

Cognitive Compensation of Speech Perception With Hearing Impairment, Cochlear Implants, and Aging: How and to What Degree Can It Be Achieved?

Trends in Hearing
2016, Vol. 20: 1–16
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2331216516670279
tia.sagepub.com


Deniz Başkent^{1,2,3}, Jeanne Clarke^{1,2,3}, Carina Pals^{1,2,3},
Michel R. Benard^{1,4}, Pranesh Bhargava^{1,2,3}, Jefta Saija^{1,2,3},
Anastasios Sarampalis^{3,5}, Anita Wagner^{1,2,3}, and
Etienne Gaudrain^{1,2,3,6}

Abstract

External degradations in incoming speech reduce understanding, and hearing impairment further compounds the problem. While cognitive mechanisms alleviate some of the difficulties, their effectiveness may change with age. In our research, reviewed here, we investigated cognitive compensation with hearing impairment, cochlear implants, and aging, via (a) phonemic restoration as a measure of top-down filling of missing speech, (b) listening effort and response times as a measure of increased cognitive processing, and (c) visual world paradigm and eye gazing as a measure of the use of context and its time course. Our results indicate that between speech degradations and their cognitive compensation, there is a fine balance that seems to vary greatly across individuals. Hearing impairment or inadequate hearing device settings may limit compensation benefits. Cochlear implants seem to allow the effective use of sentential context, but likely at the cost of delayed processing. Linguistic and lexical knowledge, which play an important role in compensation, may be successfully employed in advanced age, as some compensatory mechanisms seem to be preserved. These findings indicate that cognitive compensation in hearing impairment can be highly complicated—not always absent, but also not easily predicted by speech intelligibility tests only.

Keywords

hearing impairment, cochlear implants, speech perception, aging, cognitive compensation

Date received: 13 December 2015; revised: 30 August 2016; accepted: 31 August 2016

Introduction

Understanding speech under ideal conditions is seemingly straightforward. There is little ambiguity in the speech signal, and lexical activation is automatic, requiring minimal cognitive processing for decoding the message (e.g., Marslen-Wilson & Welsh, 1978). In real life, listening conditions are not ideal. The acoustic-phonetic cues of the speech signal are usually reduced or distorted by environmental factors, such as poor room acoustics, reverberation, and interference from background sounds and talkers (Assmann & Summerfield, 2004; Beutelmann & Brand, 2006; Bronkhorst, 2000), as well as due to speaking styles, such as casual speech and regional or foreign accents (Gordon-Salant, Yeni-

¹Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, University of Groningen, Netherlands

²Graduate School of Medical Sciences, University of Groningen, Netherlands

³Research School of Behavioural and Cognitive Neurosciences, University of Groningen, Netherlands

⁴Pento Speech and Hearing Center Zwolle, Zwolle, Netherlands

⁵Department of Psychology, University of Groningen, Netherlands

⁶Auditory Cognition and Psychoacoustics, CNRS, Lyon Neuroscience Research Center, Lyon, France

Corresponding author:

Deniz Başkent, Department of Otorhinolaryngology, University Medical Center Groningen, University of Groningen, Postbus 30.001, BB21, 9700 RB Groningen, Netherlands.
Email: d.baskent@umcg.nl



Komshian, & Fitzgibbons, 2010; Janse & Ernestues, 2011; Tamati & Pisoni, 2014). Resolving the increased ambiguity due to these factors calls for cognitive mechanisms to be engaged (e.g., attention, use of grammatical and syntactical constraints, and semantic integration from context). This disambiguation must be accomplished at a rapid pace so that the conversation can continue. As a result, these cognitive mechanisms play an important role in compensating for factors complicating, and at times reducing, daily life speech communication (Davis & Johnsruide, 2007; Evans, McGettigan, Agnew, Rosen, & Scott, 2016; Kjellberg, 2004; Mattys, Davis, Bradlow, & Scott, 2012; Stenfelt & Rönnberg, 2009; Van Engen & Peelle, 2014; Zekveld, Kramer, & Festen, 2011).

Similar to the environmental or articulation-related factors listed earlier, hearing impairment is another factor that can negatively affect speech intelligibility. This may be the consequence of missing speech cues due to reduced audibility or of distortions due to suprathreshold factors related to hearing impairment (Başkent, 2006; Ching, Dillon, & Byrne, 1997; Glasberg & Moore, 1986; Plomp, 1978). Hearing devices can also change the speech signals, for example, due to front-end processing or due to the limitations of the speech transmission to the auditory nerve, as is the case for cochlear implants (CIs; Başkent, Gaudrain, Tamati, & Wagner, 2016; Hohmann & Kollmeier, 1995; Souza, 2002; Stone & Moore, 2004). Hearing impairment is closely associated with advanced age. A further compromise may occur due to age-related changes in cognitive processes (Bishop, Lu, & Yankner, 2010; Grady, 2012; Salthouse, 1996), additionally affecting perception of speech in adverse listening conditions in hearing-impaired (HI) individuals (Divenyi & Haupt, 1997; Dubno, Dirks, & Morgan, 1984; Jerger, Jerger, & Pirozzolo, 1991).

Cognitive processes of speech perception have been of special interest to our group. The speech intelligibility tests commonly used for speech audiometry in clinical settings provide only a partial picture of an individual's speech communication skills. The intelligibility score provides a numeric value for speech perception, tested under ideal conditions of one (clearly articulated) word or sentence presented at a time, without revealing the underlying processes of the comprehension. In our research, we have employed new approaches to determine whether cognitive processes of speech comprehension for HI individuals differ from that of normal hearing (NH) individuals. For instance, we have studied whether HI listeners can benefit from top-down compensation mechanisms that facilitate speech perception by NH listeners in particular in adverse conditions—mechanisms through which listeners can fill in unheard or masked parts of the signal by relying on the context and their knowledge of the language and situation.

If these mechanisms greatly differed in HI than NH, and if a cognitive compensation is less likely to take place, this difference could be one of the factors contributing to difficulties HI listeners experience in perceiving speech in noise.

To answer these questions, new methods, in addition to the traditional intelligibility tests, are needed. A number of possibilities have been considered in the past for such tests (Gatehouse & Gordon, 1990; Gilbert, Tamati, & Pisoni, 2013; Sarampalis, Kalluri, Edwards, & Hafter, 2009; Zekveld, George, Kramer, Goverts, & Houtgast, 2007; Zekveld, Kramer, & Festen, 2010). Here, we specifically focus on three mechanisms that can potentially be used for cognitive compensation in hearing impairment. Together, they present a comprehensive picture of the differences in cognitive processes of speech perception between HI and NH (varying from use of context for resolving lexical ambiguity to time course of comprehension). Specifically, these three mechanisms are as follows: (a) top-down restoration of speech, a cognitive mechanism that helps fill in missing speech segments and hence enhance speech perception; (b) increase in cognitive processing, listening effort, to maintain or enhance speech perception performance; and (c) use of context information and its time course, to resolve ambiguity at a high pace for real-time communication. The methods to quantify these cognitive compensation mechanisms include measurements of (a) speech intelligibility via phonemic restoration paradigm, (b) response times in verbal responses or via dual-task paradigm, and (c) eye tracking via the visual world paradigm.

Top-Down Restoration of Degraded Speech

The brain receives a continuous stream of perceptual information from which it has to build a coherent representation of the real world. To achieve this, the perceived pieces of information that belong to a common object need to be segregated (from others) and grouped together (Darwin & Carlyon, 1995; Wagemans et al., 2012). Identification of perceptual objects in this manner makes perception easier and more efficient (Bregman, 1990). The tendency to form perceptual objects from perceived pieces can also enhance perception of degraded speech. As early as in the 1950s, Miller and Licklider (1950) observed that interrupted speech remains highly intelligible for a wide range of interruption rates (from very slow interruptions of 0.1 Hz to as fast as 10 kHz), despite missing a large amount of speech information. This is partially due to the acoustic and linguistic redundancy in speech signals, where speech cues are coded in multiple ways, and the rich sentential context, which provides additional information for resolving lexical ambiguity (Assmann &

Summerfield, 2004; Lippman, 1996; Wingfield, 1975). Hence, speech with missing segments can be perceptually restored using the acoustic and linguistic content of the audible speech segments, either locally or globally, using context (Bashford, Riener, & Warren, 1992; Benard, Mensink, & Başkent, 2014; Sivonen, Maess, Lattner, & Friederici, 2006; Verschuure & Brocaar, 1983). The top-down restoration can be so strong that, under specific circumstances, listeners may not even be aware of the missing part of a speech signal. Warren (1970), for the first time, demonstrated this using speech with a silent gap filled with a coughing sound. While such nonspeech filler might not contribute to speech information, it nonetheless serves to create the illusion that the speech signal continues behind the coughing sound (or noise), throughout the interruption.

