color change. Participants were encouraged to guess at sentences that they did not completely recognize.

Following each sentence, the participant verbally gave their response, and the experimenter tracked incorrect words. The response and scoring took roughly 4 to 8 sec, depending on the speed of response and scoring. Following the coding, the trial was advanced following a 2-sec delay. Baseline pupil dilation data were collected over the following 2 sec. Following the baseline period, there was a variable amount of silence because stimuli were offset aligned within a 5-sec window (i.e., 4-sec stimuli were preceded by 1 sec of extra silence, and 3-sec stimuli were preceded by 2 sec of extra silence).

Before testing, each subject participated in a brief practice session consisting of 15 trials, which included samples from each vocoder condition. The practice session procedure (including on-screen response prompt cues) was exactly the same as that for the test session, except that the stimulus was presented in written form after the participant guessed the sentence. Following the practice session, each participant completed 60 trials: 12 trials at each level of spectral resolution (the four vocoder conditions, and natural unprocessed speech). The trials were consecutive tokens from the IEEE sentence corpus that were organized into 10 blocks of six sentences; blocks were organized by spectral resolution, and the ordering of conditions was randomized. Six different groups of sentences were randomly assigned to participants so that a wide range of the corpus was used.

Analysis

Average pupil diameter was calculated for the first 2 sec of each trial; diameter for the remainder of the trial was subtracted from that baseline level to obtain a measure of relative pupil dilation. Pupil diameter tracings were "de-blinked" by detecting short (<300 msec) gaps of missing values, interpolating between diameter measurements 100 msec before the blink and 166 msec after the blink. The "gap" was expanded to account for brief underestimations of pupil diameter caused by the eyelids obscuring the pupil, consistent with techniques used by Siegle et al. (2008) and Zekveld et al. (2010). Values were smoothed using a symmetrical 17-sample running average to remove high-frequency artifacts.

Consistent with the analyses by Zekveld et al. (2010, 2011) and Koelewijn et al. (2012), mean and maximum pupil dilation was also analyzed during a brief time window centered on the pause between stimulus offset and the prompt to repeat the sentence. This time window is thought to be used by listeners for rehearsal/planning of the response and typically contains a local maximum in the pupillary response (Zekveld et al. 2010).

Pupillometry data in listening effort experiments have historically been analyzed using mean and maximum pupil dilation within a discrete window proximal to response prompt; limitations of this approach have been explained in some recent articles (c.f., Barr 2008; Mirman et al. 2008; Kuchinsky et al. 2013; Mirman 2014). For example, the binning of pupil dilation within a single time window eliminates the exquisite granularity offered by the data and discards any change over time. Reducing the densely sampled data series into a single number potentially diminishes the statistical power by ignoring changes in dilation *curve morphology* that do not result in changes in *mean* or *maximum* dilation. Furthermore, inclusion of multiple samples in a single time bin violates the assumption of independent samples, as timepoint [*t*] is related to timepoint [*t*-1], [*t*-2], etc. Finally, because each condition starts at an equivalent baseline, real differences are likely to be compressed in a long-time window. The growth of pupil diameter over time assumes a functional form, which is conveniently modeled by time-series analysis (Kuchinsky et al. 2013). It is reasonable to conclude that for very simple procedures such as sentence recognition, not all of these issues would *prevent* the identification of notable experimental results. However, the analysis presented here provides an approach that can be used to model and examine any functional form of pupillary data, especially in cases where mean and maximum pupil dilation tells only part of the story.

To address some of the limitations of conventional time-window analysis, growth curve analysis (Mirman et al. 2008, 2014) was applied to the pupil dilation data within the time window from −2000 msec to +500 msec relative to sentence offset. This window was chosen because pupil responses generally remained at baseline until roughly −2000 msec and generally reached a peak around +500 msec relative to stimulus offset.

This time window is presumably affected only by the perception of the stimulus and not by the motor planning and execution of the verbal response.

The overall time course of pupil dilation was modeled using orthogonal time polynomials, which are transformations of natural polynomials that make the individual time terms independent. The ultimate analysis used a second-order polynomial, with linear and quadratic components to model slope and acceleration/deceleration, respectively. Higher-order components did not provide a significant improvement to the model fit and were thus not included in the analysis. An illustration of polynomial curve fitting is provided in Supplemental Digital Content 1–3 (http://links.lww.com/EANDH/A176; http://links.lww.com/EANDH/A177; http://links.lww.com/EANDH/A178).

The statistical model of pupillary responses had a mixed-effects structure, where fixed effects included spectral resolution (number of vocoder channels; within participants) and its interaction with both polynomial time terms. The linear term represents growth of pupil diameter over time, and the quadratic term represents change in the rate of growth. Because these polynomials were orthogonal, the intercept term for each condition represents a time-independent overall effect, similar to "area under the curve." The notable effects, however, were the interactions between condition and polynomial time terms, that is, the growth of pupil dilation over time as a function of condition. The model also included random effects of participant and both time terms. That is, overall levels and rates of pupil dilation were allowed to vary across participants to model individual variability internal to this participant sample that is not generalizable to the independent variables themselves. The lme4 package (Bates et al. 2013) in R (R development core team 2013) was used for the analysis, and *p* values were estimated using the *z* distribution as an approximation for the *t* distribution; degrees of freedom in mixed-effects models are ill defined, so the *z* approximation is a mechanism that has been used to estimate significance in cases of hierarchical model structure (Mirman 2014).

Speech intelligibility (accuracy of listeners' responses) was measured as the number of key words (out of five) correctly identified in each sentence. Function words (e.g., a/an/the/for, etc.) did not count toward the performance score, consistent with scoring of these materials on standardized measures (c.f., Killion et al. 2004). Hierarchical (mixed-effects) linear modeling was used to model maximum pupil size predicted by fixed effects of intelligibility and condition, with random intercepts for each participant in each condition. This mixed-effects structure is appropriate when measurements are made on clusters of related data points that vary on multiple levels to avoid letting the impact of one level (e.g., condition) be misinterpreted as an effect of another level that is nested within the first (e.g., intelligibility), a la Simpson's paradox (Spellman et al. 2001).

RESULTS

There was a systematic relationship between the five conditions of spectral resolution and their corresponding pupillary responses, as shown in Figure 1. Responses from each condition were undifferentiated at stimulus onset and then grew at different rates to reach a local maximum level just after the offset of the stimulus. In all conditions, pupillary responses showed a local decline shortly after the response prompt and then increased dilation at the onset of the verbal response. The relationship between spectral resolution and pupil dilation was maintained even when analyzing only trials where participants correctly repeated the sentences (Fig. 1, right panel). Figure 2 shows the average mean and maximum pupil dilations across all trials, within the analysis window that included stimulus offset and response prompt (−500 to 2000 msec relative to sentence offset).

The statistical model fit to the data is described in detail in Table 1. As can be seen in Figure 3, the model provided an excellent fit to the data in the range between baseline and maximum dilation. For each of the successively poorer-resolution vocoder conditions, there was a significant effect of condition on the intercept term, indicating greater overall pupil dilation. Post hoc testing was performed by changing the default condition comparison, cycling through each of the five conditions. Slope increased as spectral resolution grew poorer; each condition was significantly different from each other condition on the basis of linear time ($p < 0.001$ for all comparisons). These results suggest that it is not merely the absolute size of pupil dilation that changes with spectral resolution; the *rate* of pupil dilation also changes significantly. Statistically significant differences for the quadratic term (change in rate) were more

