



Audiovisual benefits for speech processing speed among children with hearing loss

Rebecca Holt¹, Laurence Bruggeman^{1,2}, Katherine Demuth¹

¹Department of Linguistics and ARC Centre of Excellence in Cognition and its Disorders,
Macquarie University, Australia

²MARCS Institute for Brain, Behaviour and Development and ARC Centre of Excellence for the
Dynamics of Language, Western Sydney University, Australia

rebecca.holt@mq.edu.au, L.Bruggeman@westernsydney.edu.au,
katherine.demuth@mq.edu.au

Abstract

Children with hearing loss face a range of challenges when listening to and processing speech; in particular, they may process spoken language slowly in comparison to normal-hearing peers [1]. How then can speech processing speed be improved for children with hearing loss? In this study, a phoneme monitoring task was used to assess whether 7-11-year-old children with hearing loss showed faster speech processing when visual speech cues were available compared to auditory-only presentation. Children with hearing loss did receive an audiovisual benefit for processing speed, however this was primarily driven by cases in which the target phoneme in the monitoring task was visually salient. No difference was found between the performance of the children with hearing loss and a control group of children with normal hearing, however the results suggest that children with hearing loss who use hearing aids may receive a greater audiovisual benefit for processing speed than those who use cochlear implants. These findings have implications for practical interventions for children with hearing loss.

Index Terms: audiovisual benefit, processing speed, children, hearing loss, cochlear implants, hearing aids, phoneme monitoring

1. Introduction

It has been reported that children with hearing loss (HL) typically process speech more slowly than their peers with normal hearing (NH) [1,2]. Slow speech processing can hinder children's ability to keep up with what is said [3], which may in turn affect their academic achievement and wellbeing [4,5]. Here, we therefore investigate a potential way to make speech processing faster for children with HL.

For children with normal hearing (NH), presenting visual speech cues (e.g., movements of the speaker's face) concurrently with the auditory signal (audio-visual presentation; AV) has been shown to facilitate speech processing, conferring an "AV benefit" over speech presented in auditory modality alone (auditory-only presentation; AO). AV benefits are typically quantified by examining changes in speech perception *accuracy*: Children with NH are more accurate at identifying phonemes [6] and words [7] in noisy conditions or degraded speech when visual cues are provided. However, visual cues have also been found to benefit children's processing *speed* and *effort*: Children with NH process speech faster and exert less processing effort when visual cues are

present [8]. Furthermore, such speed and effort benefits have been found in both quiet and noisy listening conditions [8], suggesting that AV presentation is able to robustly facilitate many aspects of speech processing across different acoustic environments.

We therefore aimed to determine whether AV presentation also improves speech processing speed for children with HL. Among children with HL, AV benefits for *accuracy* have previously been observed [9,10], but AV benefits for *speed* have not been investigated. We hypothesised that, like for children with NH [8], speech processing would be faster in AV than AO modality for children with HL. However, AV benefits for processing *effort* have been found primarily among children with strong phonological awareness skills [8]. If this pattern were to extend to AV benefits for processing *speed*, we may not observe an AV benefit for speed among children with HL, as their phonological awareness skills are typically less well-developed than those of their NH peers [11].

We also aimed to address two secondary research questions. We asked whether the degree of AV benefit for speed shown by children with HL would differ from that shown by children with NH. Previous studies have not typically compared the degree of AV benefit shown by children with NH and HL, primarily because AV benefits for *accuracy* have been the focus of investigation. To observe AV benefits for accuracy among children with NH, it is generally necessary to present stimuli in noise to prevent ceiling performance, while stimuli presented to children with HL are usually presented in quiet to avoid floor performance. These contrasting listening conditions prevent direct comparisons (see, for example, [10]). However, the magnitude of AV benefits for accuracy received by *older adults* with NH and HL have been directly compared in several studies, with inconsistent results. [12] suggests that older adults with HL show greater AV benefits than older adults with NH, but [13] found no difference in their degree of AV benefit using a different task. We hypothesised that children with HL would show a greater AV benefit for speed than children with NH, as they have more to gain from visual speech cues than children with NH, due to their slower AO processing [1]. However, it could also be that, like for the older adults in [13], the degree of AV benefit might not differ across HL and NH groups.

We also asked whether children who use different types of devices to remedy their HL would benefit from visual cues to different extents. Typically, children with HL who use oral/aural communication (rather than signed communication) use cochlear implants (CIs) and/or hearing aids (HAs) to remediate their hearing. These devices function in very

different ways and consequently are appropriate for children with different HL diagnoses. CIs are composed of an electrode array implanted in the recipient’s cochlea, and an external sound processor. The processor converts the incoming acoustic signal into a digital signal, which is transmitted to the internal implant, which then converts the digital signal to electrical impulses to stimulate the auditory nerve. The outcome of implantation is that CI users typically lose any residual acoustic hearing, and so this device type is recommended only for those with severe to profound HL. In contrast, HAs rely on the recipient’s residual hearing, amplifying the incoming acoustic signal to a level that can be perceived by the user. HAs are therefore typically used by those with only mild to moderate HL (see [14] for further details). CIs and HAs therefore provide very different types of input: one providing electrical input and the other amplifying the acoustic input. Furthermore, the input available to CI and HA users when they are *not* using their device also differs: CI users receive no auditory input when their device is not in use (and thus typically use their device consistently and for most waking hours), while HA users are still able to access auditory input, albeit degraded, without their devices, resulting in potentially inconsistent use, especially among children and those with milder HL [15].