Adding a filler (usually a broadband noise) in the gaps of interrupted speech can also lead to an increase in intelligibility (Figure 1). In this case, the filler noise hides the spurious cues generated by the silent gaps that can erroneously lead to exclude the correct word as potential candidate. The noise filler thus maintains a higher level of ambiguity, making it possible to select the correct word, perhaps by relying more on context cues. Hence, interrupting speech signal with silent intervals and filling these gaps with noise engage compensatory cognitive mechanisms that result in enhanced intelligibility. This intelligibility improvement, called phonemic restoration benefit, has been frequently used in our research to quantify the top-down compensation, in both presence and absence of hearing impairment.

Top-Down Restoration and Hearing Impairment

In one of our earliest studies, we measured the phonemic restoration effect with NH, mildly HI, and moderately HI individuals. Our results (Figure 2, left panel) showed

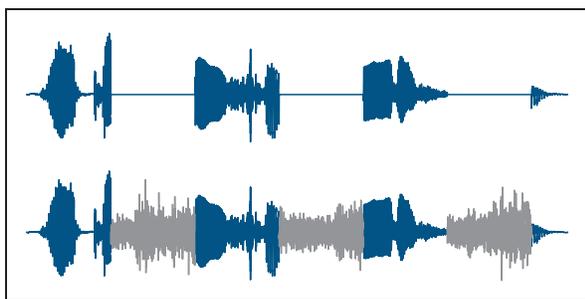


Figure 1. The waveforms are shown for the speech stimuli used in phonemic restoration experiments. The waveform in the upper part shows a sentence recording that is interrupted with periodic silent intervals. The waveform in the lower part shows the same sentence after the silent intervals are filled with filler noise bursts. The lower version is the one that induces phonemic restoration, and the difference in intelligibility as a result gives the measure of the phonemic restoration benefit.

that while mildly HI individuals could benefit from phonemic restoration, moderately HI individuals could not (Başkent, 2010; Başkent, Eiler, & Edwards, 2010). This observation implies that in mild hearing impairment (and with adequate amplification), top-down mechanisms can still be effectively used. However, as the degree of hearing impairment increases, these mechanisms seem to lose their effect. Perhaps both the loss of audibility and the distortions due to suprathreshold factors that typically accompany moderate to severe hearing loss reduce the efficacy of amplification (Başkent, 2006; Moore, 1996; Plomp, 1978). As a result, the speech cues that are necessary to trigger the compensatory cognitive mechanisms are not adequately transmitted.

Top-Down Restoration With CIs

In CIs, the speech signal is directly delivered to the auditory nerve via electric stimulation. This signal, mainly limited by the electrode–nerve interface, retains gross spectral information and the temporal envelope, while all spectrotemporal fine structure is lost (Başkent et al., 2016; Loizou, 1998; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). The relearning of the spectrotemporally degraded CI speech requires substantial adaptation following the surgery (Lazard, Innes-Brown, & Barone, 2014). While many CI users reach an acceptable level of speech perception in quiet, this is not universal, and there is a large variation in outcome measures across individual CI users (Blamey et al., 2013). Speech perception in complex environments with interfering background sounds remains a challenge (Cullington & Zeng, 2010; Eskridge, Galvin, Aronoff, Li, & Fu, 2012; Friesen, Shannon, Başkent, & Wang, 2001; Stickney, Zeng, Litovsky, & Assmann, 2004).

As CI users have to cope with degraded speech on a daily basis, top-down restoration mechanisms are especially important for this population. However, it had not been clear whether implant users could benefit from top-down restoration, given the impoverished CI-transmitted speech. A number of studies from our lab, as well as by other groups, have shown that CI users have difficulties with perception of interrupted speech (Bhargava, Gaudrain, & Başkent, 2016; Chatterjee, Peredo, Nelson, & Başkent, 2010; Nelson & Jin, 2004). Our interest was, however, in phonemic restoration, that is, if there would be an increase in intelligibility once filler noise bursts were added to the gaps of the interrupted speech. To answer this question, we first conducted a number of studies with acoustic simulations of CIs. Adding spectrotemporal degradation to speech using vocoder simulations of CIs reduced phonemic restoration benefit (Başkent, 2012). A targeted training improved the overall performance of perception of interrupted speech combined with spectrotemporal

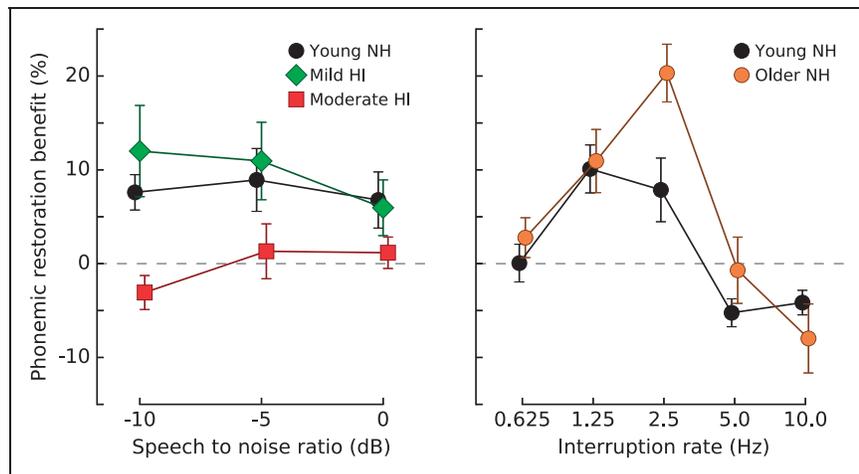


Figure 2. Phonemic restoration benefit and how it changes with hearing impairment (left panel) and aging (right panel). The left panel shows the restoration benefit for varying levels of hearing impairment, as a function of the filler noise level. The right panel shows the restoration benefit for varying ages, with no hearing impairment, as a function of interruption rate. *Note.* Source: Başkent et al. (2010) and Saija et al. (2014), respectively.

degradation, yet the phonemic restoration benefit did not show (Benard & Başkent, 2014). Adding visual lipreading cues had the same effect (Benard & Başkent, 2015). Hence, overall results with simulated CI degradations indicated that CI users would not benefit from top-down restoration.

Our data from actual CI users, however, presented a more elaborate picture than the simulation studies had suggested. In the study by Bhargava, Gaudrain, and Başkent (2014), we measured phonemic restoration by CI users for a range of parameters that varied the duty cycle of speech and the presentation level of the filler noise. When the data were averaged across the CI users, there was no phonemic restoration benefit in conditions where such benefit was observed for NH listeners. This was in line with what was expected from the previous simulation studies. However, when the data were analyzed for individual CI users, the ones with the highest baseline intelligibility scores with uninterrupted speech in quiet (i.e., good performers) showed a restoration benefit (Figure 3, left panel). Note that the causality in these data is not clear—are better CI users able to benefit more from top-down restoration mechanisms or are they good CI users because they use their top-down mechanisms better in general? Or, alternatively, are there other factors that make their devices work better in general, which then also leads to better restoration benefit in particular? Regardless, data combined from all of our studies on phonemic restoration hint at large variation in the use of cognitive mechanisms for top-down restoration within HI individuals and users of CIs, and hence at the importance of investigating individual differences, in addition to trends in group data. The individual variation likely is dominated by both the

underlying physiological limitations of hearing impairment and the cognitive resources needed for compensation (Akeroyd, 2008; Faulkner & Pisoni, 2013; Rudner & Lunner, 2014). The results, especially when compared between simulations and actual CI users, have also indicated that the use of cognitive mechanisms is very much dependent on the nature and amount of speech cues that remain in the degraded speech.