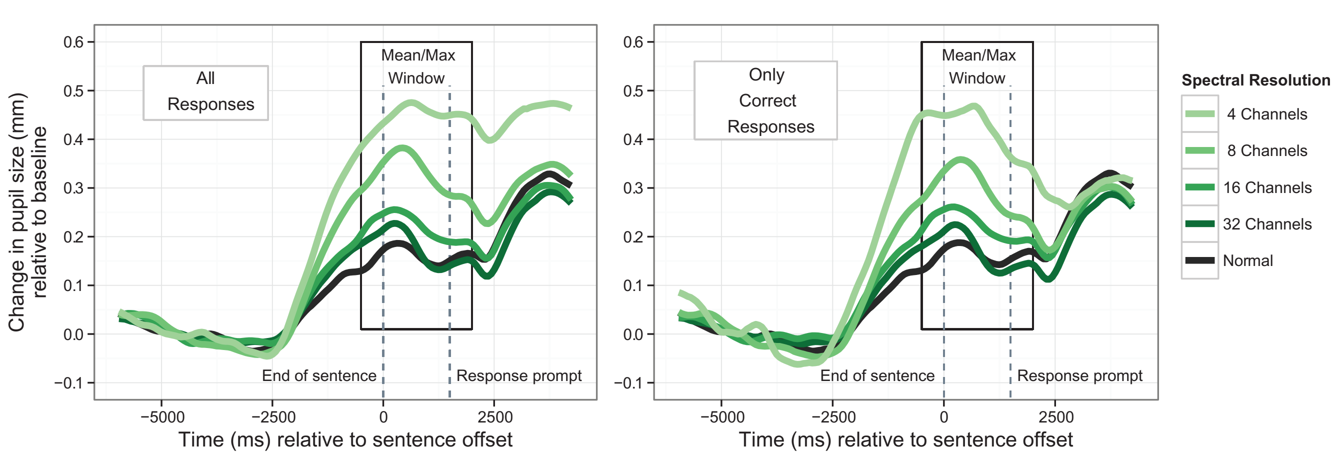

Fig. 1. Change in pupil dilation over time as a function of spectral resolution (number of vocoder channels) in Experiment 1. Time (milliseconds) is plotted relative to stimulus offset. The boxed area represents the time window traditionally used for aggregated analysis. Left panel, data from all trials. Right panel, data only from trials where the entire sentence was identified correctly.

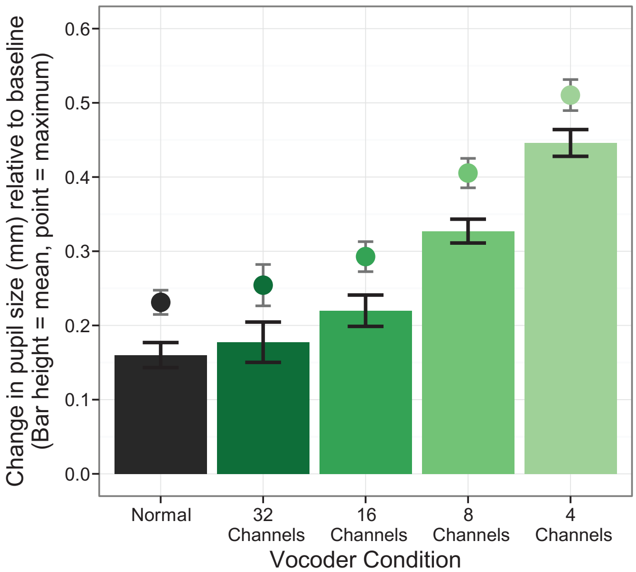

Fig. 2. Change in pupil dilation as a function of spectral resolution in Experiment 1. Data represent mean (bar height) and maximum (point) pupil dilation during the time window beginning 500 msec before and ending 2000 msec after stimulus offset (Fig. 1). Error bars represent ± 1 SE.

modest, with most differences arising between the most difficult (4 channel) condition and the other conditions.

Speech Intelligibility/Accuracy

Although speech intelligibility was poorer in conditions of degraded spectral resolution, intelligibility by itself was not found to be a significant predictor of maximum pupil dilation ($p = 0.16$). In other words, intelligibility was lower in the severely degraded conditions, but there was no further relationship between intelligibility and pupil size apart from that predicted by condition. Even when excluding the fixed effect of condition, the effect of intelligibility still did not reach significance ($t = -1.40$, $p = 0.16$). Thus, intelligibility alone did not have an independent significant effect on pupil size separate from the effect of the spectral degradation itself. Figure 4 displays the individual intelligibility scores against measures of maximum pupil dilation across listening conditions. It can be seen that while the intelligibility is lower in the two most degraded conditions, the relationship between intelligibility and pupil dilation within each condition is weak.

EXPERIMENT 1 SUMMARY

Pupil dilation was measured in response to speech processed using a vocoder where the spectral resolution was altered by changing the number of analysis and carrier channels. A systematic relationship was observed between spectral resolution and pupillary response, suggesting that spectrally degraded signals demanded more listening effort. Within each vocoder condition, pupil dilation was not correlated with sentence key word intelligibility, suggesting that intelligibility scores alone are not sufficient to predict listening effort. Furthermore, even when restricting analysis to trials in which listeners correctly reported the sentences, the effect of spectral resolution still emerged, suggesting that listeners can exert different amounts of effort when listening to speech even if their scores remain at 100%.

EXPERIMENT 2

Experiment 2 was designed specifically to better understand spectral degradation that results from CI stimulation.

TABLE 1. Generalized linear mixed-effects model formula and summary output for growth curve analysis in Experiment 1

Formula:
Pupil dilation = (Intercept + Condition) + *# intercept*
time1*(β_{time}1 + β_{time}1:Condition) + *# slope*
time2*(β_{time}2 + β_{time}2:Condition) + *# acceleration*
r(subject-level intercept) *# random error*
r(subject-level slope and acceleration) *# random error*

Formula code: PupilDilation ~ (time1 + time2) + NumChannels + time1:NumChannels + time2: NumChannels + (1 | Subject) + (time1 + time2 | Subject)

Term	Estimate	SE	*t*	*p*
Intercept	0.119	0.014	8.4	<0.001‡
Time1	0.291	0.039	7.5	<0.001‡
Time2	−0.024	0.020	−1.24	0.215
NumChannels 32	0.035	0.004	9.17	<0.001‡
NumChannels 16	0.057	0.004	14.7	<0.001‡
NumChannels 8	0.123	0.004	31.82	<0.001‡
NumChannels 4	0.196	0.004	49.7	<0.001‡
Time1:NumChannels 32	0.063	0.023	2.67	0.008†
Time1:NumChannels 16	0.123	0.024	5.22	<0.001‡
Time1:NumChannels 8	0.331	0.024	14.05	<0.001‡
Time1:NumChannels 4	0.432	0.024	18.05	<0.001‡
Time2:NumChannels 32	−0.049	0.023	−2.08	0.037*
Time2:NumChannels 16	−0.039	0.024	−1.66	0.097
Time2:NumChannels 8	−0.083	0.024	−3.55	0.000‡
Time2:NumChannels 4	−0.154	0.024	−6.45	<0.001‡

"Time 1" and "time 2" refer to linear and quadratic time polynomials, respectively. "(Intercept)," "time1," and "time2" as isolated model terms refer to the default condition, which was normal unprocessed speech. Estimates of interactions of these terms with the conditions reflect the change in the estimate of the named condition when compared against the default condition.
*$p < 0.05$.
†$p < 0.01$.
‡$p < 0.001$.

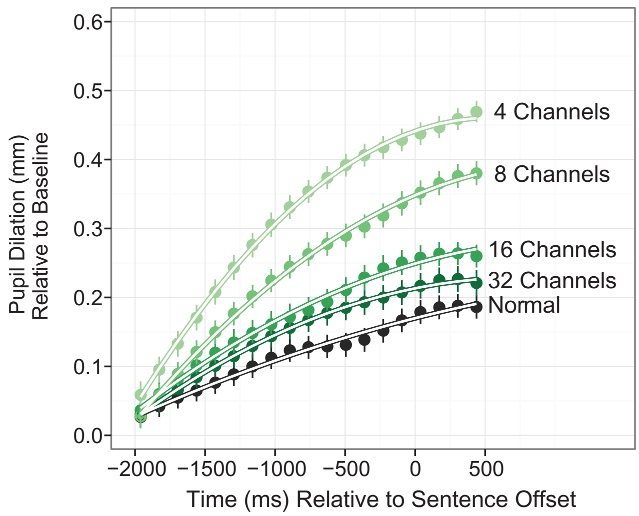

Fig. 3. Model fits (open lines) overlaid on aggregated data (points with ± SE lines) representing change in pupil diameter over time in Experiment 1. Time (milliseconds) is plotted relative to stimulus offset.

