Multisensory Integration in Cochlear Implant Recipients

Ryan A. Stevenson,[1,2] Sterling W. Sheffield,[3] Iliza M. Butera,[4]
René H. Gifford,[4–5,8] and Mark T. Wallace[4–7]

Speech perception is inherently a multisensory process involving integration of auditory and visual cues. Multisensory integration in cochlear implant (CI) recipients is a unique circumstance in that the integration occurs after auditory deprivation and the provision of hearing via the CI. Despite the clear importance of multisensory cues for perception, in general, and for speech intelligibility, specifically, the topic of multisensory perceptual benefits in CI users has only recently begun to emerge as an area of inquiry. We review the research that has been conducted on multisensory integration in CI users to date and suggest a number of areas needing further research. The overall pattern of results indicates that many CI recipients show at least some perceptual gain that can be attributable to multisensory integration. The extent of this gain, however, varies based on a number of factors, including age of implantation and specific task being assessed (e.g., stimulus detection, phoneme perception, word recognition). Although both children and adults with CIs obtain audiovisual benefits for phoneme, word, and sentence stimuli, neither group shows demonstrable gain for suprasegmental feature perception. Additionally, only early-implanted children and the highest performing adults obtain audiovisual integration benefits similar to individuals with normal hearing. Increasing age of implantation in children is associated with poorer gains resultant from audiovisual integration, suggesting a sensitive period in development for the brain networks that subserve these integrative functions, as well as length of auditory experience. This finding highlights the need for early detection of and intervention for hearing loss, not only in terms of auditory perception, but also in terms of the behavioral and perceptual benefits of audiovisual processing. Importantly, patterns of auditory, visual, and audiovisual responses suggest that underlying integrative processes may be fundamentally different between CI users and typical-hearing listeners. Future research, particularly in low-level processing tasks such as signal detection will help to further assess mechanisms of multisensory integration for individuals with hearing loss, both with and without CIs.

Key Words: Audiovisual, Cochlear implants, Integration, Multisensory, Perception, Speech

(Ear & Hearing 2017;XX;00–00)

INTRODUCTION

Over the last 2 decades, modern cochlear implant (CI) technologies have significantly improved users' auditory detection, speech perception, and quality of life (e.g., Summerfield et al 2002; Bichey & Miyamoto 2008; Bond et al. 2009; Gaylor et al. 2013). CIs are by far the most successful treatment for providing auditory perception to individuals with severe to profound hearing loss, and not surprisingly, the number of recipients worldwide has grown from over 12,000 in 1995 to latest estimates of over 324,000 (NIDCD 2014). The benefit of this success extends far beyond the simple provision of hearing, as some evidence suggests that auditory processing provides a scaffold on which the typical neurodevelopment of a wide range of cognitive processes relies (Conway et al. 2009, 2011; Kral et al. 2016).

Both CI candidacy and postoperative proficiency with a CI is primarily measured via auditory-only speech tests; however, natural speech processing is an audiovisual experience, with vision playing an integral role in shaping the intelligibility of speech signals. Thus, restricting testing for CI performance to auditory-only measures provides only a partial picture of both the benefits and limitations of these devices.

Degraded auditory input is common to all CI processors and, like hearing loss, prompts added emphasis on complementary sensory modalities. Speech processing is typically an audiovisual experience where coincident orofacial articulations can considerably boost intelligibility over unisensory auditory thresholds (Sumby & Pollack 1954; Ross et al. 2011). This is also true for typical listeners who can benefit from visual speech cues to communicate in otherwise poorly intelligible auditory signal to noise ratios (see Box 1 for an overview on audiovisual speech perception and integration in normal-hearing (NH) children and adults). Thus, when faced with impaired auditory inputs, either acoustically or electrically, the incorporation of visual cues is an effective compensatory strategy.

The body of literature describing behavioral studies of audiovisual integration by presenting audiovisual stimuli in both children and adults with CIs includes 42 published articles as of November 28, 2016. Of these 42 articles, 26 include data from adult CI recipients and 16 include data from pediatric CI recipients. Tables 1 and 2 summarize these studies in adults and children, respectively. This review is structured so as to move from the behavioral findings seen using low-level sensory processing tasks such as stimulus detection to more complex, integrative tasks such as those involving speech perception abilities and ending with a discussion of neuronal responses to multisensory integration in CI users. After these sections, we examine the trends in the findings reported in the extant literature, identify gaps in our knowledge, and highlight areas of future research.

LOW-LEVEL, NONSPEECH SENSORY PROCESSING WITH CIS

There is a paucity of studies reporting how CI users process low-level multisensory (i.e., audiovisual) stimuli. This lack of empirical work represents a large gap in the extant literature. Given that speech perception is inherently dependent on low-level sensory processing, changes at this level may have cascading impacts affecting speech perception.

Multisensory Stimulus Detection With CIs

Stimulus detection is a low-level sensory process known to benefit from multisensory information in the form of improved

[1]Department of Psychology, University of Western Ontario, London, Ontario, Canada; [2]Brain and Mind Institute, University of Western Ontario, London, Ontario, Canada; [3]Walter Reed National Military Medical Center, Audiology and Speech Pathology Center, London, Ontario, Canada; [4]Vanderbilt Brain Institute, Nashville, Tennessee; [5]Vanderbilt Kennedy Center, Nashville, Tennessee; [6]Department of Psychology, Vanderbilt University, Nashville, Tennessee; [7]Department of Psychiatry, Vanderbilt University Medical Center, Nashville, Tennessee; and [8]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, Tennessee.

<zdoi; 10.1097/AUD.0000000000000435>

The vast majority of sensory experiences one has are not limited to a single sensory modality, but instead include sensory information from multiple modalities. Typically, however, this sensory information is integrated into a single, unified perception. This process is driven primarily by two (or more) sensory inputs' temporal coincidence (Dixon & Spitz 1980; Meredith et al. 1987; Conrey & Pisoni 2004, 2006; Miller & D'Esposito 2005; van Wassenhove et al. 2007; Stevenson et al. 2010, 2011, 2012c, 2014c; Stevenson and Wallace 2013), spatial congruence (Meredith & Stein 1986a, 1996; Bertelson 1998; Stekelenburg et al. 2004), and salience (Sumby & Pollack 1954; Stein et al. 1988; Stein & Meredith 1993; Stevenson et al. 2007, 2009, 2012a; Bishop & Miller 2009; James et al. 2009, 2012b; Stevenson & James 2009; Kim & James 2010; James & Stevenson 2012a; Kim et al. 2012)—and the interactions of these factors (Macaluso et al. 2004; Stevenson et al. 2012b, 2014d; Krueger Fister et al. 2016; Nidiffer et al. 2016). Integration of information across the different senses conveys a host of behavioral and perceptual advantages. Improvements are exhibited in a wide range of paradigms, from low-level tasks such as stimulus detection (Hershenson 1962; Nelson et al. 1998; Diederich & Colonius 2004) to high-level tasks like speech perception (Sumby & Pollack 1954; Lovelace et al. 2003). For an overview of metrics used to quantify such multisensory facilitation, see Box 2, and for an in-depth tutorial review, see Stevenson et al. 2014a. These behavioral and perceptual improvements generally increase over the course of development, suggesting a major role for sensory experience in sculpting the final pattern of interactions (Wallace et al. 2004; Neil et al. 2006; Carriere et al. 2007; Wallace & Stein 2007; Polley et al. 2008; Lewkowicz & Ghazanfar 2009; Hillock et al. 2011; Ross et al., 2011; Hillock-Dunn & Wallace 2012; Wallace & Stevenson 2014; Baum et al. 2015a). For example, one of the most commonly used paradigms to assess audiovisual speech integration, the McGurk effect (described in detail later), shows increases in audiovisual integration throughout childhood and into adolescence (McGurk & MacDonald 1976; Massaro 1984; Massaro et al. 1986; Hockley & Polka 1994; Sekiyama & Burnham 2008; Wightman et al. 2006). Likewise, sensitivity to the drivers of multisensory integration also increases throughout childhood and into adolescence, such as sensitivity to the temporal relationship between auditory and visual information (Hillock et al. 2011; Hillock-Dunn & Wallace 2012).

The neural underpinnings of audiovisual speech perception in healthy adult populations are also well established (for reviews, see Calvert & Lewis 2004; Campbell 2008). This network, in addition to unisensory processing regions, includes the posterior superior temporal sulcus and gyrus, angular gyrus and supramarginal gyrus, and planum temporale and inferior frontal gyrus (including Wernicke's and Broca's areas, respectively) and ventral premotor cortex (Calvert et al. 2000, 2003; Callan et al. 2003, 2004; Jones & Callan 2003; Sekiyama et al. 2003; Wright et al. 2003; Miller & D'Esposito 2005; Ojanen et al. 2005; Skipper et al. 2005, 2007; Pekkola et al. 2006; Bernstein et al. 2008; Bishop & Miller 2009; Stevenson & James 2009; Stevenson et al. 2010, 2011). The neurodevelopment of these brain networks is substantially less studied. With that said, recent studies suggest that although the nodes recruited for audiovisual speech perception do not vary, the functional connectivity within this networks changes dramatically with maturation (Dick et al. 2010).

Box 1: An Overview of Multisensory Speech Perception and Integration

accuracy and speed of detection in NH individuals (Hershenson 1962; Nelson et al. 1998; Diederich & Colonius 2004). To our knowledge, there has only been one study to date investigating nonspeech audiovisual stimulus detection in CI users (Gilley et al. 2010). This study used a standard audiovisual detection paradigm in which individuals were presented an auditory target (1000 Hz tone), a visual target (flashed white disc), or both and asked to press a button as quickly as possible when a target was detected. Unsurprisingly, NH adults and children exhibited reaction times with audiovisual stimuli that were faster than those recorded with either of the unisensory stimuli. Children with CIs, however, had slower reaction times than NH children to all stimulus modalities. Moreover, multisensory facilitation differed depending on age of implantation. That is, only children who were implanted before the age of 4 years exhibited multisensory facilitation, albeit to a lesser extent than NH children.

This study indicated that individuals with CIs can, in fact, integrate low-level sensory information to generate perceptual gains. Furthermore, they highlight the developmental window within which these facilitative multisensory interactions mature and provide compelling support that in order for children with CIs to reap the benefits of audiovisual integration, early implantation may be a key requirement, though length of auditory experience may also contribute to these findings.

This and other studies discussed later in this review note the influence of age of implantation on the development of multisensory integration such that near-normal behavioral performance is possible with implantation between 2.5 and 4 years of age. This range is the typical period of expansive language acquisition that also corresponds with a peak in formation of cortical synapses (Huttenlocher & Dabholkar 1997). Experience-driven synaptic pruning is a critical component of shaping cortical circuits in early childhood and adolescence. Lacking sensory experience to shape this process may lead to broader processing deficits in the larger connectome (for review, see Kral et al. 2016). Early auditory intervention with a CI can ameliorate this issue, particularly if this takes place within the first 4 years of life. Indeed, the latency of the P1 wave—a measure of auditory synaptic maturation—is maximally plastic for approximately 3.5 years (Sharma et al. 2002). Accordingly, children implanted with CIs at or before 3.5 years develop age-appropriate P1 latencies within 6 months of experience with their device. In summary, the effect of age of implantation is a consistent theme in the literature, is related to critical periods in development, and will be touched upon throughout this review.

In addition to audiovisual detection in CI users, one recent study also investigated audio-tactile integration (Landry et al. 2013). Adult, pre- and postlingually deafened CI users were presented with auditory tones, vibrotactile stimuli, or both and asked to report how many vibrotactile stimuli had been presented. Importantly, on a subset of trials, a single vibrotactile stimulus was presented with 0–4 auditory tones that, when integrated, produces the perception of multiple, illusory vibrotactile sensations*. Although NH individuals showed a substantial illusory effect on the multiple tone trials, the CI group was less influenced by the multiple tones. Unfortunately, without a direct comparison of the number of perceived stimuli in the congruent condition (e.g., two vibrotactile stimuli with two tones), it is not possible to conclude whether CI users showed significant signs of integration (Box 2). Even so, these results suggest that hearing loss affects multisensory integration of auditory, visual, and somatosensory inputs.

Multisensory Temporal Perception With CIs

Temporal coincidence is one of the most salient cues indicating that two sensory inputs originated from the same external event and, thus, *should* be integrated (Meredith et al. 1987;van Atteveldt et al. 2007; van Wassenhove et al. 2007; Stevenson et al. 2012c; Stevenson & Wallace 2013). Given this, temporal

* It should be noted that this illusory paradigm was originally developed in the audiovisual domain by Shams et al (2000), known as the sound-induced flash illusion.

Multisensory gain—the perceptual benefit observed when coincident sensory information is experienced in multiple modalities—can be quantified through several methods (Stevenson et al. 2014a). Here, we will touch on three commonly used measures. First, one can measure the difference in performance between audiovisual and the best unisensory modality (for CI studies, the referent used is almost universally the auditory modality to the exclusion of the visual modality). Audiovisual gain calculated via this method can be positive (i.e., a benefit) or negative (i.e., a decrement) and is typically calculated based on the formula

$$(\mathrm{AV} - A)/(100 - A)$$

The benefit of this measure is that it shows the amount of additional information conferred by cues available from a second modality. Again, this metric is particularly useful in the context of CIs, as it enables a direct comparison between audiovisual performance and performance solely with information provided by the implant.