Interactive Effects of Voice Pitch and Spectral Resolution on Top-Down Restoration

In a separate study, we have focused on the importance of voice pitch, that is, fundamental frequency (F0), for phonemic restoration. This was motivated by the fact that cues related to voice pitch are severely degraded in the speech transmitted by the CI (Başkent et al., 2016; Moore & Carlyon, 2005; Zeng, Tang, & Lu, 2014), and as a result, CI users have difficulties in tasks that rely on the perception of voice (Chatterjee et al., 2015; Fuller et al., 2014; Fu, Chinchilla, Nogaki, & Galvin, 2005; Gaudrain & Başkent, 2015; Kovačić & Balaban, 2009). In general, F0 is an important cue for perceptual organization in auditory perception (Bregman, 1990; Darwin & Carlyon, 1995; Wagemans et al., 2012). For speech, also, voice cues were considered important for fusing speech segments, for example, in perception of interrupted speech (Chatterjee et al., 2010; Gnansia, Pressnitzer, Péan, Meyer, & Lorenzi, 2010; Nelson & Jin, 2004). On the other hand, in such processes, prior knowledge of language and use of context also play a role (Bashford et al., 1992; Oh, Donaldson, & Kong, 2016; Sivonen et al., 2006), which may be resources also available to CI users. Indirect support for this idea

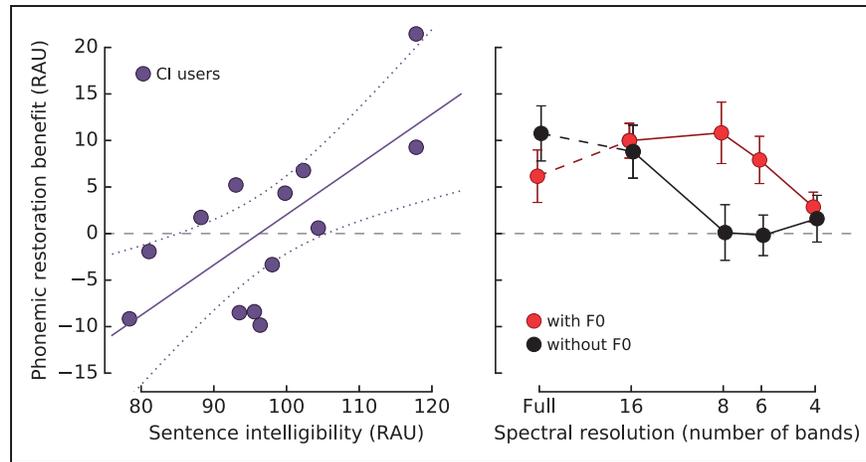


Figure 3. Phonemic restoration benefit shown for CI users (left panel) and in CI simulations with or without the voice pitch (right panel). The left panel shows the restoration benefit for individual CI users (in RAU scores), as a function of baseline sentence identification score. The right panel shows the restoration benefit averaged across NH listeners (in RAU scores) who were tested with an acoustic CI simulation, with or without the F0 cues, as a function of the number of spectral bands. *Note.* Source: Bhargava et al. (2014) and Clarke et al. (2016). CI = cochlear implant; RAU = rationalized arcsine unit; NH = normal hearing.

also came from Freyman, Griffin, and Oxenham (2012) who showed that whispered speech does not suffer from interruptions by a gated masker. Hence, while the data from CI users, as explained earlier, hinted at deficiencies in perception of interrupted speech, as well as deriving a benefit from phonemic restoration, it had not been clear how much the poorly transmitted voice cues per se may have contributed to this deficit.

Poor pitch perception with CIs likely contributes to the reduced ability to separate speech from background sounds (Cullington & Zeng, 2010; Stickney et al., 2004). The limited spectral resolution available through the implant prevents proper transmission of the harmonic structure that supports pitch perception. However, the exploration of these effects is complicated by the fact that limited spectral resolution also directly affects the representation of other speech cues, contained in the spectral envelope. To tease apart the contributions of these two components on restoration benefit, we have developed a new approach to acoustic CI simulation using TANDEM-STRAIGHT (Kawahara & Morise, 2011). This allowed us to produce speech where we could vary independently the resolution of the spectral envelope and the presence/absence of F0 (Clarke, Başkent, & Gaudrain, 2016). Removing the F0 cues by noise-exciting the spectral envelope produces speech comparable to whispered speech. When the resolution of the spectral envelope is reduced, this can be viewed as an extreme case of CI simulation where F0 cues are entirely eliminated, rather than severely degraded, as it happens in traditional methods (such as with noise-band vocoders). In contrast, preserving the original F0 cues produces natural speech when the full spectral

resolution is preserved, or, when the resolution of the spectral envelope is reduced, another extreme case of CI simulation where the F0 information is entirely preserved. The right panel of Figure 3 displays the phonemic restoration benefit with NH listeners tested using the new simulation technique. More specifically, restoration benefit was measured with varying number of spectral bands, with or without F0. The systematic comparison of the two parameters (spectral resolution, on the abscissa; presence/absence of F0 as red and black circles) shows a strong interaction. When spectral resolution is high (16 bands), where the restoration benefit is present, or low (4 bands), where the restoration benefit is absent, presence or absence of F0 does not seem to matter (overlap of red and black circles). However, in the mid ranges of spectral resolution (6 bands), the range that yields intelligibility performance most similar to actual CI users (e.g., Bhargava et al., 2016; Friesen et al., 2001), presence or absence of F0 seems to play a significant role in restoration benefit (larger benefit for red than black circles). Hence, these simulation results show that the interaction between bottom-up cues and the cognitive compensation is complex, and the exact amount and type of information provided by the hearing device can affect how much top-down restoration may occur.

Top-Down Restoration and Aging

Hearing impairment is highly correlated with aging, and as a result, many HI individuals and CI users are also of advanced age (Blamey et al., 2013; Humes, 2007; Lin, Thorpe, Gordon-Salant, & Ferrucci, 2011; Meister,

Rählmann, Walger, Margolf-Hackl, & Kiessling, 2015; Pichora-Fuller & Singh, 2006; Stevens et al., 2013). The results from the studies mentioned in the previous section showed a combined effect of hearing impairment and CI use with advanced age. To tease apart the factors of hearing status and age, we have studied the effect of age alone on phonemic restoration, without hearing impairment being involved (Saija, Akyürek, Andringa, & Başkent, 2014).

Advanced age is accompanied by changes in sensory and cognitive processes (Lindenberger, Scherer, & Baltes, 2001; Wingfield, Tun, & McCoy, 2005). On the sensory side, visual and auditory acuity declines (Caban, Lee, Gómez-Marín, Lam, & Zheng, 2005; Salthouse, 2004; Swenor, Ramulu, Willis, Friedman, & Lin, 2013). On the cognitive side, while fluid intelligence seems to be negatively affected by advanced age, crystallized intelligence may be relatively preserved or may even continue to improve with older age (Gazzaley, Cooney, Rissman, & D'Esposito, 2005; Pichora-Fuller, Schneider, & Daneman, 1995; Schieber, 2003). Perhaps as a result of these changes, even in middle-aged adults with relatively NH, lower performance was observed in perceiving speech in interfering sounds when compared with younger individuals (Başkent, van Engelshoven, & Galvin, 2014; Ruggles, Bharadwaj, & Shinn-Cunningham, 2012).

Previously, perception of interrupted speech (Figure 1, top waveform) has been shown to be negatively affected by aging (Bergman et al., 1976), an effect mostly attributed to the age-related decline in temporal processing (Gordon-Salant & Fitzgibbons, 1993). However, it was not clear whether older listeners could still effectively use the top-down restoration mechanisms, as the cognitive mechanisms, as mentioned earlier, are differentially affected by aging. If the cognitive factors such as processing speed or working memory were important for restoration, age would work against the restoration ability. If the cognitive and linguistic skills such as long-term linguistic knowledge and good use of context were important for restoration, age should not negatively affect restoration ability.