Accordingly, the style of vocoder was changed to explicitly control factors that hypothetically mimic the degradations that arise from the implant processor and electrode array.

The vocoder that was used in Experiment 1 is commonly found in experiments aimed at understanding the limitations related to speech perception by CI listeners; generally, the number of channels in a noise vocoder can be altered until performance by NH listeners is roughly equal to that of CI listeners. In spite of the popularity of this approach (including its use by the current authors in numerous previous studies), there are a number of limitations that weaken its appeal as a realistic "CI simulation" for the purposes of this study, where the intent is not to simply equate intelligibility but to specifically model the mechanism and consequence of spectral degradation that CI listeners experience. For example, the actual frequency analysis filters in a CI speech processor can be dramatically different from those found in a channel vocoder, meaning that the channel-specific envelope information in a channel vocoder does not replicate the envelope information that would be represented by a CI. More importantly, it might be the case that the focus on number of channels distracts from another known problem, that of channel interaction, which arguably describes CI spectral degradation more appropriately. CI devices available to patients are limited to a few types, with a restricted set of options for "number" of electrode channels. However, measurements of spectral resolution vary greatly (Chatterjee & Shannon 1998; Cohen et al. 2003; Saoji et al. 2009; Bierer et al. 2010; Anderson et al. 2011). This variability in spectral resolution cannot be explained merely by number of physical electrodes available in the CI devices and is likely related more to the "effective" number of channels, which is affected by the degree of interaction between channels (Boëx et al. 2003; Jones et al. 2013).

A major factor contributing to channel interaction in CIs is the predominant use of monopolar stimulation mode, which, in the context of highly conductive perilymphatic fluid surrounding the implant array, creates broad electrical fields that stimulate poorly specified regions of the cochlea (Chatterjee & Shannon 1998; Boëx et al. 2003; Abbas et al. 2004; Zhu et al. 2012). Electrode interaction is shown to relate to the ability of a CI user to resolve spectral peaks represented by the electrode array (Jones et al. 2013) and is thus understandably linked with speech recognition performance in CI listeners (Stickney et al. 2006, but see Anderson et al. 2011 for a discussion of how this trend sometimes does not hold).

The approach taken in Experiment 2 is to not simply approximate the number of recovered spectral channels in a CI but to approximate the *mechanism* by which the spectral information is degraded. Consistent with Fu and Nogaki (2005) and Litvak et al. (2007), the carrier/synthesis filters of each channel were systematically varied in width to be more or less frequency specific while the front-end "CI processing" (i.e., the number of analysis and carrier channels and channel-selection strategy) remained constant, as it does across many CI listeners. Steeply-sloping filters had more spectral specificity, while shallower filters were wider and more likely to create distortion via channel overlap and thus simulating spread of neural excitation in the implanted cochlea (Bingabr et al. 2008). Using this approach, the "effective" number channels recovered is driven primarily by the interactions between channels rather than the explicit number of frequency analysis channels.

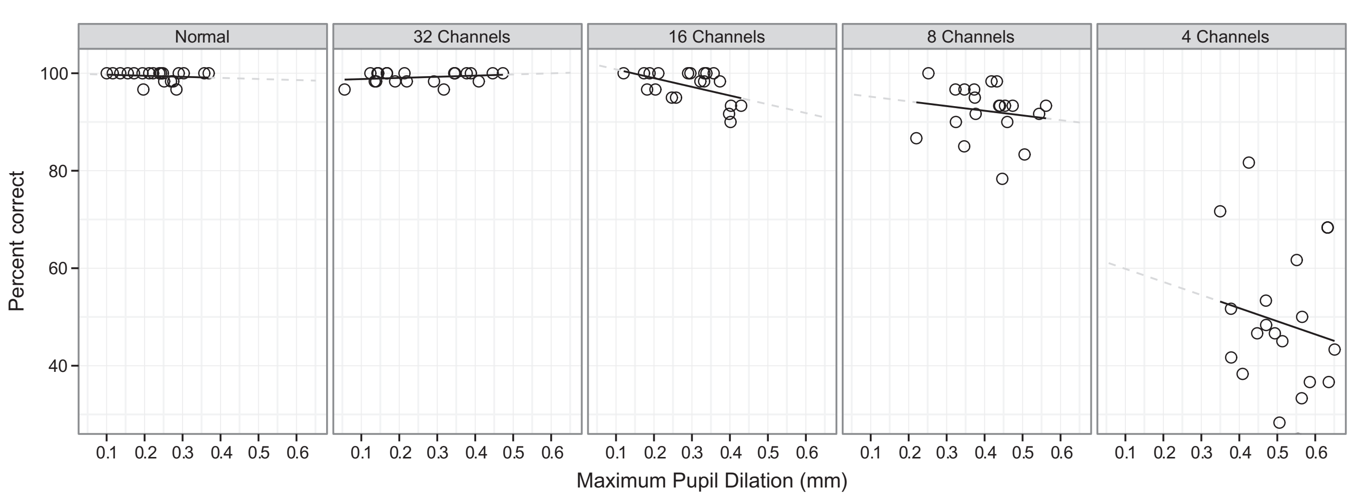

Fig. 4. Maximum pupil dilation (*x* axis) and word intelligibility (*y* axis) for each condition in Experiment 1. Regression lines are dashed when extrapolated beyond the observed data range.

Another important difference between conventional channel vocoders and the CI-simulation vocoder in Experiment 2 was the use of a channel peak-picking strategy. The most commonly used speech processing strategy for CI listeners in North America is the Advanced Combination Encoder (ACE) strategy, used in the Cochlear Nucleus family of devices. At manufacturer default settings, the ACE strategy delivers 8 peaks out of a possible 22 analysis channels for each analysis time window. This differs from conventional channel vocoders, where all analysis filters are represented in the carrier channels. By using a channel peak-picking strategy (i.e., *n-of-m* strategy) combined with systematic simulations of current spread, this experiment aimed to simulate not only the *performance* of CI listeners but also a specific factor mediating that performance in the context of a commonly used device. This vocoder will henceforth be referred to as a "CI-style" vocoder.

MATERIALS AND METHODS

Participants

Eighteen young adults with NH (age range, 18–34 years) participated in Experiment 2. The recruitment and selection criteria were the same as for Experiment 1, but no listeners in Experiment 1 participated in Experiment 2.

Stimuli and Processing

Stimuli consisted of the same basic sentence materials used in Experiment 1 but with a different vocoding technique. To use a peak-picking vocoder that approximates the ACE processing strategy, the software Praat (Boersma & Weenink 2013) was used to explicitly control various parameters of analysis and stimulation, including channel corner frequencies and rolling time windows. Out of 22 analysis channels, 8 carrier channels were activated for each time bin, consistent with the most common settings for CI clinical fitting. To select the 8 channels, the signal was first pre-emphasized to roughly flatten the spectral envelope from the standard −6 dB/octave to roughly 0 dB/octave. The intensity of each channel within each time bin was calculated, and the 8 channels with the highest intensity were retained while others were discarded. Each time bin was 30 msec and overlapped with adjacent time bins by 50% of its total duration. Carrier channels were noise bands shaped by 21-, 16-, 11-, and 7-dB/octave filters, representing best to poorest resolution. These parameter levels were chosen to approximate the level of difficulty encountered across the conditions in Experiment 1, following informal pilot listening.

Procedure

The procedure for Experiment 2 was exactly the same as for Experiment 1.

Analysis

Analysis for Experiment 2 was exactly the same as for Experiment 1.