A second commonly used measure to quantify multisensory function is based on both channels of available information. In this so-called additive model, multisensory gain is measured as the difference in response to the audiovisual stimulus when compared with the sum of the auditory and visual only responses and assumes independence between the modalities. Use of the additive model (Stevenson et al. 2014a) allows audiovisual interactions to be categorized as superadditive,

$$(\mathrm{AV} > A + V),$$

additive,

$$(\mathrm{AV} = A + V),$$

or subadditive

$$(\mathrm{AV} < A + V).$$

A third measure of integration relies more on the construction of predictive models. In such models, independence of channels is no longer a constraint, and integration is said to occur only when the accuracy/detection is greater than the sum of the two individual modalities minus any interaction between the two, or

$$p(\mathrm{AV}) \neq p(A) + p(V) - [p(A) \times p(V)].$$

An example of a model for predicting audiovisual perception based on unisensory processing often used in speech perception is the fuzzy logic model of perception (FLMP; Massaro 1987a, 1987b, 2004). The FLMP weights auditory and visual cues based on how well they represent or predict the correct stimulus over alternative stimuli to define predicted performance in the absence of integration. The FLMP can be used to determine the magnitude of multisensory integration for speech perception whether it is facilitative (superadditive) or detrimental (subadditive). However, the FLMP is limited to closed-set stimuli because of its predictive nature, restricting its use in speech perception analyses.

Box 2: Defining Audiovisual Integration.

processing across sensory modalities is vital for efficient and effective multisensory integration. Indeed, it is common for clinical populations that exhibit deficits in multisensory temporal processing to concurrently exhibit deficits in integration (Hairston et al. 2005; Bebko et al. 2006; de Boer-Schellekens et al. 2013; Woynaroski et al. 2013; Stevenson et al. 2014b, 2014c, 2015; Wallace & Stevenson 2014; Baum et al. 2015a, b; Krueger Fister et al. 2016; Noel et al., 2017; Stevenson et al. 2017). Simultaneity judgment tasks are one of the most common paradigms for measuring multisensory temporal perception. In these tasks, participants are presented auditory and visual stimuli with varying stimulus onset asynchronies and asked to indicate whether the two stimuli appeared synchronously or asynchronously. Using this paradigm, one can calculate an individual's audiovisual temporal acuity (Fig. 1). Indeed, it has been established that individuals with high temporal precision tend to show stronger integration across sensory modalities, as temporal synchrony is a reliable cue to bind (Stevenson et al. 2012c). It should be noted, however, that all individuals are tolerant of some degree of temporal offset between sensory inputs, reflecting the statistics of the natural environment (in which light and sound travel at different speeds), and which has resulted in the concept of a multisensory temporal "binding" window—within which paired audiovisual stimuli have a high likelihood of being perceived as simultaneous. Indeed, this tendency for audiovisual temporal processing to be associated with strong multisensory integration has been reported in individuals with hearing loss: the narrower an individual's temporal binding window (i.e., the more precise their temporal perception), the better their performance in an audiovisual speech-in-noise task (Baskent & Bazo 2011). Furthermore, recent evidence suggests that *visual* temporal acuity is predictive of *auditory* word and sentence recognition in CI users (Jahn et al. 2017).

To date there has been a single study of audiovisual temporal perception in CI users (Hay-McCutcheon et al. 2009), highlighting the need for more research in this area. This study investigated multisensory temporal perception in four groups: middle-aged adults with and without CIs (mean age = 47 years) and older adults with and without CIs (mean age = 73 years). The onset of hearing loss was postlingual for all participants, and averages ranged from 41 years in the elderly group of CI users to 17 years in the middle-aged group. All participants were presented with audiovisual, single-word presentations (Lachs & Hernandez 1998), and stimulus onset asynchronies ranged from the auditory stimulus, leading the visual stimulus by 300 ms to the visual stimulus leading by 500 ms. Somewhat surprisingly, this study showed no difference between NH individuals and CI users in either age group, perhaps suggesting that temporal perception remained intact both with age and with hearing loss. However, an important consideration for interpreting these results is the fact that all CI participants' hearing loss began postlingually. This implies typical formative periods of early sensory experience and development in these individuals, which may account for similar temporal binding windows between groups. Given that other studies, including the stimulus detection results highlighted earlier, show that these low-level sensory processes are disproportionately dependent on early sensory experience, it is possible that this study reflects that this sample *did* have early multisensory experience (Wallace & Stein 2007; Polley et al. 2008). Furthermore, as acknowledged by the authors, this study was only a preliminary investigation and, thus, may have been underpowered with only 10–13 participants per group. Future research, including groups of pre- and postlingually implanted CI users, as well as NH participants, could more conclusively determine if these outcomes were a result of study power or age at onset of deafness.

SPEECH PERCEPTION WITH CIS

Although there is a small body of published work focusing on low-level multisensory processing in CI users, there is substantially more focusing on multisensory speech perception in CI

TABLE 1. Summary of studies of adults with cochlear implants

Authors	Year	Pre- vs Postlingually Deafened	N	Task/Stimuli	AV Benefit Calculation
Rabinowitz et al.	1992	Post	20	Word in sentence recognition	(AV−V)/(100−V)
Agelfors	1996	Post	15 CI; 15 HA	Suprasegmental: prosody, phoneme recognition and sentence recognition	AV−V and AV−A
Tyler et al.	1997	Post	19 CI	Consonant recognition	AV−V and AV−A
Van Dijk et al.	1999	Post	37	Sentence recognition and continuous discourse tracking	AV−V and AV−A
Goh et al.	2001	Post	1 CI, 25 NH	Vocoded sentence transcription	(AV−A)/(100−A)
Kaiser et al.	2003	Post (>3 years)	20 CI, 21 NH	Sentence recognition	(AV−A)/(100−A)
Hay-McCutcheon et al.	2005	Post (5 years)	34 CI half > 65 years	Sentence recognition	(AV−A)/(100−A) and (AV−V)/(100−V) and AV/(A+V)
McKay et al.	2016	Not reported	2 CI, 10 NH	Passive sentence perception	AV, resting
Moody-Antonio et al.	2005	Pre (<2 years)	8 CI	Sentence recognition	(AV−V)/(100−V)
Rouger et al.	2007	Post	97 CI, 163 NH	Disyllabic word recognition	(AV−A)/(100−A) and model of AV integration
Rouger et al.	2008	Post	33 CI, 39 NH	Phoneme recognition, McGurk effect	AV−A, AV−V
Desai et al.	2008	Post (>14 years)	8 CI, 14 NH	Phoneme recognition, McGurk effect	AV−A, AV−V
Champoux et al.	2009	Mixed	17 CI, 17 NH	Word recognition with incongruent visual stimuli	(A−AV)/(1−A)
Hay-McCutcheon et al.	2009	Post (>3 years)	25 CI, 22 NH	Temporal synchrony detection, sentence recognition	Temporal binding window
Strelnikov et al.	2009	Post	NA	Review	NA
Leybaert & LaSasso	2010	Post	NA	Review	NA
Tremblay et al.	2010	Mixed	17 CI, 12 NH	McGurk effect	% A, V, and AV responses
Schwartz	2010	Mixed	NA	Reanalysis of McGurk data in literature	Model of audiovisual integration
Landry et al.	2012	Mixed	17 CI, 7 NH	Speechreading with auditory distractors	V−AV
Landry et al.	2013	Mixed	CI 15, NH 15	Audio-tactile detection	NA
Bernstein et al.	2014	Pre	28 CI, 43 NH	Phoneme accuracy	AV training–A training
Song et al.	2015	Post	12 CI, 12 NH	Auditory word recognition, congruent and incongruent stimuli	AV−A
Strelnikov et al.	2013	Post	10 CI, 6 NH	Word recognition	AV−V
Van Hoesel	2015	Not reported	7 CI	Sentence recognition	AV−A
Sheffield et al.	2015	Post	11 CI	Consonant identification	
Stropahl et al.	2016	Post	8 CI, 24 NH	McGurk effect	% A, V, and AV responses
Dorman et al.	2016	Post	10 CI, 16 NH	Sentence recognition	AV−A

A, auditory; AV, audiovisual; CI, cochlear implant; HA, hearing aid; NH, normal hearing; NA, not applicable; V, visual

users. Here, we will explore these studies in a hierarchical manner based on linguistic content, starting with suprasegmental feature perception, then segmental or phonemic discrimination/perception, followed by word and sentence-level perception.

Suprasegmental Feature Perception

Suprasegmental features of speech extend over multiple speech sounds, syllables, words, and sometimes sentences. They are sometimes called prosodic features and communicate intent, emotion, or speech segmentation. A common example of a suprasegmental feature is the increase or decrease in pitch or intonation at the end of a phrase to indicate a question or statement, respectively. Although generally discussed in terms of their auditory features, suprasegmental features can also be communicated through purely visual or audiovisual cues (Bernstein et al. 1989; Dohen et al. 2004; Scarborough et al. 2009; Swerts & Krahmer 2005).

Agelfors (1996) was the first to examine suprasegmental feature perception in adults with CIs and also included a group with hearing aids. The suprasegmental feature perception testing included (1) number of syllables in a stimulus, (2) long versus short vowels, (3) tone or place of accent/emphasis in a word, and (4) word emphasis in a sentence. All variables were combined to create a single suprasegmental feature perception accuracy score. Results showed no significant audiovisual benefit relative to auditory-only performance for adults with CIs or hearing aids. It should be noted, however, that the CI group in

TABLE 2. Summary of studies of children with cochlear implants

Authors	Year	Children or Infants	N	Task/Stimuli	AV Benefit Calculation
Geers & Brenner	1994	Children	13 CI, 13 HA, 13 TA	Word and sentence recognition	AV–V
Tyler et al.	1997	Children	20 CI	Consonant recognition	AV–V and AV–A
Lachs et al.	2001	Children	27 CI	Sentence recognition	(AV–A)/(100–A) and (AV–V)/(100–V)
Bergeson et al.	2003	Children	80 CI, 42 early implanted (<53 months)	Word and sentence recognition	AV–A and AV–V
Barker & Tomblin	2004	Infants	8 CI	Preferential looking to match audiovisual syllables	Looking time differences from target to nontarget
Bergeson et al.	2005	Children	80 CI, 42 early implanted (<53 months), 38 late implanted	Word and sentence recognition	AV–A and AV–V
Schorr et al.	2005	Children	36 CI, 35 NH	Consonant recognition, McGurk effect	% A, V, and AV responses
Kirk et al.	2007	Children	15 CI	Sentence recognition	A–-A and AV–V
Bergeson et al.	2010	Infants	19 CI, 20 NH, 20 HA	Preferential looking to match audiovisual syllables	Looking time differences from target to nontarget
Gilley et al.	2010	Children	16 CI children, 8 implanted <4 years, 8 NH adults, 9 NH children	Stimulus detection	Reaction time, race model
Leybaert & LaSasso	2010	Children	NA	Review	NA
Holt et al.	2011	Children	19 CI, 29 NH	Sentence recognition	AV–A and AV–V
Houston et al.	2012	Children	NA	Review	NA
Huyse et al.	2013	Children	31 CI, 31 NH	Consonant recognition, McGurk effect	(AV–A)/(100–A) and (% A, V, and AV responses
Maglione et al.	2015	Children	7 CI, 6 NH	Music perception, congruent and incongruent	AV–V
Tona et al.	2015	Children	24 CI, 12 NH	McGurk effect	% A, V, and AV responses

A, auditory; AV, audiovisual; CI, cochlear implant; HA, hearing aid; NH, normal hearing; V, visual.

this study used single-channel CIs as well as signal-processing strategies that are now outdated. As a result, these findings may not generalize to CI recipients using current technology. Unfortunately, no direct comparisons to adults with NH can be made for this study, as no control group was included.

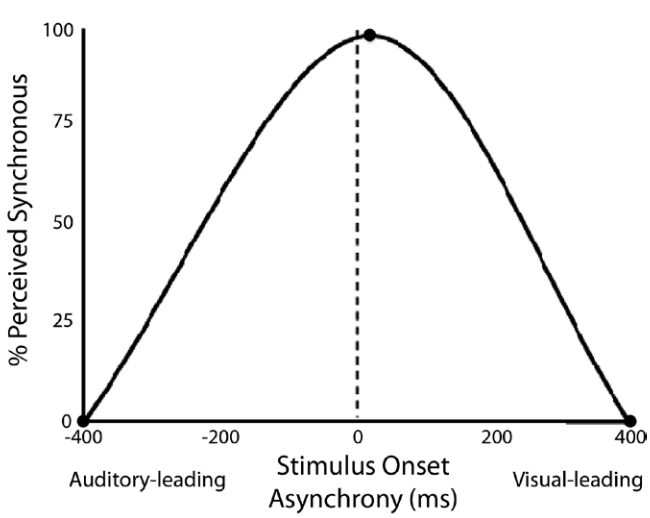

Fig. 1. Example of audiovisual temporal synchrony function. The vertical-dotted line represents presentation of the auditory and visual stimuli synchronously. Positive values on the *x* axis representing the visual stimulus presented first and negative values representing the auditory stimulus presented first.

Another aspect of suprasegmental feature perception that has been well studied from a multisensory perspective is affective or emotional content. Prosodic emotion content conveys affective content, but it is typically associated with a similarly informative facial expression. In NH listeners, the perception of emotion in voice and face is integrated, resulting in improved recognition (Busso et al. 2004; Ethofer et al. 2006; Kreifelts et al. 2007; Collignon et al. 2008; Müller et al. 2012). To date, a single study has investigated audiovisual emotion perception in CI users (Most & Aviner 2009). Unlike the previously described study by Agelfors (1996), this study examined only emotion perception as a suprasegmental feature. The participants, ranging in age from 10–17 years, were separated into four groups: NH, early-implanted CIs (<6 years), late-implanted CIs (>6 years), and those wearing hearing aids. The participants listened to a single talker producing the same sentence repeatedly with one of six emotions and were asked to identify the emotion. Results indicated that the four groups identified the correct emotion better in the visual only and audiovisual modalities than in the auditory modality. Not surprisingly, the children with NH performed significantly better in the auditory modality. Furthermore, there was no difference in performance between the groups of children with hearing loss, yet the children with NH were the only group that obtained demonstrable audiovisual benefits. These results indicate that unlike children with NH, children with hearing loss (both hearing aids and CIs) obtained no significant benefit from the information in the auditory stimulus when combined with the visual stimulus on this emotion identification task. Additionally, there was no difference between children who were early and late implanted in this study.