To answer this question, we measured phonemic restoration in young and older listeners, while eliminating hearing loss as a potential confounding factor by screening participants in both groups for (near) NH (Saija et al., 2014). The right panel of Figure 2 shows the restoration benefit of the two groups. The older group benefited more than the younger group from phonemic restoration at 2.5 Hz interruption rate. The middle, top panel of Figure 3 from Saija et al. (2014), shows that the intelligibility of interrupted speech by silence at 2.5 Hz for the older group was significantly lower than that of the younger group, in line with previous observations on interrupted speech perception. This interruption rate matches the syllabic frequency of

speech (2 to 5 Hz; Edwards & Chang, 2013; Verhoeven, Pauw, & Kloots, 2004), which means that mostly syllables were deleted by the interruptions. This omission likely increases ambiguity and results in less contextual and sentential cues compared with slower interruption rates, where larger parts of words are available, or with faster interruptions rates, where more glimpses per word and syllables are available (Bhargava et al., 2016). However, the insertion of noise in the silent segments activates the phonemic restoration mechanism, and so it is likely that this subsequently enables the older group to use their linguistic skills better, resulting in a larger improvement at 2.5 Hz. This reasoning is in line with Pichora-Fuller, Schneider, and Daneman (1995), who showed that older individuals benefit more from sentential context, and Benard et al. (2014), who confirmed that linguistic skills indeed seem to play an important role in the perception of interrupted speech in general. In parallel, using the gated speech paradigm—a form of interrupted speech where, instead of periodic interruptions, progressively longer segments of the same stimulus are presented—Moradi, Lidestam, Hällgren, and Rönnerberg (2014) found that when supportive semantic context is lacking, HI individuals need longer sample of the initial part of the speech signal than NH listeners to identify the rest of the speech signal. This implies greater reliance of HI individuals on linguistic cues than NH listeners. This lends support to the idea that linguistic training may improve the performance of HI listeners in challenging listening situation. If these findings can be corroborated with further studies, then this is good news for older and HI populations, as linguistic knowledge and skills can be improved with appropriate training.

Final Remarks on Top-Down Restoration

A general observation we were able to make, thanks to the phonemic restoration paradigm, is that there is a fine balance of how much and what type of speech degradation can be compensated by the use of cognitive mechanisms. For instance, very slow rates of interruption remove speech segments that are too long to be adequately restored by cognitive mechanism. On the other hand, very fast rates of interruption remove such short segments that intelligibility is hardly affected, thus leaving no room for improvement by engaging cognitive mechanisms. Similarly, too fine or too coarse spectral resolutions make the phonemic restoration benefit disappear. This indicates that, while cognitive compensation mechanisms likely play an essential role in everyday life situations, specific listening conditions or specific parameters on hearing devices may render them either inoperative or unnecessary.

However, our past work does not yet allow identifying which of the speech cues need to be preserved in order to be able to benefit from cognitive compensation. Stilp, Goupell, and Kluender (2013) used cochlea-scaled entropy as a measure of quantity of information to replace with noise bursts, the information-bearing segments in sentences. They found, in NH listeners, that replacing high cochlea-scaled entropy segments resulted in similar loss of intelligibility for full-spectrum stimuli and for vocoded sentences. This indicates that some aspects of the speech signal are robust to spectral degradations and remain important for speech intelligibility even through vocoding. From this, one may extrapolate that there could be at least some cues on which both NH and CI listeners rely heavily for top-down compensation. The next challenge thus seems to be to identify these cues and find ways to preserve and perhaps enhance them in hearing devices.

Increased Cognitive Processing: Listening Effort

The perception of degraded speech requires the allocation of additional cognitive resources, such as those related to verbal working memory and attention (Davis & Johnsrude, 2007; Kjellberg, 2004; Mattys & Wiget, 2011; Rönnberg et al., 2013; van Engen & Peelle, 2014; Wild et al., 2012; Zekveld et al., 2011), that is, an increase in listening effort. Allocating these extra cognitive resources to the task of speech comprehension is a necessary and useful mechanism for maintaining a high-level performance. However, this allocation may come at a cost as cognitive resources are considered to be limited and shared across tasks (Baddeley & Hitch, 1974; Kahneman, 1973). Hence, increasing resource allocation for speech understanding could reduce the resources available for other tasks that need to be conducted in parallel to speech comprehension, negatively affecting the performance in these additional tasks (e.g., remembering what is said; Rabbitt, 1968). Further, sustained listening effort may lead to fatigue, a common complaint by HI individuals (Hornsby, 2013; McGarrigle et al., 2014). A clinical tool to quantify listening effort, therefore, would be beneficial for optimizing hearing device settings.

In research on sensory perception, there is a long history of the use of response times to reflect cognitive effort, in an attempt to complement measures of detection or identification accuracy (Donders, 1868; Koga & Morant, 1923). In the field of communication systems, the value of such additional measures was soon recognized as well. Researchers showed that even if intelligibility for speech over degraded communication lines was high, further distinctions in sound quality could be made by using measures of listening effort, as measured in

response times (Hecker, Stevens, & Williams, 1966), memory functions (Rabbitt, 1966), or performance on a secondary task simultaneously conducted in a dual-task paradigm (Broadbent, 1958). In more recent years, attempts have been made to also explore listening effort in hearing impairment (Downs, 1982; Mackersie & Cones, 2011; Rudner et al., 2011; Sarampalis et al., 2009; Zekveld et al., 2010), using a range of methods including self-report as well as behavioral and psychophysiological measures, but mostly in research settings. These methods have not yet been incorporated into routine clinical procedures. In our lab, we aim to both unravel the complex interactions between cognitive processes and hearing impairment and also to evaluate the methods used for this purpose for potential clinical applicability.

Behavioral methods do not require additional equipment, other than the typical setup of a computer, sound card, headphones, or speakers. Therefore, initially, we started exploring hearing deficiencies and listening effort using the behavioral approach of the dual-task paradigm, as it has been a long- and well-established method in the field of cognitive psychology (Kahneman, 1973; Pashler, 1994). This paradigm is also of specific interest as it operates on the very assumption at the core of our working definition of listening effort; the interaction between two tasks that compete for limited cognitive resources. Variations of the dual-task paradigm have successfully shown that older listeners expend more listening effort than their younger counterparts (Gosselin & Gagné, 2011a, 2011b; Tun, Wingfield, & Stine, 1991), and so do HI listeners (Hicks & Tharpe, 2002; Hornsby, 2013; Rakerd, Seitz, & Whearty, 1996; Tun, McCoy, & Wingfield, 2009). This effect is further magnified when advanced age and hearing impairment are combined (Tun et al., 2009). These findings again illustrate the clinical relevance of reducing listening effort for HI listeners. Recent research has also examined the potential benefits of hearing device features, such as noise reduction, on decreasing listening effort. More often than not, hearing aid noise reduction does not provide an improvement in intelligibility (Nordrum, Erler, Garstecki, & Dhar, 2006). While not all dual-task studies examining the benefits of noise reduction show conclusive results (Neher, Grimm, Hohmann, & Kollmeier, 2014), some do suggest that noise reduction significantly reduces listening effort for speech perception in noise, in both NH (Sarampalis et al., 2009) and HI listeners (Desjardins & Doherty, 2014).

Listening effort, fatigue, and the need for appropriate clinical tools for quantifying these are highly relevant issues for CI users as well. Yet, at the time we started working on listening effort, there had been no study that explored this in CI users. One main challenge for such studies is the large variability in performance in CI users (Blamey et al., 2013). Therefore, we have taken a

systematic approach. In one of our first studies (Pals, Sarampalis, & Bařkent, 2013), we explored listening effort, quantified in a dual-task paradigm, with acoustic simulations of CIs. More specifically, we combined the primary task of intelligibility of degraded speech with a secondary visual mental rotation or word rhyming task. As these two tasks compete for limited cognitive resources, increased effort for the primary intelligibility task results in an increase in response times on the secondary task. We have manipulated the intelligibility of speech by changing the spectral resolution in the acoustic CI simulations (Friesen et al., 2001). As the number of channels increased, intelligibility, reflected by accuracy on the primary task, increased, and listening effort, reflected by response times on the secondary task, decreased (Figure 4, left and right panels, respectively). The core finding of this study was that, while intelligibility plateaued at six channels, listening effort continued to improve to eight channels. Thus, while clinical speech audiometry indicates the same speech performance for both six- and eight-channel settings, only the listening effort measure would indicate the additional benefit of further increasing the resolution from six to eight channels. Hence, the study confirmed that speech audiometry may indeed not be sufficient to show changes in listening effort for differing listening conditions/device settings. We now continue this line of research for further understanding of factors that can reduce or increase listening effort in CI users, with potential clinical applications in mind (Pals, Sarampalis, van Rijn, & Bařkent, 2015; Wagner, Toffanin, & Bařkent, 2016; Wagner, Pals, de Blecourt, Sarampalis, & Bařkent, 2016).