RESULTS

As in Experiment 1, there was a clear relationship between spectral resolution and pupil dilation (Fig. 5). As vocoder channel width grew wider/more degraded, pupil dilation grew larger. When analyzing only trials with correct responses (Fig. 5, right panel), the effect of spectral resolution persisted, and the most difficult condition no longer elicited elevated pupil dilation during the time corresponding to the verbal response. Figure 6 shows the average mean and maximum pupil dilations across all trials, within the analysis window that included stimulus offset and response prompt.

Statistical analysis for pupil dilation data in Experiment 2 was exactly the same as for Experiment 1, with conditions defined by channel width instead of number of channels. Results demonstrated a systematic relationship between condition and pupil dilation over time (Figs. 5 and 6). The statistical model (Table 2) provided an excellent fit to the data (Fig. 7); for each of the successively poorer-resolution vocoder conditions, there was a significant effect of condition on the intercept term, indicating greater overall pupil dilation. Post hoc testing was performed by changing the default condition comparison, cycling through each of the five conditions. Those comparisons revealed significant differences between each pair of conditions on the basis of linear time (slope; all $p < 0.001$); slope increased as spectral resolution grew poorer. Statistical differences for the quadratic term only arose in a small number of condition comparisons. Table 2 contains a detailed explanation of the statistical model and summary output. Trends in Experiment 2 were weaker than those in the first experiment,

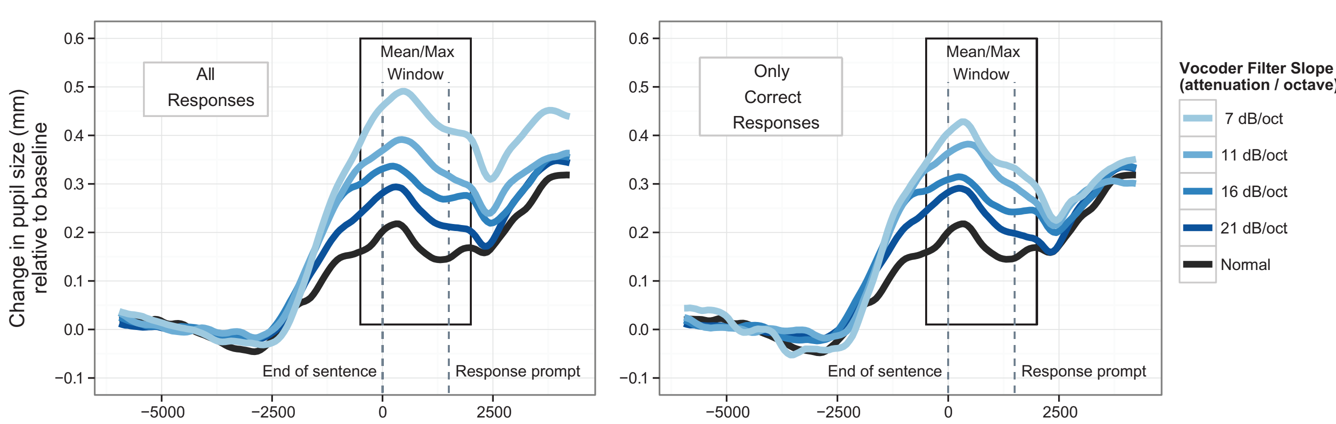

Fig. 5. Change in pupil dilation over time as a function of spectral resolution (vocoder carrier filter slope) in Experiment 2. Time (milliseconds) is plotted relative to stimulus offset. The boxed area represents the time window traditionally used for aggregated analysis. Left panel, data from all trials. Right panel, data only from trials where the entire sentence was identified correctly.

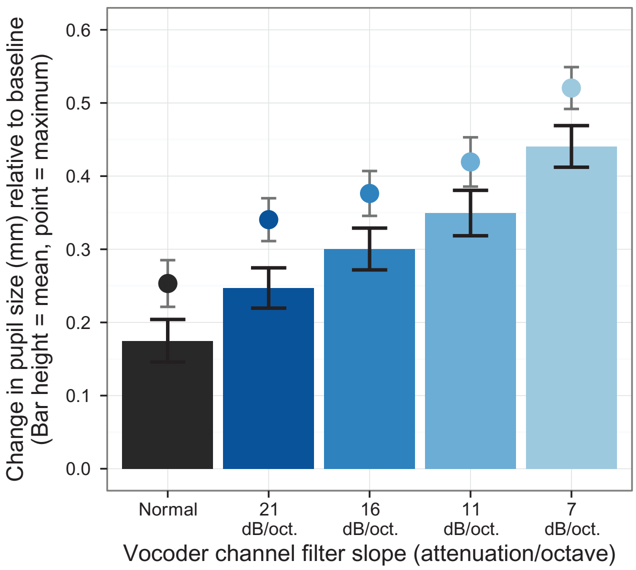

Fig. 6. Change in pupil dilation as a function of spectral resolution in Experiment 2. Data represent mean (bar height) and maximum (point) pupil dilation during the time window beginning 500 msec before and ending 2000 msec after stimulus offset (see Fig. 5). Error bars represent ± 1 SE.

likely because of the similarity of pupillary responses in the easiest three vocoder conditions (excluding the normal speech condition) in Experiment 2, which were less differentiated than those for the corresponding levels in Experiment 1.

Speech Intelligibility/Accuracy

Results were analyzed the same as for Experiment 1. Figure 8 shows the average speech intelligibility performance for each listener plotted against maximum pupil dilation (averaged across trials) in each condition. Again, intelligibility was not found to be an independent predictor of maximum pupil size; significance (p) values of 0.49, 0.51, 0.43, 0.45, and 0.50 were obtained for each successively degraded condition. Using a model without a separate fixed effect of condition, the effect of intelligibility still did not reach significance ($t = -1.40$, $p = 0.16$).

DISCUSSION

Listening effort is an everyday challenge for many people with HI and is garnering increased attention in the scientific literature. Pupillary measures offer a window into cognitive load during auditory tasks, with fine granularity and strong sensitivity. The present study suggests that spectral resolution is systematically related to listening effort. As spectral resolution was progressively degraded, listeners' pupil dilation progressively grew larger and at faster rates, indicating greater effort exerted during speech perception when signals were degraded. The same effect emerged whether spectral resolution was controlled via a conventional channel vocoder or a vocoder designed to more realistically replicate some aspects of CI processing and electrical current spread. Furthermore, listeners showed differences in pupil dilation even when analysis was restricted to trials in which all stimuli were identified correctly. That is, there were systematic differences in effort that emerged even when performance on the intelligibility task was perfect (Figs. 1 and 5).

There are a number of methodological factors that affect pupil dilation in listening tasks. Listeners in the present study were prompted to repeat sentence stimuli more quickly than that done by the listeners in the studies by Zekveld et al. (2010), who observed relatively smaller pupil dilation. Larger pupillary responses might have been observed in the present study because

TABLE 2. Generalized linear mixed-effects model formula and summary output for growth curve analysis in Experiment 2

Formula:

Pupil dilation = (Intercept + Condition) + # *intercept*
time1*($\beta_{time}1 + \beta_{time}1$:Condition) + # *slope*
time2*($\beta_{time}2 + \beta_{time}2$:Condition) + # *acceleration*
r(subject-level intercept) # *random error*
r(subject-level slope and acceleration) # *random error*

Formula code: PupilDilation ~ (time1 + time2) + Filter + time1:Filter + time2:Filter + (1 | Subject) + (time1 + time2 | Subject)

Term	Estimate	SE	*t*	*p*
Intercept	0.138	0.018	7.8	<0.001‡
Time 1	0.323	0.055	5.9	<0.001‡
Time 2	–0.052	0.025	–2.09	0.037*
Filter 21dB.oct	0.073	0.004	18.31	<0.001‡
Filter 16dB.oct	0.120	0.004	29.86	<0.001‡
Filter 11dB.oct	0.135	0.004	33.49	<0.001‡
Filter 7dB.oct	0.179	0.004	45.01	<0.001‡
Time1:Filter 21dB.oct	0.103	0.024	4.25	<0.001‡
Time1:Filter 16dB.oct	0.129	0.024	5.29	<0.001‡
Time1:Filter 11dB.oct	0.248	0.025	10.1	<0.001‡
Time1:Filter 7dB.oct	0.499	0.024	20.6	<0.001‡
Time2:Filter 21dB.oct	–0.035	0.024	–1.46	0.143
Time2:Filter 16dB.oct	–0.111	0.024	–4.54	<0.001‡
Time2:Filter 11dB.oct	–0.067	0.025	–2.74	0.006†
Time2:Filter 7dB.oct	–0.091	0.024	–3.77	0.000‡

"Time 1" and "time2" refer to linear and quadratic time polynomials, respectively. "Intercept," "time1," and "time2" as isolated model terms refer to the default condition, which was normal unprocessed speech. Estimates of interactions of these terms with the conditions reflect the change in the estimate of the named condition when compared against the default condition.