In summary, neither study of suprasegmental feature perception in individuals with CIs showed evidence of audiovisual benefit, a finding in stark contrast to that for individuals with NH. Limiting the analysis of multisensory interactions to the comparison between audiovisual and auditory-only performance presents a potential constraint in the interpretation of these studies. That is, the inclusion of a visual only condition would allow for additional analyses like audiovisual gain (See Box 2). Because enhanced speechreading ability is maintained after cochlear implantation (Rouger et al. 2007), visual only speech performance is likely to differ between CI users and NH controls. Future work examining suprasegmental processing using models derived from additive factors logic as well as from more predictive models such as the fuzzy logic model of perception would be particularly useful (see Box 2).

Phonemic Perception

Our search returned five studies to date that have tested audiovisual integration at the phoneme level in adults with CIs (Agelfors 1996; Desai et al. 2008; Rouger et al. 2008; Strelnikov et al. 2009; Leybaert & LaSasso 2010). Taken together, these studies have shown consistent improvement in phoneme perception (15- to 20-percentage points) under audiovisual conditions when compared with unimodal perception in quiet.

Two of these studies of phoneme perception included NH control groups, allowing the authors to directly compare audiovisual gains in NH and CI populations (Desai et al. 2008; Rouger et al. 2008). Rouger et al. (2008) showed no difference in the amount of audiovisual gain between groups, but suggest that this similarity may be the result of performance ceiling effects. Desai et al (2008) attempt to circumvent this issue by including conditions of 4- and 8-channel CI simulations with the NH group to match auditory performance between the groups. Both groups showed audiovisual benefits relative to auditory-only presentations, and the CI group showed benefits relative to visual performance. Notably, however, it is not stated whether any of these AV benefits were statistically significant. As a result, it is still an open question as to whether adults with CIs benefit to the same extent that NH listeners do from paired audiovisual speech.

Two additional studies have investigated audiovisual gain in prelingually deafened children implanted before the age of 8 years. Interestingly, these studies suggest a similar level of audiovisual benefit for phoneme perception in children as seen in adults, from 15- to 20-percentage points (Tyler et al. 1997; Huyse et al. 2013). To make comparisons of audiovisual benefit between children using CIs and NH children, Huyse et al (2013) also used degraded auditory stimuli to simulate CIs. When children were matched for age and unisensory visual performance, there was no difference in audiovisual benefit between the groups. This work also examined the effect of visual stimulus degradation on audiovisual benefit in each group and found such degradation to impact audiovisual benefit equally in both groups. This also again evidences the need to include visual-only measurements in studies of speech integration (see Box 2).

Audiovisual phoneme perception has also been examined in infants with CIs (11–24 months old) with a preferential-looking paradigm (Barker & Tomblin 2004). Infants were presented with an auditory vowel coupled with a congruent or incongruent visual vowel articulation. The infants looked toward the congruent audiovisual presentations more often than during incongruent presentations, implying a multisensory benefit, but only after 9 months of CI listening experience. These results suggest that infants' integrative abilities depend on an accumulation of CI listening experience. This is consistent with research showing that experience with co-varying stimuli across sensory modalities is important for multisensory integration (e.g., Xu et al. 2014; Altieri et al. 2015). The need for CI listening experience in infants is also consistent with speech perception data in children, indicating enhanced integration with earlier implantation (Bergeson et al. 2005, 2010; Gilley et al. 2010).

In summary, both adults and children with CIs obtain audiovisual benefit for congruent phoneme perception. Furthermore, when unisensory performance is matched or controlled for between groups, children with CIs have similar audiovisual gains to individuals with NH. This is in stark contrast to measures of audiovisual benefit in suprasegmental aspects of speech perception, where individuals with CIs did not show equivalent multisensory gains. Finally, this work has suggested that although experienced CI users do show typical audiovisual benefits, these benefits are not instantaneous, but require a length of listening CI experience before they emerge.

Although all of the aforementioned studies measured phoneme perception per se, one unique study measured the impact of multisensory perceptual learning on phoneme perception (Bernstein et al. 2014). This study compared perceptual learning using audiovisual training relative to auditory-only training, and how such training influenced phoneme perception in CI users and NH adults. Training phases consisted of learning novel pseudowords with or without the pairing of a novel object, and testing consisted of consonant recognition for audio-only pseudowords. Thus, in this experimental design, multisensory stimuli were only present in the training phase, not in the test phase. When NH adults were trained using vocoded speech to mimic CIs, audiovisual performance was as good as or better than with audio-only training. On the contrary, the gains CI users exhibited with audiovisual training were less than that seen with audio-only training, suggesting that the inclusion of visual stimuli impeded the effect of perceptual learning. Although performance was, in fact, higher within the training phase for audiovisual conditions compared to auditory-only conditions, there was no detectable benefit of multisensory training that translated to auditory-only performance. This is perhaps unsurprising considering that (1) the visual stimuli were objects and not visual speech cues, and (2) training paradigms using multiple different modalities are more likely to induce training effects in the trained modality as opposed to in a different modality. This final point makes it less surprising that auditory training improved auditory testing (same modality), but that multisensory training failed to improve auditory-only performance (different pairing of modalities).

Clinically, this study questions whether visual cues help new CI users gain proficiency in speech perception. This study seems to suggest that in certain circumstances, visual stimuli may impede proficiency with auditory only stimuli. It is noteworthy, however, that these visual stimuli were not articulations of the auditory speech signal itself but were visual representations of the novel objects that the participants were learning to name. Also in question is whether auditory-only or audiovisual proficiency is the primary goal, as audiovisual speech is more ecologically valid. These questions are not answered by the findings of this study, but both should be considered when designing future tests of CI users' speech proficiency and clinical speech rehabilitation programs.

Phoneme Perception: The McGurk Effect

A special case of multisensory phonemic perception is the McGurk Effect (McGurk & MacDonald 1976), and this illusion has been used as a powerful tool to assess audiovisual function in both typical and clinical populations (e.g., de Gelder et al. 1991; Williams et al. 2004; Mongillo et al. 2008; Pearl et al. 2009; Irwin et al. 2011; Bebko et al. 2014; Stevenson et al. 2014c; Baum et al. 2015a). The McGurk effect is a perceptual phenomenon in which incongruent visual and auditory syllables are presented, most commonly a visual "ga" presented with an auditory "ba." What the listener often perceives is neither the visual nor auditory tokens, but rather a novel token (frequently a "da"), representing a synthesis or "binding" of the two channels. Thus, when an individual perceives the illusory "da," it can be interpreted as evidence of audiovisual integration, but when an individual perceives a "ga" or "ba," it can be interpreted as a failure to integrate. It should be noted here that the neural mechanisms underlying integration of incongruent stimuli presented to induce the McGurk effect may only partially overlap with those underlying real-world, congruent speech. Evidence from event-related potentials suggest that early, more sensory-based integrative mechanisms (such as interactions in the N1) do not differ between congruent and incongruent audiovisual stimulus presentations, but that later interactions thought to reflect associative or semantic processing do differ (Stekelenburg & Vroomen 2007).

In general, participants are thought to weigh the reliability of the stimulus in each modality based on intrinsic and extrinsic factors (Schwartz 2010). This weighing of modalities can be driven by the specific task (e.g., clear versus degraded visual stimuli) or on sensory experience such as hearing impairment. Thus, with a degraded auditory input, individuals with CIs might place more weight on the visual modality.

Multiple studies have shown evidence that both adults and children with CIs perceive the McGurk effect less frequently than their NH peers, putting more weight on the visual modality than NH listeners (Schorr et al. 2005; Desai et al. 2008; Rouger et al. 2008; Tremblay et al. 2010; Huyse et al. 2013; Stropahl et al. 2015b). That is, they are more likely to perceive the incongruent presentation as a "ga," reflecting an over-reliance on the visual cue. Additionally, studies in both adults and children have found that when you degrade the auditory stimulus with vocoder or by adding background noise, individuals with NH respond more similarly to individuals with CIs (Desai et al. 2008; Huyse et al. 2013).

In contrast, a recent study (Tona et al. 2015) investigated perception of the McGurk effect for a Japanese group of 24 prelingually deafened pediatric CI users, aged 4–10 years, as well as an age-matched group of children with NH using a standard McGurk experiment presented with and without white noise at a +5 dB signal to noise ratio. In this study, children with CI were more likely to perceive the illusion in the presence of incongruent audiovisual stimulation than the children with NH. Additionally, an age effect was observed where older children with CI (≥6 years) were more likely to perceive the illusion than younger children with CI. Thus, older children with CI who have had longer durations of audiovisual experience were more likely to integrate. The authors theorized that the simplicity of the Japanese language likely influences the trend for greater proportion of McGurk perceivers in the CI group relative to other similar studies with English speakers. They also suggested that children with CIs may be more likely to develop audiovisual integration via longer-term exposure to audiovisual stimulation.

Further evidence regarding the effects of auditory stimulus clarity can be seen by separating individuals with CIs into above-average and below-average groups based on auditory-only performance for speech understanding (70–75% accuracy used as a cutoff). When split in this manner, below-average CI performers (both children and adults) report more visually biased responses (Champoux et al. 2009; Tremblay et al. 2010), whereas the above-average group reported more fused responses. Furthermore, Schorr et al. (2005) noted that children implanted earlier in life demonstrate greater audiovisual integration with fewer visually dominated responses than children implanted after 30 months. Again, this finding highlights that early and consistent exposure to audiovisual stimuli is extremely important to the development of multisensory systems.

In summary, both adults and children with CIs show multisensory gains when integrating audiovisual information in the context of incongruent phoneme perception (i.e., the McGurk effect), though generally less than their NH peers (Schorr et al. 2005; Desai et al. 2008; Rouger et al. 2008; Tremblay et al. 2010; Huyse et al. 2013; Stropahl et al. 2015b; but see Tona et al. 2015). When compared to individuals with NH, CI users place more weight on the visual modality, likely because of the degraded auditory input. This over-reliance on visual signals, however, appears to be lessened in individuals who are implanted before 30 months of age, providing converging evidence that an earlier age of implantation leads to more naturalistic audiovisual speech processing—again identifying age of implantation as one of the driving variables influencing multisensory integration in CI users.

Word Recognition and Sentence Perception

The majority of the research examining audiovisual integration in individuals with CIs has focused on word and sentence intelligibility. Studies in both adults and children with CIs have found audiovisual gain in most, if not all participants (Rabinowitz et al. 1992; Geers & Brenner 1994; Agelfors 1996; van Dijk et al. 1999; Lachs et al. 2001; Bergeson et al. 2003, 2005; Hay-McCutcheon et al. 2005; Kirk et al. 2007; Rouger et al. 2007; Strelnikov et al. 2009; Holt et al. 2011; Sheffield et al. 2015; van Hoesel 2015). These studies suggest that although the presence of a benefit is consistently seen, the *degree* of benefit appears to vary dramatically from individual-to-individual with broad categories of audiovisual integration including individuals who exhibit (1) no audiovisual increase relative to unisensory performance, (2) additive audiovisual performance, and (3) superadditive audiovisual benefit (performance is greater than the sum of the two individual modalities).

Word Recognition and Sentence Perception in Adults With CIs

Multiple studies have found evidence of *greater* audiovisual benefit in adults with CIs compared to adults with NH in word and sentence level perception. Of greatest interest here are those including audio-only, visual-only, and audiovisual conditions, as all three are required to calculate the most meaningful measures of audiovisual benefit (see Box 2) (Goh et al. 2001; Kaiser et al. 2003; Rouger et al. 2007). Figure 2 shows a comparison of the audiovisual speech perception benefit between adults with CIs and adults with NH in the three studies with these necessary conditions, all of which demonstrate a pattern of greater audiovisual benefit for CI users relative to NH controls.

There are a number of reasonable explanations for this consistent finding. First, a phenomenon known as "inverse effectiveness" is observed at low levels of sensory processing when stimuli are presented near an individual's perceptual threshold. In short, inverse effectiveness refers to findings that multisensory gain tends to *increase* as responses to unisensory stimuli *decrease* (Stevenson et al. 2014a). This finding has been robustly found in many areas of inquiry, from behavior (Sumby and Pollack 1954) to measures of neural populations (Stevenson & James 2009; Senkowski et al. 2011; James et al. 2012b; Stevenson et al. 2012a;) and even in single neuron activity (Meredith et al. 1986b; Carriere et al. 2008; Krueger et al. 2009; Royal et al. 2009). In terms of speech perception, inverse effectiveness is seen when the likelihood of unisensory perception declines, thus affording the opportunity for greater gains with coincident auditory and visual speech signals. This is directly applicable to CI users in that what is typically the most reliable signal in speech (i.e., the auditory signal) is impoverished compared to NH. Thus, inverse effectiveness predicts that decreased auditory performance increases the likelihood of multisensory benefit, as visual speech is generally already much less reliable than auditory speech.

An alternative possibility is that the increased visual lipreading abilities of CI users may lead to increased audiovisual gains. In all three studies referenced earlier, visual-only performance was greater for adults with CIs than for adults with NH. This superior visual-only speech perception performance in adults with CIs compared to adults with NH is present before and after implantation (Goh et al. 2001; Kaiser et al. 2003; Rouger et al. 2007). The better visual word and sentence recognition differs from the results for phoneme perception where either no differences were found or adults with CIs had poorer visual performance. In a review of their studies, Strelnikov et al. (2009) noted that the better visual performance for speechreading in individuals with hearing loss may be because these word and sentence stimuli contain more lexical context than phonemes. They also noted that visual performance does not change significantly after implantation and that better visual speechreading performance might drive the difference in audiovisual benefit between adults with CIs and adults with NH.