Use of Context and Its Time Course

Understanding speech requires the mapping of the signal onto stored mental representations. This mapping is not only affected by sensory information but also by listeners' knowledge about the context and situation. These top-down mechanisms become especially crucial in maintaining communication in adverse listening conditions, as explained earlier. To capture how quickly listeners can use this source of information, we have used eye tracking as an online measure of lexical decision making (Wagner, Pals, et al., 2016). Whereas intelligibility scores of speech audiometry reveal the end result of speech comprehension, the visual world paradigm, based on gaze fixations, can capture the time course of comprehension (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Tanenhaus, & Chambers, 2002). Knowing the time course of comprehension allows us to compare how individual stages of speech processing, as they have been identified by models of speech perception (e.g., Shortlist: Norris, 1994; Shortlist B: Norris, & McQueen, 2008; TRACE: McClelland & Elman, 1986), are affected by signal degradation. The specific stages of interest are lexical access and lexical competition. During lexical competition, listeners map the heard signal to a number of matching mental representations. This means that subconsciously they consider homonyms (e.g., 'pair' and 'pear'), words embedded in other words (e.g., 'paint' in 'painting'), and words that share sound sequences (e.g., 'two lips' and 'tulips') as potentially the matching word intended by the speaker. The measure of gaze fixations in eye tracking reflects the finding that upon

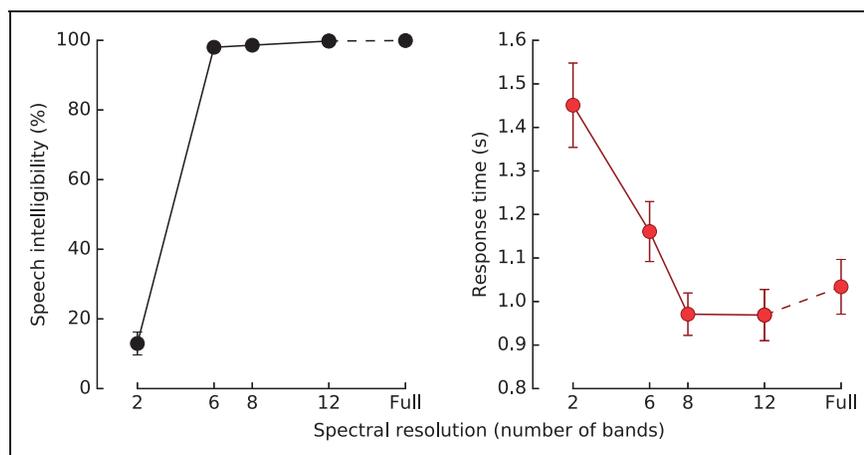


Figure 4. Speech intelligibility from primary task (left) and response time from secondary task (right), measured in a dual-task paradigm, are shown as a function of the number of spectral channels of the acoustic CI simulation. *Note.* Source: Pals et al. (2013). CI = cochlear implant.

hearing speech stimuli that refer to objects displayed on a screen, listeners fixate their gaze on the object referred to almost instantly when the acoustic input has provided enough support for that specific object (Cooper, 1974). In short, gaze fixations reflect listeners' subconscious lexical considerations even before they are aware of their lexical decision.

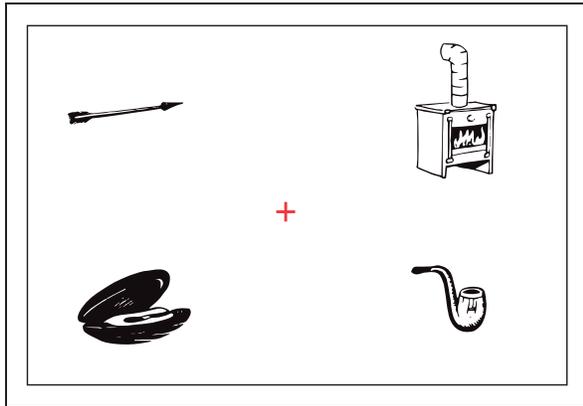


Figure 5. Visual world paradigm screenshot used by Wagner, Pals, et al. (2016) in measuring gaze fixations as a quantification of time course of speech comprehension.

As discussed before, use of linguistic context can be an effective compensation mechanism for understanding speech in adverse listening; however, the limitations of CI listening may compromise such compensation. We have used eye tracking to investigate whether degradations in CI speech would change the effectiveness of the use of sentential context. The speech signal evolves quickly, and the integration of all sources of information needs to occur timely, as delayed mapping of sound to meaning may increase the processing effort. By investigating the time course, we thus gain insight into processing bottlenecks. More specifically, we have asked the research questions if sentential context can help resolving ambiguity in word identification (lexical decision), despite the degradations of CI speech, and if so, would the time course be the same. In natural speech, listeners combine all sources of information in the sentence to predictively process later-coming information in the sentence (Dahan & Tanenhaus, 2004; Sohoglu, Peelle, Carlyon, & Davis, 2012). This means that, based on the integration of semantic information, as for instance with the thematic constraints given by the verb *crawl*, listeners can build up expectations about later-coming words. In this specific example, they would expect the words *baby* or *worm* to be more likely than words of inanimate objects.

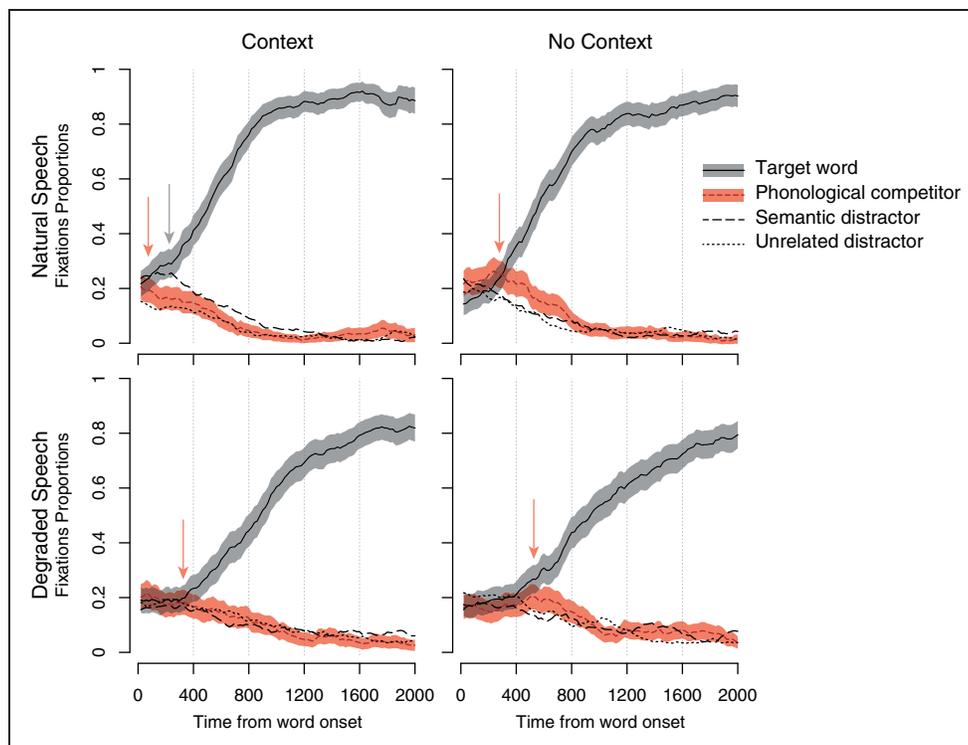


Figure 6. Gaze fixations, shown as measured by a visual world paradigm. The left and right panels show the results for high- and no-context sentences, respectively. The top and bottom panels show the results for natural and degraded speech, respectively. Note. Source: Wagner, Pals, et al. (2016).

We have measured the gaze fixations using the visual world paradigm as presented by Dahan and Tanenhaus (2004). The target word [“pijp (pipe)”] of a sentence [e.g., translated into English “At this point in the time the pipe has stopped smoking”] is presented as a picture on the screen (Figure 5), along with three other pictures, that refer to a word similar in sound [“pijl (arrow)”]; phonological competitor], a word similar in meaning [“kachel (stove)”]; semantic distractor], and a word not similar in sound nor meaning [“mossel (mussel)”]; unrelated distractor]. When no context is provided in the sentence, lexical competition would take place between the target (pijp) and the phonological competitor (pijl). When context is provided, that is, the verb *smoked* precedes the target, lexical competition between similar sounding words would be reduced and instead competition would take place between the target and the semantic distractor.