$*p < 0.05$.

†$p < 0.01$.

‡$p < 0.001$.

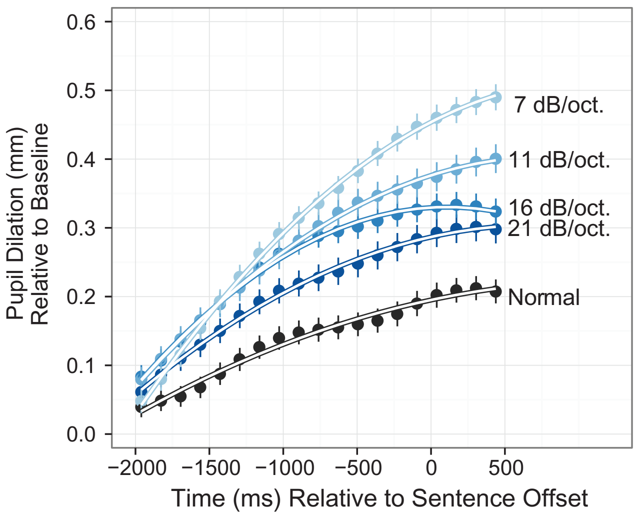

Fig. 7. Model fits (open lines) overlaid on aggregated data (points with ± SE lines) representing change in pupil diameter over time in Experiment 2. Time (milliseconds) is plotted relative to stimulus offset.

listeners were prompted to respond more quickly. Such a factor is suggested by the direct comparison of pupillary responses for unprocessed speech in quiet in the present study (where dilation reached a maximum of roughly 0.25 mm, collected after a 1.5-sec stimulus-to-prompt delay) and measurements obtained by Zekveld and Kramer (2014) for similar speech quality, which were much smaller (0.07 mm, collected during a 3-sec stimulus-to-prompt delay). Further research could determine the precise relationship between stimulus response prompt timing and the cognitive pupillary response.

To reach comparable intelligibility levels with spectrally degraded speech in quiet, listeners in the present study demonstrated relatively greater pupil dilation than that observed for noise-masked speech as measured by Zekveld et al. (2010), but less than that observed by Zekveld and Kramer (2014). Zekveld et al. (2014) measured pupil dilation in response to a prompt 4.75 sec after stimulus onset, whereas the present study measured relative to stimulus offset. Additionally, in their 2014 study, Zekveld et al. interleaved trials of different degradation types, such that vocoded sentences could be preceded and followed by sentences in quiet, sentences in noise, or silent trials. The resulting uncertainty might have added to the relatively greater pupil size observed in that study.

Beatty (1982) aggregated data from a wide variety of studies that used pupillometry as an index of cognitive load, permitting intertask comparison of listening effort. In that review, and in the present study, pupillary responses all fell within the same general range of 0 to 0.6 mm. Qualitatively different tasks such as memory, sentence disambiguation, and mental arithmetic all exhibit an orderly relationship between cognitive load and pupillary response within this restricted range of values, providing a common metric for comparison across experiments. Figure 9 displays the mean values for maximum pupil dilation in all conditions of the present study along with those of several other qualitatively different cognitive tasks and speech recognition tasks from prior literature. Despite technological advances in eye-tracking equipment as well as methodological differences and differences in scientific goals, the range of values across these studies is notably constrained. Juxtaposition of these data permits some rough estimations of the exchange rate between spectral resolution and processing load in other cognitive domains. For example, the maximum pupillary response when listening to sentences through an 8-channel vocoder was similar to that while memorizing a string of six digits or experiencing low-intensity pain. The pupillary response to the 7-dB/octave CI-style vocoder is similar to that from doing hard mental multiplication (e.g., 14 × 23), while listening to the easier 21-dB/octave CI-style vocoder is more like performing easy multiplication (e.g., 5 × 6). Although these comparisons are peculiar in nature, they are potentially useful for recognizing that effort can be understood as a common outcome that arises from many different demands; for the purposes of explaining the effects of signal degradation to patients, translation into other more familiar domains can indeed be informative.

In addition to the methodological differences highlighted above, there are multiple factors that play a role in pupil dilation, which might affect the results displayed in Figure 9. For example, differences in the pupillary response due to affective processing (Partala & Surakka 2003) could contribute to the magnitude of responses across studies that use different materials or studies that involve stress-inducing activities.

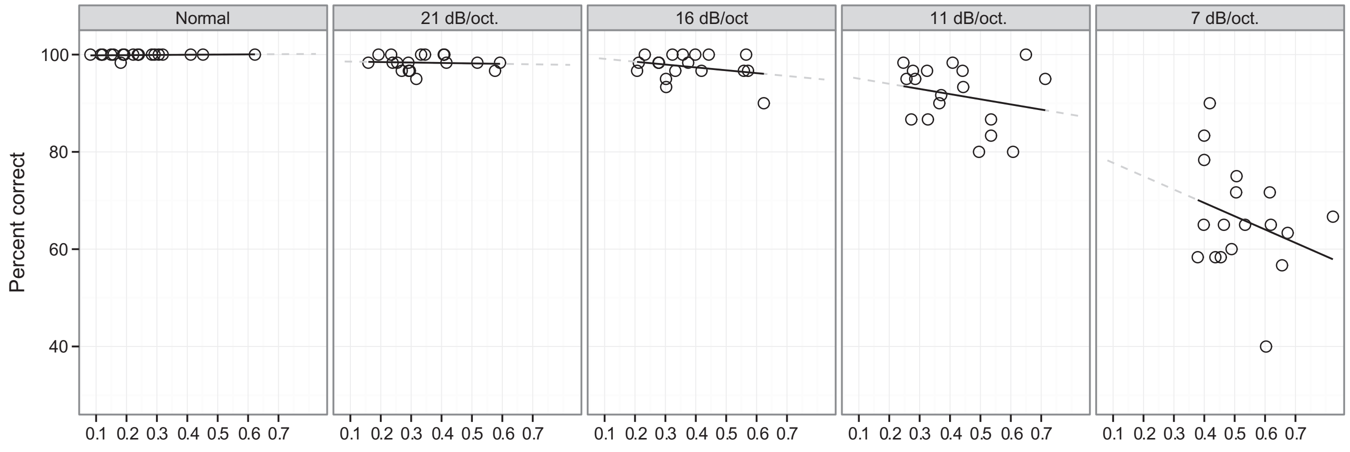

Fig. 8. Maximum pupil dilation (*x* axis) and word intelligibility (*y* axis) for each condition in Experiment 2. Regression lines are dashed when extrapolated beyond the observed data range.