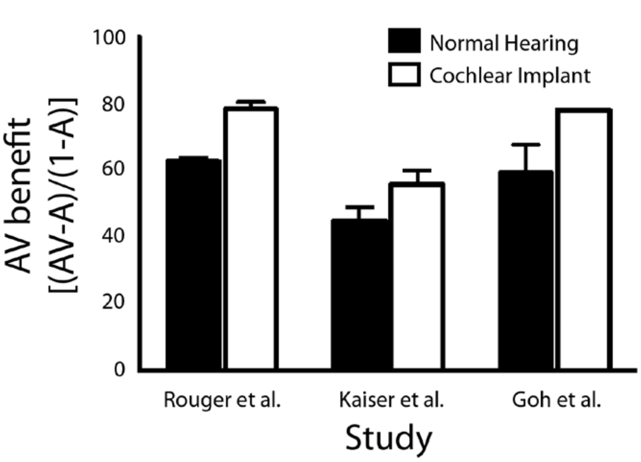

Fig. 2. Average audiovisual word and sentence recognition benefit of adults with cochlear implants (CIs) and adults with normal hearing in three separate studies on the *x* axis. Error bars represent one standard error of the mean. There is no error bar for the CI bar of the Goh et al. (2001) study because it was a case study of one adult with a CI. A, auditory; AV, audiovisual.

Regardless of whether inverse effectiveness or lipreading abilities are driving the significant increase in audiovisual gain seen in CI users, there are a number of additional factors that may influence this integrative ability. One of these is the age of the individual. Compared to younger adults, older adults with equivalent auditory-only performance exhibit poorer speechreading performance, yet still show significant gains under audiovisual conditions (Hay-McCutcheon et al. 2005). In this study, visual and auditory performance were negatively correlated in younger adults but positively correlated in older adults. A second set of factors that may affect integrative abilities includes the duration of severe to profound hearing loss—commonly referred to as duration of deafness—before implantation and age of hearing loss onset. In the same study, some of the younger adults had onsets of deafness during childhood (>6 years of age) with long durations of hearing loss. Hay-McCutcheon et al noted that these factors might have required the younger adults to improve their speechreading skills to become adequate oral communicators, unlike older adults who have shorter durations of deafness. Though older and younger adults with CIs exhibit similar audiovisual speech recognition benefit, the mechanism of audiovisual integration might differ between the groups. One factor that does not seem to affect audiovisual benefit is experience with CIs beyond the first year after activation (note that the infant study above did show experience effects under 1 year) (Barker & Tomblin 2004). When adjusting for auditory-only performance (Rouger et al. 2007), no changes were seen in audiovisual benefit between the first year and beyond, which extended to 8 years of CI experience.

In addition to age at testing, age of onset of deafness, and duration of deafness, an additional variable that may affect audiovisual abilities is whether individuals experienced severe to profound hearing loss either pre- or postlingually. All of the previously described studies of speech recognition in adults have included only postlingually deafened adults, and only one study has examined audiovisual benefit in prelingually deafened adults (Moody-Antonio et al. 2005). They found that 88% of prelingually deafened adults have audiovisual speech recognition benefits, and 38% of these individuals exhibit superadditive gains. Here again, inverse effectiveness may contribute to behavioral performance, as these prelingually deafened adults showed relatively low auditory-only (mean = 5.2%) and visual-only (mean = 25.9%) performance. It is also important to note that some of these adults do acquire audiovisual gain despite having limited to no auditory experience during development.

This research provides evidence that postlingually deafened adults with at least a year of CI experience obtain significant benefits through audiovisual word and sentence recognition. The magnitude of this benefit often eclipses that of NH listeners and is stable over time to at least 8 years after implantation. Additionally, differences in age, visual recognition, duration of deafness, and other factors may influence the magnitude of audiovisual integration. Finally, prelingually deafened adults also exhibit gains in audiovisual speech recognition despite having poor single modality performance and limited auditory experience during development.

Word Recognition and Sentence Perception in Children With CIs

Although most studies of word and sentence recognition in children using CIs report significant audiovisual benefit, there is also substantial variability (Geers & Brenner 1994; Lachs et al. 2001; Bergeson et al. 2003, 2005; Houston et al. 2012). Unlike adult studies, most children with CIs have congenital, prelingual deafness. Thus, there is an inherent difference in developmental experience between typical clinical populations of adult and child CI users. It is during this developmental period for children that the most noticeable difference between the groups can be found. In adults, after a single year of CI experience, no changes in audiovisual benefit have been found (Rouger et al. 2007). This is not the case with children. Instead, children require up to 1 year of experience with an implant before they exhibit significant audiovisual benefit in word and sentence recognition. Furthermore, the magnitude of that benefit continues to increase up to at least 5 years post implantation (Geers & Brenner 1994; Bergeson et al. 2003, 2005; Houston et al. 2012). Although these results differ drastically from what is seen with adults, it should be noted that this discrepancy may not be strictly related to CI usage, as maturation of audiovisual processing extends into late adolescence for typical listeners as well (Hillock et al. 2011; Ross et al. 2011; Hillock-Dunn & Wallace 2012).

Age of implantation is, as mentioned previously, another important aspect of CI usage, particularly with congenital deafness. On average, children that are implanted earlier show substantially better outcomes and more normative speech perception abilities, including increased audiovisual binding (Schorr et al. 2005). Bergeson et al. (2005) tested the development of audiovisual benefit longitudinally, starting preoperatively and repeated every 6 months from 1 to 3 years after implantation. Testing included the Pediatric Speech Intelligibility test, which measures both word recognition and sentence comprehension (Jerger et al. 1980). This study distinguished between children implanted before 53 months old and after 53 months old, although these groups were not matched for age at the time of evaluation. Surprisingly, this study found that children implanted *later* tended to perform better across conditions, a stark contrast to the extant literature. However, as the authors note, this surprising result was likely driven by two other factors rather than the age of implantation. First, children implanted later were also an average of 3 years older. Second, and perhaps most importantly, the children implanted later tended to have better preoperative hearing and, thus, did have some early acoustic auditory experience unlike the earlier implanted group. It should also be noted that a cutoff of 53 months (i.e., 4 years and 5 months) old was based on a median split of participants, as opposed to a more standard division (e.g., prelingual children less than 2 years of age).

One new factor that this study investigated was the developmental environment in which children were raised. Specifically, Bergeson et al dissociate between auditory/oral communication backgrounds (auditory/oral communication only) and so-called total communication backgrounds (auditory/oral communication in addition to manual communication). Children raised with auditory–oral communication generally outperformed those raised in total communication environments. Children raised in auditory/oral environments performed better in auditory-only conditions across all three years after implantation. Children in total communication environments did improve, yet at a slower rate. Furthermore, children in auditory/oral communication environments showed overall increases in visual-only and audiovisual performance, but these differences were mitigated by the second year after implantation. Speculatively, the authors suggested that this decrease in performance in the total communication group relative to the auditory/oral group may be because of the need to divide visual attention between a speaker's mouth and hands. This would, in effect, reduce the environmental exposure to congruous facial articulations with auditory speech. Further research is needed to address this issue.

Another study measuring audiovisual benefit in children with CIs was carried out by Lachs et al. (2001). This study also showed that, on average, children with CIs (aged 2–5 years) benefit from having concurrent access to both auditory and visual speech signals. Also of note, participants with high levels of auditory-only performance were most likely to show strong multisensory gain, contrary to what would be expected in terms of inverse effectiveness. Novel to this study, the intelligibility of participants' speech production was also measured. Interestingly, participants who showed both the highest audio-only performance and multisensory benefit also exhibited higher speech *intelligibility*. This finding underscores the close link between speech perception and production and demonstrates that this link carries through to CI users.

Finally, audiovisual benefit was also measured in children with CIs between the ages of 4 and 10 years old. Results again confirmed that children using CIs showed multisensory gain with audiovisual presentations relative to both auditory- and visual-only stimuli (Lachs et al. 2001). This effect was confirmed both for lexically difficult and easy words. Unfortunately, the level of benefit for easy and difficult words was not directly compared.

In summary, children with CIs benefit from audiovisual presentation of words and sentences and that benefit increases for at least 3 to 5 years after implantation. Children with CIs can perceive words and sentences better in the visual modality than children with NH. Unlike adults with CIs, however, prelingually deafened children seem to exhibit the same degree of audiovisual benefit as children with NH in the perception of words and sentences.

NEURAL RESPONSES TO MULTISENSORY STIMULI IN CI USERS

Although extensive research has been dedicated to the study of neural plasticity after sensory loss in deaf individuals (Nishimura et al. 1999; Petitto et al. 2000; MacSweeney et al. 2002; Lee et al. 2007; Petersen et al. 2013), much less has been devoted to neural plasticity of multisensory processing after cochlear implantation. This is partially because of technical limitations imposed by the magnetic and electrical components of the implant that cause significant artifacts in common noninvasive neuroimaging techniques (e.g., functional magnetic resonance imaging [fMRI], electroencephalography [EEG], and magnetoencephalography). To circumvent these limitations, several recent studies have used positron emission tomography (PET) to investigate neural responses to audiovisual speech (Barone et al. 2013; Strelnikov et al. 2013, 2015b; Song et al. 2015).

These studies primarily highlight two regions of interest; posterior superior temporal sulcus (pSTS) and inferior frontal gyrus (IFG), which includes Broca's area. Myriad studies have shown pSTS is involved in bottom-up audiovisual integration (Calvert et al. 2000, 2001; Beauchamp et al. 2004a, b; 2008; Beauchamp 2005; Miller & D'Esposito 2005; Stevenson et al. 2007, 2009, 2010, 2011; Bishop & Miller 2009; James et al. 2009; 2012a, b; 2011; Stevenson & James 2009; Werner & Noppeney 2009; Powers et al. 2012; Song et al. 2015), whereas frontal regions including IFG and Broca's area may be associated with top-down control of speech integration (Rodd et al. 2005; Zekveld et al. 2006; Davis & Johnsrude 2007; Song et al. 2015). Bilateral superior temporal sulcus (STS) has also been shown to respond to auditory-only voice stimuli in proficient, but not nonproficient CI users, highlighting its relevance here (Coez et al. 2008).

The first PET study investigating audiovisual processing in CI users presented audio-only, congruent audiovisual, and incongruent audiovisual (mismatched auditory and visual) stimuli to adults with and without CIs (Song et al. 2015). Relative to their activity in response to auditory-only presentations, NH participants showed increased activity in pSTS when presented with congruent audiovisual speech. CI participants, on the other hand, showed little increase in activation of pSTS, suggesting that the underlying neural mechanisms for audiovisual integration in CI users may differ from those in NH populations.

A different activation pattern was seen in response to *in*congruent audiovisual presentations relative to auditory-only presentations. As is typically seen, NH participants did not show as great of a response relative to the congruent presentation in pSTS. In contrast, CI users showed a greater increase in neural activity than did their matched controls in pSTS, IFG, and premotor cortex. The authors hypothesize that pSTS and Broca's area modulate cortical plasticity in deaf individuals after CI implantation based on the extent to which each region has been coopted by visual processing during deafness (Song et al. 2015).

In this same study, the researchers also investigated activity within early visual areas, with the hypothesis being that CI users may rely more heavily on visual processing. In short, CI users who showed more activity in early visual areas to either congruent or incongruent audiovisual speech also exhibited lower levels of behavioral speech comprehension, as measured outside of the scanner by a word perception test 1 year after implantation. That is, the stronger the response in visual areas, the less proficient the CI user. Two possible explanations for this are that CI users with poorer speech performance may still heavily rely on visual information or that learned overreliance on visual aspects of speech actively inhibits an individual's ability to later become a successful CI user. Further work is needed to dissociate between these two hypotheses.

A second PET study presenting audiovisual and visual word stimuli found corroborating results (Strelnikov et al. 2013). Relative to their NH peers, CI users showed an increase in top-down modulation from IFG and bottom-up integration. CI users also showed an increase in activity in visual cortices, and in CI users, the amount of activity in visual cortices was negatively correlated with speech comprehension scores, again suggesting that reliance on visual cortices in CI users may be negatively predictive of speech outcomes.

A third study using PET to investigate the neural correlates of audiovisual integration extended these findings to longitudinally map neuroplastic changes after implantation (Strelnikov et al. 2015b). In a cohort of adult CI users, experimenters scanned individuals directly after implantation and again once after they had achieved normative auditory speech recognition scores several months after implantation (Rouger et al. 2007). Each individual was presented with both visual-only speech and audiovisual speech at both time points. After even these few months of audiovisual experience, CI users showed increased neural activation to audiovisual stimuli relative to their activation observed shortly after implantation. This activation was centered in the medial temporal lobe, extending into STS and inferior parietal cortex. Additionally, this study suggested that experience with a CI led to enhanced coupling of activation between lower-level visual regions and multisensory regions, specifically in STS and the surrounding areas.

Taken together, these PET studies suggest that there is extensive neuroplasticity involved in the acquisition of audiovisual integration after cochlear implantation. Although canonical multisensory regions, such as STS, may be recruited for such processing, reliance on early visual cortices remains a distinguishing feature relative to integration in NH individuals. With that said, only two such studies have been published on the topic to date, and much work remains to be done to further clarify this picture.

A fourth neuroimaging study, using the emerging technology of functional near-infrared spectroscopy (fNIRS), sought to test neural activations to visual and auditory words, as well as auditory and audiovisual sentences (McKay et al. 2016). It should be noted at the outset that this was a preliminary report, testing only two CI users, one with high and one with low speech-perception performance. This study reported that the single CI user with good speech understanding showed activation patterns to AV sentences similar to NH listeners, although the poor performer showed little activity outside of primary and association auditory regions. No direct comparison was made between audiovisual and unisensory responses. Although the sample size prohibits any definitive conclusions from being drawn, this study does show that fNIRS, as a noninvasive technique that is compatible with CIs, is a promising technique for future studies.