Figure 6 shows the data from gaze fixations, with high- and no-context sentences (left and right panels, respectively), and without and with acoustic CI simulation (top and bottom panels, respectively). Displayed are the proportions of fixations averaged across items and participants and the 95% confidence intervals. The most important result is the disambiguation point (marked with red arrows), where the proportion of target fixation (shown in gray in upper part of each panel) splits from phonological competitor. In natural speech, the disambiguation occurs much faster with context than with no context (comparison of the left to the right panel on top row). With degraded speech, a similar effect is observed, but the disambiguation point comes at a significantly later time (lower panels). This observation implies that context is still helpful in dissolving the ambiguity despite the degradation. However, the caveat is that the semantic distractor does not show an effect with degraded speech while it does with natural speech (top-left panel, indicated by the gray arrow). This observation implies that speech degradation reduces the efficiency of the semantic integration and also delays it considerably. This likely causes problems in real-life conversations, as they need to be carried out at a fast pace. In short, while in NH listeners, the use of semantic integration leads to a relief of resources needed for lexical access, this source of relief is not functioning properly when processing degraded speech. As a result, the degraded speech cues at the early stages of speech processing seem to affect the later stages, possibly (and negatively) affecting higher level cognitive functions. For example, the delayed processing will likely draw more on memory resources in CI users relative to NH listeners.

Conclusions

Overall, our studies suggest that there seems to be a fine balance and a complex interaction between the amount

and type of bottom-up speech cues available in the case of degraded speech and how effectively the degradation can be compensated using cognitive mechanisms. Our overall results have shown that this balance can be disturbed in hearing impairment and use of hearing devices, making this population extra vulnerable in real-life noisy listening environments.

Our results with top-down restoration show that the restoration mechanism does not seem to be as robust for HI individuals and CI users as it is for NH individuals. A mild hearing impairment does not seem to hinder restoration, but a moderate hearing impairment does. The degradations caused by vocoder simulations seem to prevent restoration almost entirely, which could not be recovered by a short targeted training or added speech information, such as visual speech cues. The results from the systematic manipulation of voice pitch and spectral resolution in a new form of CI simulation show a strong interactive effect, where only a certain combination of the conditions produce a restoration benefit. In actual CI users, only the better-performing individuals seem to benefit from restoration. All combined, these observations imply that the top-down restoration mechanism can only be achieved in a more limited manner in HI and CI listeners compared to NH listeners.

Our results on listening effort indicate indeed that high-level speech intelligibility can be maintained when speech degradation becomes more severe; however, an increase in cognitive processing as compensation can also come at a cost. Our results on the use of context support this idea. Degradations to speech, applied via a vocoder to mimic CI signal degradation, still allow the effective use of sentential context. Yet this seems to come at the cost of delayed processing, which may have negative consequences in real-life communication where the conversation needs to be conducted at a rapid pace.

The cognitive compensation abilities seem to be highly variable across individuals. This is perhaps a combined effect from the underlying etiology and other physiological limitations caused by hearing impairment and the individual differences in the cognitive resources (linguistic knowledge, vocabulary, cognitive capacity, etc.) needed for compensation.

There is also a strong effect of age, an important factor as many HI individuals tend to be older. However, this effect is not easy to predict. While in some situations, such as in perception of interrupted speech, age has a negative effect, in some others, such as in phonemic restoration, older listeners seem to be able to effectively use cognitive compensation mechanisms. The latter is a very positive finding, as we have attributed the lack of a negative age effect to lexical and linguistic knowledge that seem to be retained in advanced age, and these are entities that can potentially be improved with proper training. Hence, our results

also indicate the importance of potential training tools for improving perception of degraded speech in HI individuals.

Our results have also confirmed that the complex and interactive effects of cognitive factors in speech perception with hearing loss cannot be readily captured with the existing traditional speech tests used in the audiological practice. Measures for online speech processes and for cognitive factors may reveal more to speech comprehension and communication, especially in real-life conditions, than intelligibility scores alone. New methods (such as proposed by Pals et al., 2015; Wagner, Pals, et al., 2016; Wagner, Toffanin, Başkent 2016; but also by Winn, Edwards, & Litovsky, 2015; Zekveld et al., 2010) need to be incorporated into these practices, as well as into research and development of new hearing devices. While dual-task paradigm is proven a robust psychophysical measure of listening effort in the lab, also based on a large number of studies in cognitive psychology (see Pashler & Johnston, 1998), it can be relatively difficult to set up. The two tasks have to interact in just the right way. If one is too easy or too difficult, no or only minimal effects will be observed. Further, a dual task can be too taxing for an older HI person. In contrast, eye tracking and pupillometry are robust neurophysiological methods for quantifying cognitive mechanisms of speech perception and listening effort. While these require additional equipment, for populations where behavioral measures may be difficult to apply (such as in very young children), eye tracker still remains a good option, especially given that such technology is becoming more affordable. With such methods, device features may be better optimized and customized for individuals, by taking into account more complex mechanisms of speech perception. Similarly, manufacturers may be able to better assess new device features. There is a possibility that some device features are currently underassessed, due to lack of such measures, and are perhaps unnecessarily or prematurely discarded when they do not show a clear benefit in speech intelligibility. And finally, new rehabilitation and training programs can be developed that take into account the cognitive processes of speech, once we know which of these processes are robust or fragile in the face of hearing impairment.

Author Note

The work presented here summarizes a number of selected projects conducted at the Department of Otorhinolaryngology of University Medical Center Groningen, in collaboration with the Cochlear Implant Center Northern Netherlands (CINN) at University Medical Center Groningen (UMCG), Department of Psychology of University of Groningen (RUG), Pento

Speech and Hearing Center Zwolle, CNRS Lyon, and as part of the research program Healthy Aging and Communication of Otorhinolaryngology Department of UMCG. Portions of the results have been previously presented at a number of conferences, including Midwinter Meeting of Association of Research in Otolaryngology (ARO), Conference on Implantable Auditory Prostheses (CIAP), International Symposium on Hearing (ISH), and this article is based on a presentation given by the first author at the International Symposium on Auditory and Audiological Research (ISAAR 2015).

Acknowledgments

We would like to thank Frans Cornelissen for providing the eye-tracker; Jop Luberti for providing the pictures of the objects used in the eye-tracking study; Paolo Toffanin and Frits Leemhuis for technical and research support; Ria Woldhuis for general support; and Annemieke ter Harmsel, Nico Leenstra, Marije Sleurink, Floor Burgerhof, Maraike Coenen, Esmée van der Veen, Charlotte de Blecourt, Wydineke Boels, Margriet Bekendam, and Lisette Stuifzand for help with collecting data and transcribing participant responses. We also greatly appreciate the efforts of our participants.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work was supported by a VIDI Grant (Grant No. 016.093.397) from the Netherlands Organization for Scientific Research (NWO), the Netherlands Organization for Health Research and Development (ZonMw), a Marie Curie Intra-European Fellowship (FP7-PEOPLE-2012-IEF 332402), Rosalind Franklin Fellowship from UMCG, and funds from Heinsius Houbolt Foundation, Cochlear Ltd., Pento Speech and Hearing Center Zwolle, Doorhout Mees, Stichting Steun Gehoorgestoorde Kind as well as through the LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-0060/ANR-11-IDEX-0007) operated by the French National Research Agency.

References

- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology, 47*, S53–S71.
- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