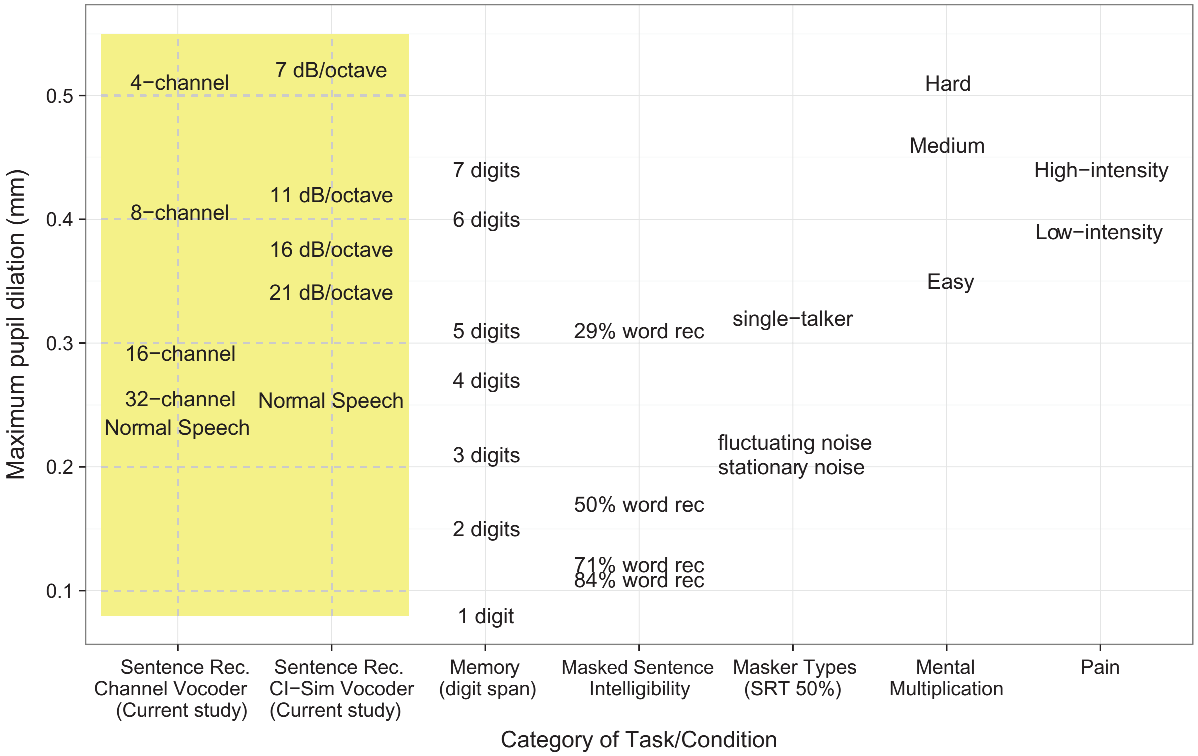

Fig. 9. Maximum pupil size elicited by various tasks. The first two columns show data from the present study. Digit span memory data aggregated by Beatty (1982) are from Ahern (1978), Kahneman and Beatty (1966), Kahneman et al. (1968), and Peavler (1974). Masked sentence intelligibility scores for young normal-hearing listeners for 50, 71, and 84% are from Zekveld et al. (2010), and 29% is from Zekveld et al. (2013). Data for different masker types to elicit 50% sentence recognition thresholds (SRT) are from Koelewijn et al. (2012). Data for mental multiplication are from Ahern and Beatty (1979). Data for pain are from Höfle et al. (2008). CI indicates cochlear implant.

Complexity of stimulus materials can also play a role in pupil size (Just & Carpenter 1993). Pupillary responses are known to become smaller with age (Winn et al. 1994), which can affect direct comparisons between subject groups. Pupillary responses change throughout the hormonal cycle, which is an effect that is modulated by the use of contraceptives (Laeng & Falkenberg 2007). Perhaps most importantly, the pupillary response to sound stimuli can be modulated by selective attention (Beatty 1982) and perceived self-efficacy (Hutchinson et al. 2008), implying that participants' engagement with the experimental control should be a noteworthy methodological consideration.

Pupillary responses are time-series data and are therefore by definition multilevel, functional data (e.g., data that take the form of a polynomial function consisting of linear, quadratic, and cubic components); each successive data point is related to samples before and after it. Although analysis techniques that identify differences in mean dilation within a specified window are clear and straightforward, they do not account for the time-varying morphology of the pupillary response and potentially violate assumptions of independence that underlie many statistical analyses. Following the approach of Kuchinsky et al. (2013), growth curve analysis in the present study addressed both of these issues and revealed that spectral degradation affected not only the mean and maximum pupil dilation but also the shape of growth of pupil dilation over time. In the end, it is likely that for this study and similar studies, a simpler analysis of mean and maximum pupil dilation would have been sufficient to identify differences across conditions. However, toward the goal of establishing the most powerful and most valid measurement technique that holds potential for identifying trends more complex than those found in this study, a move toward functional data analysis is desirable.

The relationship between spectral resolution and listening effort in this study was not captured by intelligibility scores alone. For all but the most difficult condition in each experiment, intelligibility scores were at or near 100%, minimizing any potential relationship between intelligibility and pupil dilation. Such high scores for audiology patients would likely be considered unremarkable, and the effort expended to achieve them would consequently be overlooked by traditional analyses of intelligibility. A typical solution to this "problem" would be to create more challenging listening conditions (e.g., adding masking noise) that would prevent ceiling effects. There are two potential problems in such an approach. First, the experimenter could be interested in effects that *actually* exist among the "easy" conditions for which intelligibility is not a sensitive-enough metric. Second, perception of speech in noise could rely on auditory abilities that are entirely separate from listening in quiet (e.g., auditory stream segregation, recruitment of pitch/harmonicity cues for auditory object formation, which likely play a role for speech in noise to a greater extent than for speech in quiet). Redesigning the experimental listening condition because of the limitations of the measurement tool (intelligibility) is less appealing than choosing a different measurement tool that is sensitive to differences that were previously unresolved.

During these experiments, there was both a listening and a speaking component. Pupil dilation during the spoken response in trials with correct responses were nearly indistinguishable across conditions, suggesting that spectral degradation affected the perception of stimuli and preparation of responses, but not the actual production of speech. When analyzing trials that included incorrect responses, the elevated pupil dilation observed in the most difficult conditions persisted through to the time allocated to the verbal response, unlike the pattern seen for the correct trials. Thus, it is possible that listeners drew upon cognitive resources for a longer time to form a response when the stimulus was too impoverished to be easily recognized.

Evidence for the relationship between spectral resolution and listening effort was previously indicated by Pals et al. (2013), who varied the number of channels in a noise vocoder and measured listening effort as interference in performance of a secondary task. Although listening effort decreased as the number of channels was increased from two to eight in 2-channel increments, no effect was identified for resolution better than 8 channels. In the present study, significant differences were found for two different resolutions better than 8 channels (including 12, 16, and 24 channels), suggesting that pupillometry might be a more sensitive measure of the impact of spectral resolution on listening effort than the dual-task paradigm. In the present study, even for the vocoded conditions with the best resolution in both experiments (32 channels/21-dB/octave slope), pupil size was significantly different than in the normal speech condition. Although the effect was modest, it should be considered that this was a short experiment; the long-term impact of increased difficulty could potentially accumulate and become problematic for a person with HI. Furthermore, impoverished auditory input could potentially have cumulative adverse consequences on conceptual/semantic integration as is the case for impoverished visual input and its effects on reading (Gao et al. 2011).

Pupillometry and dual-task paradigms can both lead to similar conclusions (cf. Kahneman et al. 1967), but each offer advantages and disadvantages. While pupillometric measures are a well-established physiological index of cognitive load, dual-task paradigms are considerably more likely to be adopted by clinicians who do not have access to eye-tracking equipment. Furthermore, even when eye-tracking hardware is available, it can be unwieldy and demand complicated data processing that delays interpretation. However, dual-task experiments can be performed with basic computer software and thus are more readily accessible by a greater population of scientists and researchers. It is conceivable, however, that the ability to multitask might interact with the primary measured ability (speech perception) in undesirable ways; a poor ability to multitask might give the same results as the experience of elevated listening effort. It is likely the case that the variability across individuals in terms of the ability to multitask compounds the variability in the primary auditory abilities and the susceptibility to cognitive demands. In line with this speculation, Recarte et al. (2008) suggested that measures of effort in dual-task situations reflect the total amount of effort spent for both tasks, rather than the effort involved solely in the primary task (i.e., the speech task). To support this notion, they observed that the correlation between task load and pupil size was weakened with the addition of a secondary task, presumably because the measured cognitive load was no longer driven by a single factor. Assuming that the primary task is the sole interest to the researcher, this is a disadvantage that does not affect single-task pupillometric measures.