A fifth quite different study of multisensory integration in CI users investigated the integration of music in children using EEG (Maglione et al. 2015). In this study, participants were presented with visual stimuli of a 4-minute clip from Disney's *Fantasia* without sound (visual-only), with the original score (congruent audiovisual), and with the reversed score (incongruent). The experimenters were particularly interested in alpha waves originating in the frontal lobes, a biomarker of musical appreciation (one of the most common complaints reported by CI users). Normal-hearing controls showed significant differences in neural responses to all three classes of stimuli. In contrast, CI users did not show any differences between congruent and incongruent audiovisual presentations. Furthermore, only bilaterally implanted users had a significant difference between either of the audiovisual conditions and the visual-only condition. This later finding suggests that bilateral implants may yield more naturalistic audiovisual integration of music. Because of the small number of participants (seven CI users and six controls), these data require further testing in a larger sample before strong conclusions can be made. Additionally, this is, to our knowledge, the only study of audiovisual integration of music in CI users, despite the ubiquity of complaints in reference to quality of music perception postimplantation, particularly in postlingually deafened individuals. This area of inquiry is in great need of attention.

CROSS-MODAL PLASTICITY/ANIMAL MODELS

Given that substantial reorganizational changes take place in brain networks after periods of sensory deprivation (Neville et al. 1998; Bavelier et al. 2001; Wallace et al. 2004; Carriere et al. 2007; Sharma et al. 2009), cross-modal plasticity likely contributes to the unique development of brain networks in CI users. More specifically, a growing body of literature is focused on defining the role of vision in speech recovery through behavioral and neuroanatomical markers of audiovisual fusion, as well as unique visually driven cortical activation in the auditory-deprived brain (Giraud et al. 2001; Bavelier et al. 2006; Merabet & Pascual-Leone 2010; Kral & Sharma 2012). Converging work in this area has investigated cross-modal plasticity in both humans and animals. For example, a causal link between cross-modal plasticity in auditory cortex and specific visual improvements has been demonstrated in the congenitally deaf cat (Lomber et al. 2010). Visual enhancements in these cats are specific to movement and localization tasks that are, respectively, supported by the Dorsal Zone and Posterior Auditory Field–regions typically specialized for higher-order auditory processing. Later anatomical studies further confirmed underlying structural reorganization resulting in the formation of novel connections between nonauditory inputs and auditory fields that are likely to support the perceptual enhancements (Barone et al. 2013; Kok et al. 2014). Together, these studies indicate both structural and functional reorganization in deaf cats and auditory regions that are recruited for specific visual functions. The relationship between these reorganized auditory areas and later auditory habilitation via a CI was further investigated in translational work, which showed that auditory areas responsive to visual cues still retain their responsiveness to native auditory inputs, although they seem to lack bimodal interactions (Land et al. 2016). This, like other findings in humans, suggests that auditory processing can be successfully engaged while multisensory integrative abilities are altered or absent. Interestingly, recent behavioral and physiological experiments in early deafened, bilaterally implanted ferrets have demonstrated that intermodal training (i.e., using interleaved auditory and visual cues on separate trials) can improve auditory-alone localization abilities (Isaiah et al. 2014). This intermodal training paradigm was also seen to enhance neural sensitivity to sound localization cues within the auditory cortex. Thus, vision may facilitate the restoration of auditory function after cochlear implantation, likely through top-down modulations of responses within the auditory cortex (Isaiah et al. 2014; Isaiah & Hartley 2015). In summary, auditory cortical regions are capable of taking on new visual functions in deaf animals, while maintaining the ability of conveying electrical stimulation from auditory neuroprotheses.

Similar to work in animal models of deafness, human behavioral and neuroimaging experiments on deaf subjects have detailed visual reorganization at the structural and functional levels. These include reduced visual reaction times (Bottari et al. 2010, 2011), greater visual discrimination accuracy in the periphery (Bottari et al. 2010), as well as improvements in motion detection and on selective attentional measures (Bavelier et al. 2006; Bottari et al. 2008). Neuroimaging studies of cross-modal plasticity after cochlear implantation have used fNIRS (Saliba et al. 2016), PET (Strelnikov et al. 2015a), and EEG (Sharma et al. 2015) to investigate the neural correlates of these visual perceptual enhancements. Here, we will briefly discuss several representative studies from this substantial body of work (for a thorough review of the topic of cross-modal plasticity in CI users, see Anderson et al. 2016; Stropahl et al. 2016).

Preoperative fMRI of CI candidates allows for high spatial resolution of functional measures that can be correlated to later behavioral assessments of auditory speech proficiency. Two studies have used this strategy to investigate neural activity patterns in CI candidates performing a visual word rhyming task (Lazard et al. 2010) or evoking auditory imagery of environmental sounds from memory (Lazard et al. 2013). Their findings indicate that more proficient CI users exhibit dorsal phonological processing compared to more ventral processing in poor performing CI users (Lazard et al. 2010). Although the distinction between dorsal and ventral processing streams has been well described in the visual domain since the 1980s (see Ungerleider & Haxby 1994), a similar concept in the auditory system is relatively new (Hickok & Poeppel 2007). In both modalities, the functional distinction is between the dorsal "where pathway" of object identification and the ventral "what pathway" of spatial processing. Interestingly, this dorsal–ventral processing distinction made by Lazard et al. (2010) also corroborates prior studies relating higher resting state ventro-temporal metabolism to lower CI proficiency (i.e., in contrast to higher outcomes with dorsolateral prefrontal activity; Giraud & Lee 2007). Despite these consistencies, it should be noted that there is not conclusive evidence to date suggesting that ventral processing streams are more susceptible to reorganization than dorsal ones (Vachon et al. 2013).

Higher CI outcomes postimplantation are seen when the typical left lateralization of phonological processing is preserved via a rhyming task comparing the endings of visually presented words (Lazard et al. 2010), as well as the typical right lateralization of environmental sound imagery (Lazard et al. 2013). These findings seem to align with studies in other imaging modalities that correlate broader cortical activation (including across both hemispheres) to both auditory (Olds et al. 2016) and visual stimuli (Doucet et al. 2006) with lower CI outcomes. Instead, more focal activity (either intramodal or cross-modal) may indicate more efficient, localized processing without the need for more expansive cortical recruitment. The aforementioned future CI users also exhibit an overall reduction in temporal lobe activation to auditory phonemic processing that further decreases with longer durations of auditory deprivation (Lazard et al. 2013). Taken together, these findings suggest that functional reorganization increases with longer periods of auditory deprivation, which negatively impacts auditory rehabilitation. Similarly, EEG source localization in postlingually deafened CI users has indicated so-called maladaptive plasticity, whereby activation of right auditory cortex in response to visual, flashing checkerboards was inversely related to speech recognition (Sandmann et al. 2012). Indeed, the more hypometabolic or unrecruited the auditory cortex is during deafness, the more successful speech recognition appears to be after implantation (Lee et al. 2001). From these publications, it seems that more extensive reorganization of auditory cortex negatively predicts success with a CI.

When examining beyond auditory cortex, Strelnikov et al. (2015a) make an important distinction that postlingual CI users seem to have high functional activity in visual cortex that is positively associated with speech outcomes. These findings are more in line with the notion that visual proficiency could serve a compensatory role, particularly in multisensory speech recovery. That is, higher functional connectivity between multisensory STS and visual cortex may better facilitate audiovisual integration to benefit comprehension (Strelnikov et al. 2013). Interestingly, a study using PET to study speech-induced activity also reported that CI users had altered functional specificity of the superior temporal cortex such that an increasing contribution of visual cortex to speech recognition was positively correlated with speechreading ability (Giraud et al. 2001). This suggests that CI users were actively using enhanced audiovisual integration to facilitate their learning of the novel and degraded speech sounds from a CI. In a recent dual EEG-functional near-infrared spectroscopy (fNIRS) study, more efficient auditory processing in NH controls and more efficient *visual* processing in CI users was observed during a sensory adaptation task, whereby percent signal change is measured during the repetition of an identical stimulus over several seconds (i.e., tones or visual checkerboards) (Chen et al. 2016). Future work is required to conclude whether such intramodal visual enhancements benefit speech outcomes more in postlingual than prelingually deafened CI users for whom multisensory integration may be underdeveloped. Recently, higher auditory cortex activation during a visual discrimination task was also positively related to face recognition abilities in postlingually deafened CI users (Stropahl et al. 2015a). Because this finding was specific to faces (and not images of houses), it also supports the idea that functionally selective plasticity may preferentially stem from the processing of the highly ecologically relevant stimuli needed for speech understanding (Heimler et al. 2014; Stropahl et al. 2015a).

Finally, it should be noted that cross-modal plasticity is not specific to profound deafness but has also been identified even with partial hearing loss. Notably, changes in functional connectivity of individuals with only high frequency hearing loss have been reported (Puschmann & Thiel 2016). This study implemented an fMRI task of auditory stimulus categorization in which audible low frequency sounds were paired with matched or mismatched visual motion cues (i.e., ascending or descending in space and frequency). Interestingly, increased functional connectivity between auditory cortex and the right middle temporal visual area was found to be a function of the degree of hearing loss. Similarly, a recent study using EEG source localization (Campbell & Sharma 2014) reported more visually evoked activity in the temporal cortex of older adults with mild to moderate hearing loss. Thus, cross-modal plasticity may begin during the earliest stages of hearing impairment, may expand as deafness progresses, and appears to persist after cochlear implantation.

In conclusion, a great deal of work has provided compelling evidence for the activation of auditory cortex by visual stimuli in the hearing impaired. These include responses to visual speech (Stropahl et al. 2015a), motion (Finney et al. 2001), checkerboards (Sandmann et al. 2012), and an apparent motion illusion (Doucet et al. 2006). Careful distinctions should be made between pre- and postlingual CI users when interpreting the influence of such reorganizational changes on ongoing or later auditory re/habilitation.

AUDIOVISUAL SPEECH INTEGRATION IN OTHER HEARING-IMPAIRED POPULATIONS

Although this review is primarily focused on CI users, we would be remiss to not touch on audiovisual integration in hearing-impaired (HI) individuals who use other forms of auditory prostheses, including hearing aids, or no prostheses. In a well-designed study of adults with acquired hearing loss, Tye-Murray et al (2007) found that the audiovisual integrative processes underlying speech perception in noise was remarkably similar for NH and HI individuals when auditory-only performance was matched. These results held for individual consonants, words, and sentences. It should be noted here that such individuals with acquired hearing loss did have NH in their formative developmental years, and thus, the typical integrative processes underlying audiovisual enhancement remained unaltered after hearing loss when auditory performance was matched—a finding that has been replicated (Walden et al. 2001; Bernstein & Grant 2009). Like NH and CI users, however, there is a wide range of integrative abilities in HI listeners (Grant et al. 1998). In addition to varying across HI listeners, the ability to benefit from the inclusion of visual speech is dependent upon the complementary information provided by the visual speech cues. Visual speech has widely been found to provide a great deal of information on place of articulation, whereas auditory speech more reliably provides voicing and manner of articulation information. In a seminal study, Walden et al (1974) showed that multisensory enhancement in HI listeners was greatest when auditory and visual information were complementary as opposed to redundant. Thus, the relative comparability of enhancement between NH and HI listeners may vary depending on which consonants are being presented (Busacco 1988; Grant & Walden 1995, 1996; Walden et al. 2001). These findings of how the complementary (versus redundant) nature of auditory and visual information drive multisensory enhancement in HI highlights the need for similar studies in CI users, an area that has been understudied.

HI listeners who use prostheses other than CIs, such as hearing aids, have also been the focus of a number of studies of audiovisual speech integration. HI listeners using hearing aids have been shown to benefit from being presented with both auditory and visual speech information (Moradi et al. 2016; Walden et al. 2001) and, in some instances, even to show greater multisensory enhancement than their NH peers (Moradi et al. 2016). The use of hearing aids specifically provides a boost in audiovisual enhancement when the salience of manner of articulation and voicing are increased through amplification, providing auditory information that is more complementary to the visual speech cues (Walden et al. 2001). Though, to our knowledge, not previously studied with CI users, it seems logical that these findings could also be applied to CI device configuration—by configuring device settings to specifically amplify manner of articulation and voicing, it may be possible to increase the complementarity of auditory and visual information, leading to greater multisensory gains for the listener.

AREAS OF FUTURE INQUIRY

Throughout this review, we have pointed out areas in need of research. Here, we will highlight those that seem in most need, and which we believe will have the most significant effect. Clearly, as this review reveals, there is an extreme dearth of work on low-level multisensory sensory perception in CI users. Speech processing is an inherently multisensory process and is very much dependent on lower-level processing. Abilities in these low-level processes contribute to relative abilities in speech perception per se, and thus, this is an area that is extremely important in terms of knowledge of the mechanisms through which CIs improve speech perception. Studies comparing pre- and postlingually deafened CI users, as well as NH listeners, are needed and should be structured so as to investigate a wide variety of low-level multisensory tasks. Canonical redundant-target detection tasks would be an excellent place to begin. Furthermore, no studies to date have measured multisensory spatial perception and its influence on integration in CI users.

Also, as highlighted earlier, little research has been completed toward understanding how CI users perceive multisensory speech signals outside of phoneme/word recognition. Whether individuals with CIs gain audiovisual benefit for suprasegmental feature perception with current technology and in tasks that individuals with NH do gain audiovisual benefit is unknown. For example, most research showing audiovisual benefit with suprasegmental features in NH individuals is in language acquisition, both in infants and in second-language learning in adults (Busso et al. 2004; Cunillera et al. 2010). Recently implanted individuals with CIs must adapt to a very different incoming auditory signal comprising envelope cues in current pulses and might be compared to adults learning a second language. Research using a wider variety of suprasegmental tasks with modern CI technology in new and experienced CI recipients needs to be conducted to determine if individuals with CIs have atypical audiovisual integration for suprasegmental feature perception.