- Assmann, P., & Summerfield, Q. (2004). The perception of speech under adverse conditions. In S. Greenberg, & W. A. Ainsworth (Eds.), *Speech processing in the auditory system* (vol. 18, pp. 231–308). New York, NY: Springer.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.) *The psychology of learning and motivation: Advances in research and theory* (vol. 8, pp. 47–89). New York, NY: Academic Press.
- Bashford, J. A., Riener, K. R., & Warren, R. M. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. *Attention, Perception & Psychophysics*, *51*, 211–217.
- Başkent, D. (2006). Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels. *The Journal of the Acoustical Society of America*, *120*, 2908–2925.
- Başkent, D. (2010). Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores. *The Journal of the Acoustical Society of America*, *128*, EL169–EL174.
- Başkent, D. (2012). Effect of speech degradation on top-down repair: Phonemic restoration with simulations of cochlear implants and combined electric–acoustic stimulation. *Journal of the Association for Research in Otolaryngology: JARO*, *13*, 683–692.
- Başkent, D., Eiler, C. L., & Edwards, B. (2010). Phonemic restoration by hearing-impaired listeners with mild to moderate sensorineural hearing loss. *Hearing Research*, *260*, 54–62.
- Başkent, D., Gaudrain, E., Tamati, T. N., & Wagner, A. (2016). Perception and psychoacoustics of speech in cochlear implant users. In A. T. Cacace, E. de Kleine, A. Genene-Holt, & P. van Dijk (Eds.), *Scientific Foundations of Audiology: Perspectives from Physics, Biology, Modeling, and Medicine* (pp. 285–319).
- Başkent, D., van Engelshoven, S., & Galvin, J. J. (2014). Susceptibility to interference by music and speech maskers in middle-aged adults. *The Journal of the Acoustical Society of America*, *135*, EL147–EL153.
- Benard, M. R., & Başkent, D. (2014). Perceptual learning of temporally interrupted and spectrally degraded speech. *The Journal of the Acoustical Society of America*, *136*, 1344–1351.
- Benard, M. R., & Başkent, D. (2015). The effect of visual cues on top-down restoration of temporally interrupted speech, with and without further degradations. *Hearing Research*, *328*, 24–33.
- Benard, M. R., Mensink, J. S., & Başkent, D. (2014). Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities. *The Journal of the Acoustical Society of America*, *135*, EL88–EL94.
- Bergman, M., Blumenfeld, V. G., Cascardo, D., Dash, B., Levitt, H., & Margulies, M. K., et al (1976). Age-related decrement in hearing for speech: Sampling and longitudinal studies. *Journal of Gerontology*, *31*, 533–538.
- Beutelmann, R., & Brand, T. (2006). Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *120*, 331–342.
- Bhargava, P., Gaudrain, E., & Başkent, D. (2014). Top-down restoration of speech in cochlear-implant users. *Hearing Research*, *309*, 113–123.
- Bhargava, P., Gaudrain, E., & Başkent, D. (2016). The intelligibility of interrupted speech: Cochlear implant users and normal hearing listeners. *Journal of the Association for Research in Otolaryngology*, *17*, 475–491.
- Bishop, N. A., Lu, T., & Yankner, B. A. (2010). Neural mechanisms of ageing and cognitive decline. *Nature*, *464*, 529–535.
- Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., . . . Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology & Neurotology*, *18*, 36–47.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: Bradford Books, MIT Press.
- Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY: Pergamon Press.
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica United with Acustica*, *86*, 117–128.
- Caban, A. J., Lee, D. J., Gómez-Marín, O., Lam, B. L., & Zheng, D. D. (2005). Prevalence of concurrent hearing and visual impairment in US adults: The National Health Interview Survey, 1997–2002. *American Journal of Public Health*, *95*(11), 1940–1942.
- Chatterjee, M., Peredo, F., Nelson, D., & Başkent, D. (2010). Recognition of interrupted sentences under conditions of spectral degradation. *The Journal of the Acoustical Society of America*, *127*, EL37–EL41.
- Chatterjee, M., Zion, D., Deroche, M. L., Burianek, B., Limb, C., Goren, A., . . . Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research*, *322*, 151–162.
- Ching, T., Dillon, H., & Byrne, D. (1997). Prediction of speech recognition from audibility and psychoacoustic abilities of hearing-impaired listeners. In W. Jesteadt (Ed.), *Modeling sensorineural hearing loss* (pp. 433–445). Mahwah, NJ: Lawrence Erlbaum.
- Clarke, J., Başkent, D., & Gaudrain, E. (2016). Pitch and spectral resolution: A systematic comparison of bottom-up cues for top-down repair of degraded speech. *The Journal of the Acoustical Society of America*, *139*, 395–405.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107.
- Cullington, H., & Zeng, F. (2010). Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. *Ear and Hearing*, *31*, 70–73.
- Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *30*, 498–513.

- Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47(2), 292–314.
- Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Hearing* (pp. 387–424). San Diego, CA: Academic Press.
- Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, 229, 132–147.
- Desjardins, J. L., & Doherty, K. A. (2014). The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear and Hearing*, 35, 600–610.
- Divenyi, P. L., & Haupt, K. M. (1997). Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and lateral asymmetry effects. *Ear and Hearing*, 18, 42–61.
- Donders, F. C. (1868). *Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868–1869*. Second issue, II, 92–120. [Translated by Koster, W. G. and reprinted: Donders, F. C. (1969)]. On the speed of mental processes. In Koster, W. G. (Ed.) *Attention and Performance II. Acta Psychologica*, 30, 412–431.
- Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disease*, 47, 189–193.
- Dubno, J., Dirks, D., & Morgan, D. (1984). Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America*, 76, 87–96.
- Edwards, E., & Chang, E. F. (2013). Syllabic (~2–5 Hz) and fluctuation (~1–10 Hz) ranges in speech and auditory processing. *Hearing Research*, 305, 113–134.
- Esckridge, E. N., Galvin, J. J. III, Aronoff, J. M., Li, T., & Fu, Q.-J. (2012). Speech perception with music maskers by cochlear implant users and normal-hearing listeners. *Journal of Speech, Language, and Hearing Research*, 55, 800–810.
- Evans, S., McGettigan, C., Agnew, Z. K., Rosen, S., & Scott, S. K. (2016). Getting the cocktail party started: Masking effects in speech perception. *Journal of Cognitive Neuroscience*, 28, 483–500.
- Faulkner, K. F., & Pisoni, D. B. (2013). Some observations about cochlear implants: Challenges and future directions. *Neuroscience Discovery*, 1(1), 9.
- Freyman, R. L., Griffin, A. M., & Oxenham, A. J. (2012). Intelligibility of whispered speech in stationary and modulated noise maskers. *The Journal of the Acoustical Society of America*, 132, 2514–2523.
- Friesen, L. M., Shannon, R. V., Başkent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, 110, 1150–1163.
- Fu, Q.-J., Chinchilla, S., Nogaki, G., & Galvin, J. J., III (2005). Voice gender identification by cochlear implant users: The role of spectral and temporal resolution. *The Journal of the Acoustical Society of America*, 118, 1711–1718.
- Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q.-J., Free, R. H., ... Başkent, D. (2014). Gender categorization is abnormal in cochlear implant users. *Journal of the Association for Research in Otolaryngology: JARO*, 15, 1037–1048.
- Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology*, 24, 63–68.
- Gaudrain, E., & Başkent, D. (2015). *Discrimination of vocal characteristics in cochlear implants*. Paper presented at the 38th Annual Mid-winter Meeting of the Association for Research in Otolaryngology, Baltimore, MD.
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, 8, 1298–1300.
- Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, reliability and validity of PRESTO: A new high-variability sentence recognition test. *Journal of the American Academy of Audiology*, 24, 26–36.
- Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, 79, 1020–1033.
- Gnansia, D., Pressnitzer, D., Péan, V., Meyer, B., & Lorenzi, C. (2010). Intelligibility of interrupted and interleaved speech for normal-hearing listeners and cochlear implantees. *Hearing Research*, 265, 46–53.
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech and Hearing Research*, 36, 1276–1285.
- Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010). Recognition of accented English in quiet by younger normal-hearing listeners and older listeners with normal-hearing and hearing loss. *The Journal of the Acoustical Society of America*, 128, 444–455.
- Gosselin, P. A., & Gagné, J.-P. (2011a). Older adults expend more listening effort than young adults recognizing audio-visual speech in noise. *International Journal of Audiology*, 50, 786–792.
- Gosselin, P. A., & Gagné, J.-P. (2011b). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research*, 54, 944–958.
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, 13, 491–505.
- Hecker, M. H., Stevens, K. N., & Williams, C. E. (1966). Measurements of reaction time in intelligibility tests. *The Journal of the Acoustical Society of America*, 39, 1188–1189.
- Hicks, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, 45, 573–584.
- Hohmann, V., & Kollmeier, B. (1995). The effect of multi-channel dynamic compression on speech intelligibility. *The Journal of the Acoustical Society of America*, 97, 1191–1195.
- Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing*, 34, 523–534.
- Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology*, 18, 590–603.