It is worthwhile to consider the aspects of CI listening that were *not* simulated by our stimuli. Devices manufactured by the Med-El and Advanced Bionics corporations are not well simulated by our approach because they offer a different number of channels, electrode spacing, input frequency analysis range, and processing strategies. However, the channel-filtering vocoder in Experiment 2 is akin to the types of current focusing approaches used by Srinivasan (2010) and Landsberger et al. (2012). There was no attempt in this study to simulate pulsatile stimulation, staggered channel selection, and the nearly ubiquitous upward shifting of all frequency energy (cf. Başkent & Shannon 2005). We also did not simulate properties of electrical loudness growth (Chatterjee 1999), limited dynamic range (Stafford et al. 2013), place-frequency nonmonotonicity (Donaldson & Nelson 2000), inconsistent electrode–neuron interface, and atrophy of cells in the cochlea and spiral ganglion (Nadol 1997).

Having established that spectral resolution has a systematic effect on pupil dilation, the method used by the present study could be applied to situations in which different processing strategies for CIs could be compared with regard to their ability to provide better spectral resolution. More generally, this approach could be used to capture reports of benefit or detriment in situations where changes in intelligibility scores are negligible. Such situations have been reported in the CI population for situations such as the use of foveated frequency-electrode allocation (Fourakis et al. 2004), self-selected frequency-electrode allocation (Jethanamest et al. 2010; Fitzgerald et al. 2013), the use of current focusing (Mens & Berenstein 2005; Bierer 2007) and, for some patients, the use of a second CI (Litovsky et al. 2006; Summerfield et al. 2006; Wackym et al. 2007). Most importantly, the anecdotal reports of clinicians and their patients can now be better served by experimental approaches that are aimed at revealing what traditional tests sometimes fail to capture.

ACKNOWLEDGMENTS

Portions of this article were presented at the 162nd meeting of the Acoustical Society of America, San Francisco, California, December 6, 2013, and at the 37th Mid-Winter meeting of the Association for Research in otolaryngology, February 24, 2014.

The work was supported by grants from the National Institutes of Health–National Institute on Deafness and Other Communication Disorders (NIH-NICHD) (R01 DC003083, Litovsky; R01 DC02932, Edwards), by a core grant to the Waisman Center from the NIH-NICHD (P30 HD03352), and the University of Wisconsin-Madison Department of Surgery.

The authors declare no other conflict of interest.

Address for correspondence: Matthew B. Winn, Waisman Center, 1500 Highland Avenue, University of Wisconsin-Madison, Madison, WI 53705, USA. E-mail: mwinn83@gmail.com

Received March 12, 2014; accepted December 18, 2014.

REFERENCES

Abbas, P. J., Hughes, M. L., Brown, C. J., et al. (2004). Channel interaction in cochlear implant users evaluated using the electrically evoked compound action potential. *Audiol Neurootol*, *9*, 203–213.

Ahern, S. (1978). *Activation and intelligence: Pupillometric correlates of individual differences in cognitive abilities.* Unpublished Doctoral Dissertation. University of California, Los Angeles, CA.

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science*, *205*, 1289–1292.

Anderson, E., Oxenham, A., Kreft, H., et al. (2011). Comparing spectral tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *J Acoust Soc Am*, *133*, 364–375.

Barr, D. (2008) Analyzing "visual world" eyetracking data using multilevel logistic regression. *J Mem Lang*, *59*, 457–474.
Başkent, D., & Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-place mapping. *J Acoust Soc Am*, *117*(3 Pt 1), 1405–1416.
Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol Bull*, *91*, 276–292.
Bierer, J. A. (2007). Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. *J Acoust Soc Am*, *121*, 1642–1653.
Bierer, J. A., Bierer, S. M., Middlebrooks, J. C. (2010). Partial tripolar cochlear implant stimulation: Spread of excitation and forward masking in the inferior colliculus. *Hear Res*, *270*, 134–142.
Bingabr, M., Espinoza-Varas, B., Loizou, P. C. (2008). Simulating the effect of spread of excitation in cochlear implants. *Hear Res*, *241*, 73–79.
Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.56. Retrieved September 15, 2013, from http://www.praat.org/.
Boëx, C., de Balthasar, C., Kós, M. I., et al. (2003). Electrical field interactions in different cochlear implant systems. *J Acoust Soc Am*, *114*(4 Pt 1), 2049–2057.
Chapman, C. R., Oka, S., Bradshaw, D. H., et al. (1999). Phasic pupil dilation response to noxious stimulation in normal volunteers: Relationship to brain evoked potentials and pain report. *Psychophysiology*, *36*, 44–52.
Chatterjee, M. (1999). Effects of stimulation mode on threshold and loudness growth in multielectrode cochlear implants. *J Acoust Soc Am*, *105*(2 Pt 1), 850–860.
Chatterjee, M., & Shannon, R. (1998). Forward masked excitation patterns in multielectrode cochlear implants. *J Acoust Soc Am*, *103*, 2565–2572.
Cohen, L. T., Richardson, L. M., Saunders, E., et al. (2003). Spatial spread of neural excitation in cochlear implant recipients: Comparison of improved ECAP method and psychophysical forward masking. *Hear Res*, *179*, 72–87.
Danermark, B., & Gellerstedt, L. C. (2004). Psychosocial work environment, hearing impairment and health. *Int J Audiol*, *43*, 383–389.
Demorest, M. E., & Erdman, S. A. (1987). Development of the communication profile for the hearing impaired. *J Speech Hear Disord*, *52*, 129–143.
Donaldson, G. S., & Nelson, D. A. (2000). Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies. *J Acoust Soc Am*, *107*, 1645–1658.
Dubno, J. R., Dirks, D. D., Langhofer, L. R. (1982). Evaluation of hearing-impaired listeners using a Nonsense-syllable Test. II. Syllable recognition and consonant confusion patterns. *J Speech Hear Res*, *25*, 141–148.
Edwards, B. (2007). The future of hearing aid technology. *Trends Ampl*, 11, 31–46.
Feuerstein, J. F. (1992). Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort. *Ear Hear*, *13*, 80–86.
Fitzgerald, M. B., Sagi, E., Morbiwala, T. A., et al. (2013). Feasibility of real-time selection of frequency tables in an acoustic simulation of a cochlear implant. *Ear Hear*, *34*, 763–772.
Fourakis, M. S., Hawks, J. W., Holden, L. K., et al. (2004). Effect of frequency boundary assignment on vowel recognition with the Nucleus 24 ACE speech coding strategy. *J Am Acad Audiol*, *15*, 281–299.
Friesen, L. M., Shannon, R. V., Baskent, D., et al. (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. (2013) AngelSim: Cochlear implant and hearing loss simulator [Computer program]. Version 1.07.01. Retrieved April 15, 2013, from http://www.tigerspeech.com/angelsim/angelsim_about.html.
Fu, Q.-J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *J Assoc Res Otolaryngol*, *6*, 19–27.
Gao, X., Stine-Morrow, E. A., Noh, S. R., et al. (2011). Visual noise disrupts conceptual integration in reading. *Psychon Bull Rev*, *18*, 83–88.
Glasberg, B. R., & Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am*, *79*, 1020–1033.
Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *J Acoust Soc Am*, *87*, 2592–2605.
Grimby, A., & Ringdahl, A. (2000). Does having a job improve the quality of life among post-lingually deafened Swedish adults with severe-profound hearing impairment? *Br J Audiol*, *34*, 187–195.
Henry, B. A., Turner, C. W., Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *J Acoust Soc Am*, *118*, 1111–1121.
Hillyard, S. A., Hink, R. F., Schwent, V. L., et al. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
Höfle, M., Kenntner-Mabiala, R., Pauli, P., et al. (2008). You can see pain in the eye: Pupillometry as an index of pain intensity under different luminance conditions. *Int J Psychophysiol*, *70*, 171–175.
Hutchinson, J., Sherman, T., Martinovic, N. (2008). The effect of manipulated self-efficacy on perceived and sustained effort. *J Appl Sport Psych*, *20*, 457–472.
Järvelin, M. R., Mäki-Torkko, E., Sorri, M. J., et al. (1997). Effect of hearing impairment on educational outcomes and employment up to the age of 25 years in northern Finland. *Br J Audiol*, *31*, 165–175.
Jethanamest, D., Tan, C. T., Fitzgerald, M. B., et al. (2010). A new software tool to optimize frequency table selection for cochlear implants. *Otol Neurotol*, *31*, 1242–1247.
Jones, G., Won, J. H., Drennan, W., et al. (2013). Relationship between channel interaction and spectral-ripple discrimination in cochlear implant users. *J Acoust Soc Am*, *133*, 425–433.
Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Can J Exp Psychol*, *47*, 310–339.
Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.
Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583–1585.
Kahneman, D., Beatty, J., Pollack, I. (1967). Perceptual deficit during a mental task. *Science*, *157*, 218–219.
Kahneman, D., Onuska, L., Wolman, R. E. (1968). Effects of grouping on the pupillary response in a short-term memory task. *Q J Exp Psychol*, *20*, 309–311.
Killion, M. C., Niquette, P. A., Gudmundsen, G. I., et al. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, *116*(4 Pt 1), 2395–2405.
Koelewijn, T., Zekveld, A. A., Festen, J. M., et al. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear Hear*, *33*, 291–300.
Kramer, S., Kapteyn, T., Festen, J., et al. (1997). Assessing aspect of hearing handicap by means of pupil dilation. *Audiology*, *36*, 155–164.
Kramer, S. E., Kapteyn, T. S., Houtgast, T. (2006). Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *Int J Audiol*, *45*, 503–512.
Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Jr., et al. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology*, *50*, 23–34.
Laeng, B., & Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormonal cycle. *Horm Behav*, *52*, 520–530.
Landsberger, D. M., Padilla, M., Srinivasan, A. G. (2012). Reducing current spread using current focusing in cochlear implant users. *Hear Res*, *284*, 16–24.
Litovsky, R., Parkinson, A., Arcaroli, J., et al. (2006). Simultaneous bilateral cochlear implantation in adults: A multicenter clinical study. *Ear Hear*, *27*, 714–731.
Litvak, L. M., Spahr, A. J., Saoji, A. A., et al. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *J Acoust Soc Am*, *122*, 982–991.
Mens, L. H., & Berenstein, C. K. (2005). Speech perception with mono- and quadrupolar electrode configurations: A crossover study. *Otol Neurotol*, *26*, 957–964.
Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. New York, NY: CRC Press.
Mirman, D., Dixon, J. A., Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *J Mem Lang*, *59*, 475–494.
Munson, B., Donaldson, G., Allen, S., et al. (2003). Patterns of phoneme misperceptions by individual with cochlear implants. *J Acoust Soc Am*, *113*, 925–935.
Nachtegaal, J., Kuik, D. J., Anema, J. R., et al. (2009). Hearing status, need for recovery after work, and psychosocial work characteristics: Results from an internet-based national survey on hearing. *Int J Audiol*, *48*, 684–691.
Nadol, J. B., Jr. (1997). Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation. *Otolaryngol Head Neck Surg*, *117*(3 Pt 1), 220–228.