Another area that could benefit from a more multisensory approach would be in postimplantation clinical models. One study to date has investigated the efficacy of audiovisual training relative to auditory-only training. The typical auditory-only model assumes that by actively forcing CI recipients to use the auditory information only, it will be learned more quickly. With that said, it is possible that by presenting auditory and visual information together, CI recipients will be able to link their experience with speechreading to the statistical regularities of auditory speech and improving their speech perception abilities. Only a single study has tested these options empirically, with mixed results (Bernstein et al. 2014). Auditory-only testing, when following auditory-only training, showed greater improvements than when following audiovisual training. Audiovisual training produced higher levels of audiovisual speech perception (as measured during the training itself), begging the question as to which should in fact be the tested metric, audiovisual speech or auditory-only speech? Regardless, there is a great need for such studies, and the clinical effect of this work will be quite significant.

There are, of course, many other areas in need of more research. How multisensory integration develops in children with CIs, the effect of prelingual implantation in children, the effect that preimplant visual performance has on postimplant perception, and the effect that duration of deafness in postlingually deafened adults has on integration to name a few. In every case, however, there is a strong need for a rigorous definition of age of implantation (in terms of early or late implantation), consistent inclusion of audio-only, visual-only, and audiovisual conditions as opposed to only including a single unisensory baseline, and the ubiquitous inclusion of a well-matched control group of NH listeners.

Furthermore, in the vast majority of the CI literature, CI users are treated as a homogenous group, whereas in reality, there is quite a diversity of hearing experience within the CI population. Although this has been mentioned throughout the manuscript, it is worth addressing here. Though studies have segregated cohorts into having received early or late implantations, or pre- and post-lingual implantations, even within these delineated groups, there will inherently be a wide range of individual differences in hearing experience. For example, many children with congenital hearing loss retain some level of residual hearing. Likewise, individuals who received CIs later in life have varying levels and types of hearing loss, as well as various experience with other hearing-enhancing prostheses, such as hearing aids. This oversimplification of either being a CI user of NH listener undoubtedly overlooks many important nuances in auditory, visual, and audiovisual speech perception abilities within individuals and should be a focus of future research.

CONCLUSION

This review of audiovisual integration in CI users reveals a number of trends in the field. First of all, it is clear that CI users, regardless of listening history, are typically able to show at least some perceptual gain from multisensory integration. The extent of this gain, however, varied based on the age of implantation and varied between components of speech (phoneme–word–suprasegmental). A number of consistent findings were observed in adults who had received CIs:

- High-performing adults with CIs are able to obtain multisensory integration benefits similar to or in excess of individuals with NH. However, this effect may be affected by age and CI experience and also does not extend to lower-performing adult CI users.
- Though multisensory gains in adults with CIs are often similar to those in NH controls, the pattern of auditory, visual, and audiovisual responses suggests that the underlying integrative processes may differ between these two groups.
- Comparisons in multisensory gain between adults with and without CIs varied according to stimulus property—adults with CIs showed multisensory gain in word and phoneme recognition, but not in suprasegmental feature processing or low-level stimulus detection.
- High levels of visual speechreading proficiency before implantation may lead to a reduction in CI proficiency because of a decreased neuroplasticity in what are typically auditory brain networks reorganized to process visual inputs.

In children with CIs, there were also a number of consistent findings:

- Age of implantation was of paramount importance. Children who were implanted early showed multisensory benefits similar to NH listeners. This finding highlights the need for early detection of, and intervention for, hearing loss, not only in terms of auditory perception per se, but also in terms of multisensory processing and the associated behavioral and perceptual benefits.

- CI experience is influential in children, where increases in CI experience are associated with increased audiovisual integration.
- Communication environment influences proficiency in children using CIs, with auditory/oral communication leading to better outcomes than total communication environments, both in terms of auditory-only and audiovisual speech comprehension.
- Increases in speech perception may be linked to improved speech production.
- Future research, particularly in low-level processing tasks such as signal detection, will help to further define a sensitive period of audiovisual integration and the mechanism for differences in audiovisual integration for individuals with and without CIs.

ACKNOWLEDGMENTS

I. B. was supported by the National Institutes of Health under the award numbers T32 MH 064913 and F31 DC 015956. R. S. was funded by a Banting Fellowship from the Canadian National Science and Engineering Research Counsel, the Autism Research Training Program funded by the Canadian Health Institutes of Health Research, and the University of Western Ontario Faculty Development fund.

R.G. was a member of the Audiology Advisory Board for Advanced Bionics and Cochlear Americas at the time of publication. No conflicts of interest or sources of funding are declared for any of the other authors.

Address for correspondence: Ryan A. Stevenson, Department of Psychology, University of Western Ontario, Ontario, Canada. E-mail: rsteve28@uwo.ca

Received May 2, 2016; accepted January 30, 2017.

REFERENCES

Agelfors, E. (1996). A comparison between patients using cochlear implants and hearing aids. Part I: Results on speech tests. *Quarterly Progress and Status Report*.

Altieri, N., Stevenson, R. A., Wallace, M. T., et al. (2015). Learning to associate auditory and visual stimuli: behavioral and neural mechanisms. *Brain Topogr*, *28*, 479–493.

Anderson, C. A., Lazard, D. S., Hartley, D. E. (2017). Plasticity in bilateral superior temporal cortex: Effects of deafness and cochlear implantation on auditory and visual speech processing. *Hear Res*, *343*, 138–149.

Barker, B. A., & Tomblin, J. B. (2004). Bimodal speech perception in infant hearing aid and cochlear implant users. *Arch Otolaryngol Head Neck Surg*, *130*, 582–586.

Barone, P., Strelnikov, K., Déguine, O. (2013). Role of audiovisual plasticity in speech recovery after adult cochlear implantation. In *AVSP* (pp. 99–104).

Başkent, D., & Bazo, D. (2011). Audiovisual asynchrony detection and speech intelligibility in noise with moderate to severe sensorineural hearing impairment. *Ear Hear*, *32*, 582–592.

Baum, S. H., Stevenson, R. A., Wallace, M. T. (2015a). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Prog Neurobiol*, *134*, 140–160.

Baum, S. H., Stevenson, R. A., Wallace, M. T. (2015b). Testing sensory and multisensory function in children with autism spectrum disorder. *J Vis Exp*, e52677.

Bavelier, D., Brozinsky, C., Tomann, A., et al. (2001). Impact of early deafness and early exposure to sign language on the cerebral organization for motion processing. *J Neurosci*, *21*, 8931–8942.

Bavelier, D., Dye, M. W., Hauser, P. C. (2006). Do deaf individuals see better? *Trends Cogn Sci*, *10*, 512–518.

Beauchamp, M. S. (2005). Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics*, *3*, 93–113.

Beauchamp, M. S., Argall, B. D., Bodurka, J., et al. (2004). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci*, *7*, 1190–1192.

Beauchamp, M. S., Lee, K. E., Argall, B. D., et al. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*, 809–823.

Beauchamp, M. S., Yasar, N. E., Frye, R. E., et al. (2008). Touch, sound and vision in human superior temporal sulcus. *Neuroimage*, *41*, 1011–1020.

Bebko, J. M., Schroeder, J. H., Weiss, J. A. (2014). The McGurk effect in children with autism and Asperger syndrome. *Autism Res*, *7*, 50–59.

Bebko, J. M., Weiss, J. A., Demark, J. L., et al. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *J Child Psychol Psychiatry*, *47*, 88–98.

Bergeson, T. R., Houston, D. M., Miyamoto, R. T. (2010). Effects of congenital hearing loss and cochlear implantation on audiovisual speech perception in infants and children. *Restor Neurol Neurosci*, *28*, 157–165.

Bergeson, T. R., Pisoni, D. B., Davis, R. A. (2003). A longitudinal study of audiovisual speech perception by children with hearing loss who have cochlear implants. *Volta Rev*, *103*, 347–370.

Bergeson, T. R., Pisoni, D. B., Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear Hear*, *26*, 149–164.

Bernstein, J. G., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, *125*, 3358–3372.

Bernstein, L. E., Eberhardt, S. P., Auer, E. T. Jr. (2014). Audiovisual spoken word training can promote or impede auditory-only perceptual learning: prelingually deafened adults with late-acquired cochlear implants versus normal hearing adults. *Front Psychol*, *5*, 934.

Bernstein, L. E., Eberhardt, S. P., Demorest, M. E. (1989). Single-channel vibrotactile supplements to visual perception of intonation and stress. *J Acoust Soc Am*, *85*, 397–405.

Bernstein, L. E., Lu, Z. L., Jiang, J. (2008). Quantified acoustic-optical speech signal incongruity identifies cortical sites of audiovisual speech processing. *Brain Res*, *1242*, 172–184.

Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In M. Sabourin, F. I. M. Craik, M. Robert (Eds.), *Advances in Psychological Science, Vol. 2: Biological and Cognitive Aspects* (pp. 419–439). Hove, UK: Psychological Press.

Bichey, B. G., & Miyamoto, R. T. (2008). Outcomes in bilateral cochlear implantation. *Otolaryngol Head Neck Surg*, *138*, 655–661.

Bishop, C. W., & Miller, L. M. (2009). A multisensory cortical network for understanding speech in noise. *J Cogn Neurosci*, *21*, 1790–1805.

Bond, M., Mealing, S., Anderson, R., et al. (2009). The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: a systematic review and economic model. *Health Technol Assess*, *13*, 1–330.

Bottari, D., Caclin, A., Giard, M. H., et al. (2011). Changes in early cortical visual processing predict enhanced reactivity in deaf individuals. *PLoS One*, *6*, e25607.

Bottari, D., Nava, E., Ley, P., et al. (2010). Enhanced reactivity to visual stimuli in deaf individuals. *Restor Neurol Neurosci*, *28*, 167–179.

Bottari, D., Turatto, M., Bonfioli, F., et al. (2008). Change blindness in profoundly deaf individuals and cochlear implant recipients. *Brain Res*, *1242*, 209–218.

Busacco, D. A. (1988). *The Effects of Age on the Benefit Derived from Visual Cues in Auditory-visual Speech Recognition by the Hearing Impaired*. UMI. Doctoral dissertation, Columbia University, Chicago, Illinois.

Busso, C., Deng, Z., Yildirim, S., et al. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 205–211). ACM.

Callan, D. E., Jones, J. A., Munhall, K., et al. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport*, *14*, 2213–2218.

Callan, D. E., Jones, J. A., Munhall, K., et al. (2004). Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. *J Cogn Neurosci*, *16*, 805–816.

Calvert, G., & Lewis, J. (2004). Hemodynamic studies of audiovisual interaction. In G. A. Calvert, C. Spence, B. E. Stein (Eds), *The Handbook of Multisensory Perception* (pp. 483–502). Cambridge, MA: MIT Press.

Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: the neural substrates of visible speech. *J Cogn Neurosci*, *15*, 57–70.

Calvert, G. A., Campbell, R., Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol*, *10*, 649–657.