- Janse, E., & Ernestus, M. (2011). The roles of bottom-up and top-down information in the recognition of reduced speech: Evidence from listeners with normal and impaired hearing. *Journal of Phonetics*, *39*, 330–343.
- Jerger, J., Jerger, S., & Pirozzolo, F. (1991). Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly. *Ear and Hearing*, *12*, 103–109.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kawahara, H., & Morise, M. (2011). Technical foundations of TANDEM-STRAIGHT, a speech analysis, modification and synthesis framework. *Sadhana—Academy Proceedings in Engineering Sciences*, *36*, 713–722.
- Kjellberg, A. (2004). Effects of reverberation time on the cognitive load in speech communication: Theoretical considerations. *Noise and Health*, *7*, 11–21.
- Koga, Y., & Morant, G. (1923). On the degree of association between reaction times in the case of different senses. *Biometrika*, *15*, 346–372.
- Kovačić, D., & Balaban, E. (2009). Voice gender perception by cochlear implantees. *The Journal of the Acoustical Society of America*, *126*, 762–775.
- Lazard, D. S., Innes-Brown, H., & Barone, P. (2014). Adaptation of the communicative brain to post-lingual deafness: Evidence from functional imaging. *Hearing Research*, *307*, 136–143.
- Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors among older adults in the United States. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *66*, 582–590.
- Lindenberger, U., Scherer, H., & Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: Not due to sensory acuity reductions operating during cognitive assessment. *Psychology and Aging*, *16*, 196–205.
- Lippman, R. P. (1996). Accurate consonant perception without mid-frequency speech energy. *IEEE Transactions on Speech and Audio Processing*, *4*, 66–69.
- Loizou, P. C. (1998). Mimicking the human ear: An overview of signal-processing strategies for converting sound into electrical signals in cochlear implants. *IEEE Signal Processing Magazine*, *15*, 101–130.
- Mackersie, C. L., & Cones, H. (2011). Subjective and psychophysiological indices of listening effort in a competing-talker task. *Journal of the American Academy of Audiology*, *22*, 113.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*, 953–978.
- Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, *65*, 145–160.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.
- McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., ... Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British society of audiology cognition in hearing special interest group ‘white paper’. *International Journal of Audiology*, *53*, 433–440.
- Meister, H., Rähmann, S., Walger, M., Margolf-Hackl, S., & Kiessling, J. (2015). Hearing aid fitting in older persons with hearing impairment: The influence of cognitive function, age, and hearing loss on hearing aid benefit. *Clinical Interventions in Aging*, *10*, 435–443.
- Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *22*, 167–173.
- Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing*, *17*, 133–161.
- Moore, B. C. J., & Carlyon, R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper (Eds.), *Pitch: Neural coding and perception* (pp. 234–277). New York, NY: Springer.
- Moradi, S., Lidestam, B., Hällgren, M., & Rönnerberg, J. (2014). Gated auditory speech perception in elderly hearing aid users and elderly normal-hearing individuals: Effects of hearing impairment and cognitive capacity. *Trends in Hearing*, *18*, 1–12.
- Neher, T., Grimm, G., Hohmann, V., & Kollmeier, B. (2014). Do hearing loss and cognitive function modulate benefit from different binaural noise-reduction settings? *Ear and Hearing*, *35*, e52–e62.
- Nelson, P. B., & Jin, S. H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *115*, 2286–2294.
- Nordrum, S., Eler, S., Garstecki, D., & Dhar, S. (2006). Comparison of performance on the hearing in noise test using directional microphones and digital noise reduction algorithms. *American Journal of Audiology*, *15*, 81–91.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.
- Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357–395.
- Oh, S. H., Donaldson, G. S., & Kong, Y.-Y. (2016). Top-down processes in simulated electric-acoustic hearing: The effect of linguistic context on bimodal benefit for temporally interrupted speech. *Ear and Hearing*, *37*, 582–592.
- Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research*, *56*, 1075–1084.
- Pals, C., Sarampalis, A., van Rijn, H., & Başkent, D. (2015). Validation of a simple response-time measure of listening effort. *The Journal of the Acoustical Society of America*, *138*, EL187–EL192.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220–244.
- Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155–189). Hove, England: Psychology Press.
- Pichora-Fuller, K. M. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up

- and top-down information processing. *International Journal of Audiology*, 47, 72–82.
- Pichora-Fuller, M., & Singh, G. (2006). Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification*, 10, 29–59.
- Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, 63, 533–549.
- Rabbitt, P. M. (1966). Recognition: Memory for words correctly heard in noise. *Psychometric Science*, 6, 383–384.
- Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology*, 20, 241–248.
- Rakerd, B., Seitz, P., & Whearty, M. (1996). Assessing the cognitive demands of speech listening for people with hearing losses. *Ear and Hearing*, 17, 97–106.
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... Rudner, M. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, 7, 31.
- Rudner, M., & Lunner, T. (2014). Cognitive spare capacity and speech communication: A narrative overview. *BioMed Research International*, article ID 869726.
- Rudner, M., Ng, H. N., Rönnberg, N., Mishra, S., Rönnberg, J., Lunner, T., ... Stenfelt, S. (2011). Cognitive spare capacity as a measure of listening effort. *Journal of Hearing Sciences*, 1, 47–49.
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology*, 22, 1417–1422.
- Saija, J. D., Akyürek, E. G., Andringa, T. C., & Başkent, D. (2014). Perceptual restoration of degraded speech is preserved with advancing age. *Journal of the Association for Research in Otolaryngology: JARO*, 15, 139–148.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, 13, 140–144.
- Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research*, 52, 1230–1240.
- Schieber, F. (2003). Human factors and aging: Identifying and compensating for age-related deficits in sensory and cognitive function. In K. W. Schaie, & N. Charness (Eds.), *Impact of Technology on Successful Aging* (pp. 42–84). New York, NY: Springer.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303–304.
- Sivonen, P., Maess, B., Lattner, S., & Friederici, A. D. (2006). Phonemic restoration in a sentence context: Evidence from early and late ERP effects. *Brain Research*, 1121, 177–189.
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *The Journal of Neuroscience*, 32, 8443–8453.
- Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends in Amplification*, 6, 131–165.
- Stenfelt, S., & Rönnberg, J. (2009). The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology*, 50, 385–393.
- Stevens, G., Flaxman, S., Brunskill, E., Mascarenhas, M., Mathers, C. D., Finucane, M. (2013). Global and regional hearing impairment prevalence: An analysis of 42 studies in 29 countries. *The European Journal of Public Health*, 23, 146–152.
- Stickney, G. S., Zeng, F. G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *The Journal of the Acoustical Society of America*, 116, 1081–1091.
- Stilp, C. E., Goupell, M. J., & Kluender, K. R. (2013). Speech perception in simulated electric hearing exploits information-bearing acoustic change. *The Journal of the Acoustical Society of America*, 133, EL136–EL141.
- Stone, M. A., & Moore, B. C. J. (2004). Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, 116, 2311–2323.
- Swenor, B. K., Ramulu, P. Y., Willis, J. R., Friedman, D., & Lin, F. R. (2013). The prevalence of concurrent hearing and vision impairment in the United States. *JAMA Internal Medicine*, 173, 312–313.
- Tamati, T. N., & Pisoni, D. B. (2014). Non-native listeners' recognition of high-variability speech using PRESTO. *Journal of the American Academy of Audiology*, 25, 869–892.
- Tun, P. A., McCoy, S. L., & Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging*, 24, 761–766.
- Tun, P. A., Wingfield, A., & Stine, E. A. (1991). Speech-processing capacity in young and older adults: A dual-task study. *Psychology and Aging*, 6, 3–9.
- Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience*, 8, 577.
- Verhoeven, J., Pauw, G. D., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech*, 47, 297–308.
- Verschuure, J., & Brocaar, M. P. (1983). Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise. *Perception & Psychophysics*, 33, 232–240.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., von der, ... Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, 138, 1218–1252.
- Wagner, A., Pals, C., de Blecourt, C. M., Sarampalis, A., & Başkent, D. (2016). Does signal degradation affect top-down processing of speech? In P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner, & C. P. Lanting (Eds.), *Advances in experimental medicine and biology: Physiology, psychoacoustics and cognition in normal and impaired hearing* (pp 297–306). New York, NY: Springer.

- Wagner, A. E., Toffanin, P., & Başkent, D. (2016). The timing and effort of lexical access in natural and degraded speech. *Frontiers in Psychology, 7*, 398.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*, 392–393.
- Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *The Journal of Neuroscience, 32*, 14010–14021.
- Wingfield, A. (1975). Acoustic redundancy and the perception of time-compressed speech. *Journal of Speech, Language, and Hearing Research, 18*, 96–104.
- Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Current Directions in Psychological Science, 14*(3), 144–148.
- Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing, 36*, e153–e165.
- Zekveld, A., George, E., Kramer, S., Goverts, S., & Houtgast, T. (2007). The development of the text reception threshold test: A visual analogue of the speech reception threshold test. *Journal of Speech, Language, and Hearing Research, 50*, 576.
- Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and Hearing, 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 and Hearing, 32*, 498–510.
- Zeng, F.-G., Tang, Q., & Lu, T. (2014). Abnormal pitch perception produced by cochlear implant stimulation. *PLoS One, 9*, e88662.