Nelson, D., Van Tasell, D., Schroder, A., et al. (1995). Electrode ranking of "place-pitch" and speech recognition in electrical hearing. *J Acoust Soc Am*, *98*, 1987–1999.

Pals, C., Sarampalis, A., Baskent, D. (2013). Listening effort with cochlear implant simulations. *J Speech Lang Hear Res*, *56*, 1075–1084.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *Int J Hum-Comp Stud*, *59*, 185–198.

Parving, A., & Christensen, B. (1993). Training and employment in hearing-impaired subjects at 20–35 years of age. *Scand Audiol*, *22*, 133–139.

Peavler, W. (1974). Individual differences in pupil size and performance. In M. Janisse (Ed.), *Pupillary Dynamics and Behavior* (pp. 159–175). New York, NY: Plenum Press.

Picou, E. M., Ricketts, T. A., Hornsby, B. W. (2011). Visual cues and listening effort: Individual variability. *J Speech Lang Hear Res*, *54*, 1416–1430.

Polt, J. M. (1970). Effect of threat of shock on pupillary response in a problem-solving situation. *Percept Mot Skills*, *31*, 587–593.

Quentin Summerfield, A., Barton, G. R., Toner, J., et al. (2006). Self-reported benefits from successive bilateral cochlear implantation in post-lingually deafened adults: Randomised controlled trial. *Int J Audiol*, *45*(Suppl 1), S99–S107.

Rakerd, B., Seitz, P. F., Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear Hear*, *17*, 97–106.

Recarte, M. A., Pérez, E., Conchillo, A., et al. (2008). Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *Span J Psychol*, *11*, 374–385.

Saoji, A. A., Litvak, L. M., Hughes, M. L. (2009). Excitation patterns of simultaneous and sequential dual-electrode stimulation in cochlear implant recipients. *Ear Hear*, *30*, 559–567.

Sarampalis, A., Kalluri, S., Edwards, B., et al. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *J Speech Lang Hear Res*, *52*, 1230–1240.

Schluroff, M. (1983). In the eye of the beholder: Cognitive effort during sentence processing. In R. Bäuerle, Ch. Schwarze, A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Language* (pp. 302–323). New York, NY: de Gruyter.

Shannon, R., Fu, Q.-J., Galvin, J. (2004) The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Oto-Laryngologica Suppl*, *552*, 50–54

Shannon, R. V., Zeng, F. G., Kamath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science*, *270*, 303–304.

Siegle, G. J., Ichikawa, N., Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, *45*, 679–687.

Spellman, B. A., Price, C. M., Logan, J. M. (2001). How two causes are different from one: The use of (un)conditional information in Simpson's paradox. *Mem Cognit*, *29*, 193–208.

Srinivasan, A. G., Landsberger, D. M., Shannon, R. V. (2010). Current focusing sharpens local peaks of excitation in cochlear implant stimulation. *Hear Res*, *270*, 89–100.

Stafford, R., Stafford, J., Wells, J., et al. (2013). Vocoder simulations of highly focused cochlear implant stimulation with limited dynamic range and discriminable steps. *Ear Hear*, *35*, 262–270.

Stickney, G., Loizou, P., Mishra, L., et al. (2006). Effects of electrode design and configuration on channel interactions. *Hear Res*, *211*, 33–45.

Wackym, P. A., Runge-Samuelson, C. L., Firszt, J. B., et al. (2007). More challenging speech-perception tasks demonstrate binaural benefit in bilateral cochlear implant users. *Ear Hear*, *28*(2 Suppl), 80S–85S.

Winn, B., Whitaker, D., Elliott, D. B., et al. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Invest Ophthalmol Vis Sci*, *35*, 1132–1137.

Won, J. H., Drennan, W. R., Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J Assoc Res Otolaryngol*, *8*, 384–392.

Xu, L., Thompson, C. S., Pfingst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am*, *117*, 3255–3267.

Zekveld, A. A., Festen, J., Kramer, S. (2013). Task difficulty affects two measures of processing load: The pupil response during sentence processing and delayed cued recall of the sentences. *J Speech Lang Hear Res*, *56*, 1156–1165.

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., et al. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*, *101*, 76–86.

Zekveld, A. A., & 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., Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear Hear*, *31*, 480–490.

Zekveld, A. A., Kramer, S. E., 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.

Zhu, Z., Tang, Q., Zeng, F. G., et al. (2012). Cochlear-implant spatial selectivity with monopolar, bipolar and tripolar stimulation. *Hear Res*, *283*, 45–58.