Calvert, G. A., Hansen, P. C., Iversen, S. D., et al. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, *14*, 427–438.
Campbell, R. (2008). The processing of audio-visual speech: empirical and neural bases. *Philos Trans R Soc Lond B Biol Sci*, *363*, 1001–1010.
Campbell, J., & Sharma, A. (2014). Cross-modal re-organization in adults with early stage hearing loss. *PLoS One*, *9*, e90594.
Carriere, B. N., Royal, D. W., Perrault, T. J., et al. (2007). Visual deprivation alters the development of cortical multisensory integration. *J Neurophysiol*, *98*, 2858–2867.
Carriere, B. N., Royal, D. W., Wallace, M. T. (2008). Spatial heterogeneity of cortical receptive fields and its impact on multisensory interactions. *J Neurophysiol*, *99*, 2357–2368.
Champoux, F., Lepore, F., Gagné, J. P., et al. (2009). Visual stimuli can impair auditory processing in cochlear implant users. *Neuropsychologia*, *47*, 17–22.
Chen, L. C., Stropahl, M., Schönwiesner, M., et al. (2017). Enhanced visual adaptation in cochlear implant users revealed by concurrent EEG-fNIRS. *Neuroimage*, *146*, 600–608.
Coez, A., Zilbovicius, M., Ferrary, E., et al. (2008). Cochlear implant benefits in deafness rehabilitation: PET study of temporal voice activations. *J Nucl Med*, *49*, 60–67.
Collignon, O., Girard, S., Gosselin, F., et al. (2008). Audio-visual integration of emotion expression. *Brain Res*, *1242*, 126–135.
Conrey, B., & Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *J Acoust Soc Am*, *119*, 4065–4073.
Conrey, B. L., & Pisoni, D. B. (2004). Detection of Auditory-Visual Asynchrony in Speech and Nonspeech Signals. In D. B. Pisoni (Ed.), *Research on Spoken Language Processing* (pp. 71–94). Bloomington: Indiana University.
Conway, C. M., Pisoni, D. B., Anaya, E. M., et al. (2011). Implicit sequence learning in deaf children with cochlear implants. *Dev Sci*, *14*, 69–82.
Conway, C. M., Pisoni, D. B., Kronenberger, W. G. (2009). The Importance of Sound for Cognitive Sequencing Abilities: The Auditory Scaffolding Hypothesis. *Curr Dir Psychol Sci*, *18*, 275–279.
Cunillera, T., Camara, E., Laine, M., et al. (2010). Speech segmentation is facilitated by visual cues. *Q J Exp Psychol (Hove)*, *63*, 260–274.
Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear Res*, *229*, 132–147.
de Boer-Schellekens, L., Eussen, M., Vroomen, J. (2013). Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Front Integr Neurosci*, *7*, 8.
de Gelder, B., Vroomen, J., Van der Heide, L. (1991). Face recognition and lip-reading in autism. *Eur J Cogn Psychol*, *3*, 69–86.
Desai, S., Stickney, G., Zeng, F. G. (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *J Acoust Soc Am*, *123*, 428–440.
Dick, A. S., Solodkin, A., Small, S. L. (2010). Neural development of networks for audiovisual speech comprehension. *Brain Lang*, *114*, 101–114.
Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept Psychophys*, *66*, 1388–1404.
Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*, 719–721.
Dohen, M., Lœvenbruck, H., Cathiard, M.-A., et al. (2004). Visual perception of contrastive focus in reiterant French speech. *Speech Communication*, *44*, 155–172.
Dorman, M. F., Liss, J., Wang, S., et al. (2016). Experiments on Auditory-Visual Perception of Sentences by Users of Unilateral, Bimodal, and Bilateral Cochlear Implants. *J Speech Lang Hear Res*, *59*, 1505–1519.
Doucet, M. E., Bergeron, F., Lassonde, M., et al. (2006). Cross-modal reorganization and speech perception in cochlear implant users. *Brain*, *129*(pt 12), 3376–3383.
Ethofer, T., Pourtois, G., Wildgruber, D. (2006). Investigating audiovisual integration of emotional signals in the human brain. *Prog Brain Res*, *156*, 345–361.
Finney, E. M., Fine, I., Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nat Neurosci*, *4*, 1171–1173.
Gaylor, J. M., Raman, G., Chung, M., et al. (2013). Cochlear implantation in adults: a systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg*, *139*, 265–272.
Geers, A., & Brenner, C. (1994). Speech Perception Results: Audition and Lipreading Enhancement. *Volta Review*, *96*, 97–108.
Gilley, P. M., Sharma, A., Mitchell, T. V., et al. (2010). The influence of a sensitive period for auditory-visual integration in children with cochlear implants. *Restor Neurol Neurosci*, *28*, 207–218.
Giraud, A. L., & Lee, H. J. (2007). Predicting cochlear implant outcome from brain organisation in the deaf. *Restor Neurol Neurosci*, *25*, 381–390.
Giraud, A. L., Price, C. J., Graham, J. M., et al. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*, 657–663.
Goh, W. D., Pisoni, D. B., Kirk, K. I., et al. (2001). Audio-visual perception of sinewave speech in an adult cochlear implant user: a case study. *Ear Hear*, *22*, 412–419.
Grant, K., & Walden, B. (1995). Predicting auditory-visual speech recognition in hearing-impaired listeners. In *Proceedings of the XIIIth International Congress of Phonetic Sciences* (pp. 122–129).
Grant, K. W., & Walden, B. E. (1996). Evaluating the articulation index for auditory-visual consonant recognition. *J Acoust Soc Am*, *100*(4 pt 1), 2415–2424.
Grant, K. W., Walden, B. E., Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: consonant recognition, sentence recognition, and auditory-visual integration. *J Acoust Soc Am*, *103*(5 pt 1), 2677–2690.
Hairston, W. D., Burdette, J. H., Flowers, D. L., et al. (2005). Altered temporal profile of visual-auditory multisensory interactions in dyslexia. *Exp Brain Res*, *166*, 474–480.
Hay-McCutcheon, M. J., Pisoni, D. B., Hunt, K. K. (2009). Audiovisual asynchrony detection and speech perception in hearing-impaired listeners with cochlear implants: a preliminary analysis. *Int J Audiol*, *48*, 321–333.
Hay-McCutcheon, M. J., Pisoni, D. B., Kirk, K. I. (2005). Audiovisual speech perception in elderly cochlear implant recipients. *Laryngoscope*, *115*, 1887–1894.
Heimler, B., Weisz, N., Collignon, O. (2014). Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience*, *283*, 44–63.
Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *J Exp Psychol*, *63*, 289–293.
Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*, *8*, 393-402.
Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Dev Sci*, *15*, 688–696.
Hillock, A. R., Powers, A. R., Wallace, M. T. (2011). Binding of sights and sounds: age-related changes in multisensory temporal processing. *Neuropsychologia*, *49*, 461–467.
Hockley, N. S., Polka, L. (1994). A developmental study of audiovisual speech perception using the McGurk paradigm. *J Acoust Soc Am*, *96*, 3309–3309.
Holt, R. F., Kirk, K. I., Hay-McCutcheon, M. (2011). Assessing multimodal spoken word-in-sentence recognition in children with normal hearing and children with cochlear implants. *J Speech Lang Hear Res*, *54*, 632–657.
Houston, D. M., Beer, J., Bergeson, T. R., et al. (2012). The ear is connected to the brain: some new directions in the study of children with cochlear implants at Indiana University. *J Am Acad Audiol*, *23*, 446–463.
Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *J Comp Neurol*, *387*, 167–178.
Huyse, A., Berthommier, F., Leybaert, J. (2013). Degradation of labial information modifies audiovisual speech perception in cochlear-implanted children. *Ear Hear*, *34*, 110–121.
Irwin, J. R., Tornatore, L. A., Brancazio, L., et al. (2011). Can children with autism spectrum disorders "hear" a speaking face? *Child Dev*, *82*, 1397–1403.
Isaiah, A., & Hartley, D. E. (2015). Can training extend current guidelines for cochlear implant candidacy? *Neural Regen Res*, *10*, 718–720.
Isaiah, A., Vongpaisal, T., King, A. J., et al. (2014). Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J Neurosci*, *34*, 11119–11130.
Jahn, K. N., Stevenson, R. A., Wallace, M. T. (2017). Visual temporal acuity is related to auditory speech perception abilities in cochlear implant users. *Ear and Hearing*, *38*, 236–243.
James, T. W., & Stevenson, R. A. (2012). The use of fMRI to assess multisensory integration. In M. M. Murray, M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton, FL.

James, T. W., Kim, S., Stevenson, R. A. (2009). Assessing multisensory interaction with additive factors and functional MRI. In *The International Society for Psychophysics*. Galway, Ireland.

James, T. W., Stevenson, R. A., Kim, S. (2012). Inverse effectiveness in multisensory processing. In B. E. Stein (Ed.), *The New Handbook of Multisensory Processes*. Cambridge, MA: MIT Press.

James, T. W., VanDerKlok, R. M., Stevenson, R. A., et al. (2011). Multisensory perception of action in posterior temporal and parietal cortices. *Neuropsychologia*, *49*, 108–114.

Jerger, S., Lewis, S., Hawkins, J., et al. (1980). Pediatric speech intelligibility test. I. Generation of test materials. *Int J Pediatr Otorhinolaryngol*, *2*, 217–230.

Jones, J. A., & Callan, D. E. (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport*, *14*, 1129–1133.

Kaiser, A. R., Kirk, K. I., Lachs, L., et al. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *J Speech Lang Hear Res*, *46*, 390–404.

Kim, S., & James, T. W. (2010). Enhanced effectiveness in visuo-haptic object-selective brain regions with increasing stimulus salience. *Hum Brain Mapp*, *31*, 678–693.

Kim, S., Stevenson, R. A., James, T. W. (2012). Visuo-haptic neuronal convergence demonstrated with an inversely effective pattern of BOLD activation. *J Cogn Neurosci*, *24*, 830–842.

Kirk, K. I., Hay-McCutcheon, M. J., Holt, R. F., et al. (2007). Audiovisual Spoken Word Recognition by Children with Cochlear Implants. *Audiol Med*, *5*, 250–261.

Kok, M. A., Chabot, N., Lomber, S. G. (2014). Cross-modal reorganization of cortical afferents to dorsal auditory cortex following early- and late-onset deafness. *J Comp Neurol*, *522*, 654–675.

Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci*, *35*, 111–122.

Kral, A., Kronenberger, W. G., Pisoni, D. B., et al. (2016). Neurocognitive factors in sensory restoration of early deafness: a connectome model. *Lancet Neurol*, *15*, 610–621.

Kreifelts, B., Ethofer, T., Grodd, W., et al. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage*, *37*, 1445–1456.

Krueger, J., Royal, D. W., Fister, M. C., et al. (2009). Spatial receptive field organization of multisensory neurons and its impact on multisensory interactions. *Hear Res*, *258*, 47–54.

Krueger Fister, J., Stevenson, R. A., Nidiffer, A. R., et al. (2016). Stimulus intensity modulates multisensory temporal processing. *Neuropsychologia*, *88*, 92–100.

Lachs, L., & Hernandez, L. R. (1998). Update: The Hoosier Audiovisual Multitalker Database. In D. B. Pisoni (Ed.), *Research on Spoken Language Processing* (pp. 377–388). Bloomington, IN: Speech Research Laboratory, Indiana University.

Lachs, L., Pisoni, D. B., Kirk, K. I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: a first report. *Ear Hear*, *22*, 236–251.

Land, R., Baumhoff, P., Tillein, J., et al. (2016). Cross-Modal Plasticity in Higher-Order Auditory Cortex of Congenitally Deaf Cats Does Not Limit Auditory Responsiveness to Cochlear Implants. *J Neurosci*, *36*, 6175–6185.

Landry, S. P., Guillemot, J. P., Champoux, F. (2013). Temporary deafness can impair multisensory integration: a study of cochlear-implant users. *Psychol Sci*, *24*, 1260–1268.

Lazard, D. S., Lee, H. J., Gaebler, M., et al. (2010). Phonological processing in post-lingual deafness and cochlear implant outcome. *Neuroimage*, *49*, 3443–3451.

Lazard, D. S., Lee, H. J., Truy, E., et al. (2013). Bilateral reorganization of posterior temporal cortices in post-lingual deafness and its relation to cochlear implant outcome. *Hum Brain Mapp*, *34*, 1208–1219.

Lee, D. S., Lee, J. S., Oh, S. H., et al. (2001). Cross-modal plasticity and cochlear implants. *Nature*, *409*, 149–150.

Lee, H. J., Truy, E., Mamou, G., et al. (2007). Visual speech circuits in profound acquired deafness: a possible role for latent multimodal connectivity. *Brain*, *130*(pt 11), 2929–2941.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends Cogn Sci*, *13*, 470–478.

Leybaert, J., & LaSasso, C. J. (2010). Cued speech for enhancing speech perception and first language development of children with cochlear implants. *Trends Amplif*, *14*, 96–112.

Lomber, S. G., Meredith, M. A., Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat Neurosci*, *13*, 1421–1427.

Lovelace, C. T., Stein, B. E., Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res Cogn Brain Res*, *17*, 447–453.

Macaluso, E., George, N., Dolan, R., et al. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage*, *21*, 725–732.

MacSweeney, M., Calvert, G. A., Campbell, R., et al. (2002). Speechreading circuits in people born deaf. *Neuropsychologia*, *40*, 801–807.

Maglione, A. G., Scorpecci, A., Malerba, P., et al. (2015). Alpha EEG Frontal Asymmetries during Audiovisual Perception in Cochlear Implant Users. A Study with Bilateral and Unilateral Young Users. *Methods Inf Med*, *54*, 500–504.

Massaro, D. W. (1984). Children's perception of visual and auditory speech. *Child Dev*, *55*, 1777–1788.

Massaro, D. W. (1987a). In B. Dodd, B. A. Campbell (Eds.), *Hearing by Eye: The Psychology of Lip Reading* (pp. 53–84). Lawrence Erlbaum Associates Ltd., Publishers, Mahwah, NJ, USA.

Massaro, D. W. (1987b). Categorical partition: A fuzzy logical model of categorization behavior. *Categorical Perception: The Groundwork of Cognition*, 254–283.

Massaro, D. W. (2004). In G. Calvert, C. Spence, B. E. Stein (Eds.), *The Handbook of Multisensory Processes* (pp. 153–176). Cambridge, MA: MIT Press.

Massaro, D. W., Thompson, L. A., Barron, B., et al. (1986). Developmental changes in visual and auditory contributions to speech perception. *J Exp Child Psychol*, *41*, 93–113.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

McKay, C. M., Shah, A., Seghouane, A. K., et al. (2016). Connectivity in Language Areas of the Brain in Cochlear Implant Users as Revealed by fNIRS. In *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* (pp. 327–335).

Merabet, L. B., & Pascual-Leone, A. (2010). Neural reorganization following sensory loss: the opportunity of change. *Nat Rev Neurosci*, *11*, 44–52.

Meredith, M. A., & Stein, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Res*, *365*, 350–354.

Meredith, M. A., & Stein, B. E. (1986b). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol*, *56*, 640–662.

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *J Neurophysiol*, *75*, 1843–1857.

Meredith, M. A., Nemitz, J. W., Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci*, *7*, 3215–3229.

Miller, L. M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *J Neurosci*, *25*, 5884–5893.

Mongillo, E. A., Irwin, J. R., Whalen, D. H., et al. (2008). Audiovisual processing in children with and without autism spectrum disorders. *J Autism Dev Disord*, *38*, 1349–1358.

Moody-Antonio, S., Takayanagi, S., Masuda, A., et al. (2005). Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol Neurotol*, *26*, 649–654.

Moradi, S., Lidestam, B., Rönnberg, J. (2016). Comparison of gated audiovisual speech identification in elderly hearing aid users and elderly normal-hearing individuals effects of adding visual cues to auditory speech stimuli. *Trends in Hearing*, *20*, 2331216516653355.

Most, T., & Aviner, C. (2009). Auditory, visual, and auditory-visual perception of emotions by individuals with cochlear implants, hearing AIDS, and normal hearing. *J Deaf Stud Deaf Educ*, *14*, 449–464.

Müller, V. I., Cieslik, E. C., Turetsky, B. I., et al. (2012). Crossmodal interactions in audiovisual emotion processing. *Neuroimage*, *60*, 553–561.

Neil, P. A., Chee-Ruiter, C., Scheier, C., et al. (2006). Development of multisensory spatial integration and perception in humans. *Dev Sci*, *9*, 454–464.

Nelson, W. T., Hettinger, L. J., Cunningham, J. A., et al. (1998). Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Hum Factors*, *40*, 452–460.

Neville, H. J., Bavelier, D., Corina, D., et al. (1998). Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. *Proc Natl Acad Sci U S A*, *95*, 922–929.
NIDCD. (2014). Cochlear Impants. Retrieved 2016, 2016 from https://www.nidcd.nih.gov/health/cochlear-implants.
Nidiffer, A. R., Stevenson, R. A., Krueger Fister, J., et al. (2016). Interactions between space and effectiveness in human multisensory performance. *Neuropsychologia*, *88*, 83–91.
Nishimura, H., Hashikawa, K., Doi, K., et al. (1999). Sign language 'heard' in the auditory cortex. *Nature*, *397*, 116.
Noel, J. P., De Niear, M. A., Stevenson, R. A., et al. (2017). Atypical rapid audio visual temporal recalibration in autism spectrum disorders. *Autism Res*, *10*, 121–129.
Ojanen, V., Möttönen, R., Pekkola, J., et al. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage*, *25*, 333–338.
Olds, C., Pollonini, L., Abaya, H., et al. (2016). Cortical activation patterns correlate with speech understanding after cochlear implantation. *Ear Hear*, *37*, e160–e172.
Pearl, D., Yodashkin-Porat, D., Katz, N., et al. (2009). Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Compr Psychiatry*, *50*, 186–192.
Pekkola, J., Ojanen, V., Autti, T., et al. (2006). Attention to visual speech gestures enhances hemodynamic activity in the left planum temporale. *Hum Brain Mapp*, *27*, 471–477.
Petersen, B., Gjedde, A., Wallentin, M., et al. (2013). Cortical plasticity after cochlear implantation. *Neural Plast*, *2013*, 318521.
Petitto, L. A., Zatorre, R. J., Gauna, K., et al. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: implications for the neural basis of human language. *Proc Natl Acad Sci*, *97*, 13961–13966.
Polley, D. B., Hillock, A. R., Spankovich, C., et al. (2008). Development and plasticity of intra- and intersensory information processing. *J Am Acad Audiol*, *19*, 780–798.
Powers, A. R. 3rd, Hevey, M. A., Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *J Neurosci*, *32*, 6263–6274.
Puschmann, S., & Thiel, C. M. (2017). Changed crossmodal functional connectivity in older adults with hearing loss. *Cortex*, *86*, 109–122.
Rabinowitz, W. M., Eddington, D. K., Delhorne, L. A., et al. (1992). Relations among different measures of speech reception in subjects using a cochlear implant. *J Acoust Soc Am*, *92*(4 pt 1), 1869–1881.
Rodd, J. M., Davis, M. H., Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb Cortex*, *15*, 1261–1269.
Ross, L. A., Molholm, S., Blanco, D., et al. (2011). The development of multisensory speech perception continues into the late childhood years. *Eur J Neurosci*, *33*, 2329–2337.
Rouger, J., Fraysse, B., Deguine, O., et al. (2008). McGurk effects in cochlear-implanted deaf subjects. *Brain Res*, *1188*, 87–99.
Rouger, J., Lagleyre, S., Fraysse, B., et al. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc Natl Acad Sci U S A*, *104*, 7295–7300.
Royal, D. W., Carriere, B. N., Wallace, M. T. (2009). Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Exp Brain Res*, *198*, 127–136.
Saliba, J., Bortfeld, H., Levitin, D. J., et al. (2016). Functional near-infrared spectroscopy for neuroimaging in cochlear implant recipients. *Hear Res*, *338*, 64–75.
Sandmann, P., Dillier, N., Eichele, T., et al. (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain*, *135*(Pt 2), 555–568.
Scarborough, R., Keating, P., Mattys, S. L., et al. (2009). Optical phonetics and visual perception of lexical and phrasal stress in English. *Lang Speech*, *52*(pt 2-3), 135–175.
Schorr, E. A., Fox, N. A., van Wassenhove, V., et al. (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proc Natl Acad Sci U S A*, *102*, 18748–18750.
Schwartz, J. L. (2010). A reanalysis of McGurk data suggests that audiovisual fusion in speech perception is subject-dependent. *J Acoust Soc Am*, *127*, 1584–1594.
Sekiyama, K., & Burnham, D. (2008). Impact of language on development of auditory-visual speech perception. *Dev Sci*, *11*, 306–320.
Sekiyama, K., Kanno, I., Miura, S., et al. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neurosci Res*, *47*, 277–287.
Senkowski, D., Saint-Amour, D., Höfle, M., et al. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage*, *56*, 2200–2208.
Shams, L., Kamitani, Y., Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, *408*, 788.
Sharma, A., Campbell, J., Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: evidence from the P1 and N1 event related potentials in cochlear implanted children. *Int J Psychophysiol*, *95*, 135–144.
Sharma, A., Dorman, M. F., Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear*, *23*, 532–539.
Sharma, A., Nash, A. A., Dorman, M. (2009). Cortical development, plasticity and re-organization in children with cochlear implants. *J Commun Disord*, *42*, 272–279.
Sheffield, B. M., Schuchman, G., Bernstein, J. G. (2015). Trimodal speech perception: how residual acoustic hearing supplements cochlear-implant consonant recognition in the presence of visual cues. *Ear Hear*, *36*, e99–e112.
Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., et al. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain Lang*, *101*, 260–277.
Skipper, J. I., Nusbaum, H. C., Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage*, *25*, 76–89.
Song, J. J., Lee, H. J., Kang, H., et al. (2015). Effects of congruent and incongruent visual cues on speech perception and brain activity in cochlear implant users. *Brain Struct Funct*, *220*, 1109–1125.
Stein, B., & Meredith, M. A. (1993). *The Merging of the Senses*. Boston, MA: MIT Press.
Stein, B. E., Huneycutt, W. S., Meredith, M. A. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain Res*, *448*, 355–358.
Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J Cogn Neurosci*, *19*, 1964–1973.
Stekelenburg, J. J., Vroomen, J., de Gelder, B. (2004). Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neurosci Lett*, *357*, 163–166.
Stevenson, R. A., & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage*, *44*, 1210–1223.
Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Exp Brain Res*, *227*, 249–261.
Stevenson, R. A., Altieri, N. A., Kim, S., et al. (2010). Neural processing of asynchronous audiovisual speech perception. *Neuroimage*, *49*, 3308–3318.
Stevenson, R. A., Baum, S. H., Segers, M., Ferber, S., Barense, M. D., & Wallace, M. T. (2017). Multisensory speech perception in autism spectrum disorder: From phoneme to whole-word perception. *Autism Res*, DOI: 10.1002/aur.1776.
Stevenson, R. A., Bushmakin, M., Kim, S., et al. (2012a). Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topogr*, *25*, 308–326.
Stevenson, R. A., Fister, J. K., Barnett, Z. P., et al. (2012b). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Exp Brain Res*, *219*, 121–137.
Stevenson, R. A., Geoghegan, M. L., James, T. W. (2007). Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Exp Brain Res*, *179*, 85–95.
Stevenson, R. A., Ghose, D., Fister, J. K., et al. (2014a). Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr*, *27*, 707–730.
Stevenson, R. A., Kim, S., James, T. W. (2009). An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Exp Brain Res*, *198*, 183–194.
Stevenson, R. A., Park, S., Cochran, C., et al. (2017). The associations between multisensory temporal processing and symptoms of schizophrenia. *Schizophr Res*, *179*, 97–103.
Stevenson, R. A., Segers, M., Ferber, S., et al. (2014b). The impact of multisensory integration deficits on speech perception in children with autism spectrum disorders. *Front Psychol*, *5*, 379.

Stevenson, R. A., Segers, M., Ferber, S., et al. (2016). Keeping time in the brain: Autism spectrum disorder and audiovisual temporal processing. *Autism Res*, *9*, 720–738.
Stevenson, R. A., Siemann, J. K., Schneider, B. C., et al. (2014c). Multisensory temporal integration in autism spectrum disorders. *J Neurosci*, *34*, 691–697.
Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., et al. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *Neuroimage*, *55*, 1339–1345.
Stevenson, R. A., Wallace, M. T., Altieri, N. (2014d). The interaction between stimulus factors and cognitive factors during multisensory integration of audiovisual speech. *Front Psychol*, *5*, 352.
Stevenson, R. A., Zemtsov, R. K., Wallace, M. T. (2012c). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *J Exp Psychol Hum Percept Perform*, *38*, 1517–1529.
Strelnikov, K., Marx, M., Lagleyre, S., et al. (2015a). PET-imaging of brain plasticity after cochlear implantation. *Hear Res*, *322*, 180–187.
Strelnikov, K., Rouger, J., Barone, P., et al. (2009). Role of speechreading in audiovisual interactions during the recovery of speech comprehension in deaf adults with cochlear implants. *Scand J Psychol*, *50*, 437–444.
Strelnikov, K., Rouger, J., Demonet, J. F., et al. (2013). Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain*, *136*(Pt 12), 3682–3695.
Strelnikov, K., Rouger, J., Lagleyre, S., et al. (2015b). Increased audiovisual integration in cochlear-implanted deaf patients: independent components analysis of longitudinal positron emission tomography data. *Eur J Neurosci*, *41*, 677–685.
Stropahl, M., Chen, L. C., Debener, S. (2017). Cortical reorganization in postlingually deaf cochlear implant users: Intra-modal and cross-modal considerations. *Hear Res*, *343*, 128–137.
Stropahl, M., Plotz, K., Schönfeld, R., et al. (2015a). Cross-modal reorganization in cochlear implant users: Auditory cortex contributes to visual face processing. *Neuroimage*, *121*, 159–170.
Stropahl, M., Schellhardt, S., Debener, S. (2015b). McGurk stimuli for the investigation of multisensory integration in cochlear implant users: The Oldenburg Audio Visual Speech Stimuli (OLAVS). *Psychonomic Bull Rev*, 1–10.
Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J Acoust Soc Am*, *26*, 212–215.
Summerfield, A. Q., Marshall, D. H., Barton, G. R., et al. (2002). A cost-utility scenario analysis of bilateral cochlear implantation. *Arch Otolaryngol Head Neck Surg*, *128*, 1255–1262.
Swerts, M., Krahmer, E. (2005). Cognitive processing of audiovisual cues to prominence. In *AVSP* (pp. 29–30).
Tona, R., Naito, Y., Moroto, S., et al. (2015). Audio-visual integration during speech perception in prelingually deafened Japanese children revealed by the McGurk effect. *Int J Pediatr Otorhinolaryngol*, *79*, 2072–2078.
Tremblay, C., Champoux, F., Lepore, F., et al. (2010). Audiovisual fusion and cochlear implant proficiency. *Restor Neurol Neurosci*, *28*, 283–291.
Tye-Murray, N., Sommers, M. S., Spehar, B. (2007). Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. *Ear Hear*, *28*, 656–668.
Tyler, R., Parkinson, A. J., Fryauf-Bertchy, H., et al. (1997). Speech perception by prelingually deaf children and postlingually deaf adults with cochlear implant. *Scand Audiol Suppl*, *46*, 65–71.
Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Curr Opin Neurobiol*, *4*, 157–165.
Vachon, P., Voss, P., Lassonde, M., et al. (2013). Reorganization of the auditory, visual and multimodal areas in early deaf individuals. *Neuroscience*, *245*, 50–60.
van Atteveldt, N. M., Formisano, E., Blomert, L., et al. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb Cortex*, *17*, 962–974.
van Dijk, J. E., van Olphen, A. F., Langereis, M. C., et al. (1999). Predictors of cochlear implant performance. *Audiology*, *38*, 109–116.
van Hoesel, R. J. (2015). Audio-visual speech intelligibility benefits with bilateral cochlear implants when talker location varies. *J Assoc Res Otolaryngol*, *16*, 309–315.
van Wassenhove, V., Grant, K. W., Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*, 598–607.
Walden, B. E., Grant, K. W., Cord, M. T. (2001). Effects of amplification and speechreading on consonant recognition by persons with impaired hearing. *Ear Hear*, *22*, 333–341.
Walden, B. E., Prosek, R. A., Worthington, D. W. (1974). Predicting audiovisual consonant recognition performance of hearing-impaired adults. *J Speech Hear Res*, *17*, 270–278.
Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *J Neurophysiol*, *97*, 921–926.
Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105–123.
Wallace, M. T., Perrault, T. J. Jr, Hairston, W. D., et al. (2004). Visual experience is necessary for the development of multisensory integration. *J Neurosci*, *24*, 9580–9584.
Werner, S., & Noppeney, U. (2010). Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cereb Cortex*, *20*, 1829–1842.
Wightman, F., Kistler, D., Brungart, D. (2006). Informational masking of speech in children: auditory-visual integration. *J Acoust Soc Am*, *119*, 3940–3949.
Williams, J. H., Massaro, D. W., Peel, N. J., et al. (2004). Visual-auditory integration during speech imitation in autism. *Res Dev Disabil*, *25*, 559–575.
Woynaroski, T. G., Kwakye, L. D., Foss-Feig, J. H., et al. (2013). Multisensory speech perception in children with autism spectrum disorders. *J Autism Dev Disord*, *43*, 2891–2902.
Wright, T. M., Pelphrey, K. A., Allison, T., et al. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cereb Cortex*, *13*, 1034–1043.
Xu, J., Yu, L., Rowland, B. A., et al. (2014). Noise-rearing disrupts the maturation of multisensory integration. *Eur J Neurosci*, *39*, 602–613.
Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., et al. (2006). Top-down and bottom-up processes in speech comprehension. *Neuroimage*, *32*, 1826–1836.