Thus, substantial differences exist between CI and HA users, both in terms of the types of input provided by their respective devices (electric vs. acoustic signals, absent vs. present residual hearing), and the characteristics of the users themselves (e.g., severe/profound vs. mild/moderate HL, consistent vs. potentially inconsistent device users). We therefore hypothesised that these groups may differ in the extent to which they benefit from visual speech cues. Potentially, children who use CIs may show greater AV benefits due to their more severe HL: They may rely on visual information more to compensate for their more degraded auditory input. Alternatively, children who use HAs may show greater AV benefits as they make use of their residual hearing: At times when HA users are not using their devices, they may rely heavily on visual speech cues, leading to improved AV integration skill as they supplement their auditory input with visual cues.

To answer these questions, we employed a phoneme monitoring task identical to that used in [8]. Participants with HL, aged between 7 and 11 years old, heard sentences presented in either AV or AO modality and were instructed to make a button-press as quickly as possible when they heard pre-specified target phonemes. Reaction times (RTs) in this task were used as a measure of speech processing speed. The study had a (2)x2 design, with two levels of presentation modality (AV vs. AO) and two groups (children with HL vs. children with NH). Data from a control group of children with NH were drawn from [8]. This enabled us to address our first two research questions (do children with HL show an AV benefit for speed, and does this benefit differ from the benefit shown by children with NH?). The third question (do children who use CIs and HAs differ in their degree of AV benefit for processing speed?) was addressed via an additional comparison between CI and HA users following the main analysis. Here we present results from a preliminary sample of 12 children with HL, however data collection is ongoing.

2. Methods

2.1. Participants

Twelve children with HL aged 7-11 years participated in the study ($M_{age} = 9$ years, 6 months, $SD = 1$ year, 5 months; 10M, 2F). Six were bilateral CI recipients and six used HAs. Several types of HA were represented in the HA group: Four participants used bilateral behind-the-ear HAs, one used a bone-anchored hearing aid (BAHA) and one used a contralateral routing of signals (CROS) aid. Further characteristics of participants with HL are shown in Table 1. The control group of 19 children with NH were also aged 7-11 years old ($M_{age} = 8$ years, 11 months, $SD = 1$ year, 5 months; 10M, 9F). All participants (NH and HL) were monolingual native speakers of English and none had any reported language or cognitive impairment, or any vision impairment not corrected by glasses. Parental written informed consent was obtained for all participants.

Table 1: *Characteristics of participants with HL*

Age	Gender	Device type	Severity of HL
8;6	M	HAs	Moderate
8;7	M	BAHA	Moderate (unilateral)
8;7	M	HAs	Moderate
8;9	M	CROS	Profound (unilateral)
10;7	M	HAs	Moderate/severe
11;8	F	HAs	Mild/moderate
7;2	F	CIs	Profound
9;5	M	CIs	Profound
9;11	M	CIs	Profound
10;2	M	CIs	Profound
10;6	M	CIs	Profound
11;9	M	CIs	Profound

2.2. Stimuli

The stimuli used in this study were adapted from [1] and were the same as used in [8] (see Table 2 for examples). The target phonemes were /b, p, g, k/. Each occurred as a word-initial singleton in twelve target words. Manner of articulation, word position, and singleton vs. cluster effects on phoneme monitoring RT were thus controlled [16,17,18]. All target words were familiar to children, with a log frequency greater than 3.00 in the CBBC section of the SUBTLEX database [19]. Each target word occurred in a sentence of 9-12 syllables that did not contain any other occurrences of the target phoneme. The target phoneme always occurred in the fifth or sixth syllable of the sentence, to control for sentence-position effects on RT [18]. When recorded, each sentence was produced with the target word prosodically focused, to facilitate phoneme monitoring and control for prosodic effects on RT [16]. Sixteen catch sentences were also created, containing no instances of any target phonemes but matching the test sentences in length and prosodic structure. As the position of the target phoneme did not vary across the test sentences, the catch sentences were included to prevent participants from automatically responding at the same position in each sentence without paying attention.

All test and catch sentences were spoken by a female native speaker of Australian English and recorded using a Sony HXR-NX30P digital HD video camera with a Sony ECM-XMI electret condenser microphone. The speaker’s whole face and shoulders were visible and centred in the recording. The

speaker wore a black t-shirt and was shown in front of a solid grey background. Video recordings were segmented and the audio tracks extracted. The mean intensity of each sentence was normalised in Praat [20]. For the AV stimuli, the audio and video tracks were recombined. For the AO stimuli, the audio tracks were matched with a static frame taken from an AV stimulus in which the speaker looked at the camera with her mouth closed and had a smiling expression. Each stimulus sentence appeared in the AV condition for half the participants and AO for the other half.

Table 2: Sample sentences from the phoneme monitoring task (target phoneme in bold)

Target phoneme	Sample sentence
/b/	Grandma gave a bone to the puppy.
/p/	The boy used a pen to write his name.
/g/	The farmer fed the geese in the afternoon.
/k/	The tourist loved the castle in the valley.

2.2. Procedure

2.2.1. Discrimination task

The test session began with a discrimination task, adapted from [1]. This was to confirm that all participants were able to successfully discriminate the phonemes /b, p, g, k/, so that they could perform the phoneme monitoring task. In each trial, participants were presented with two images that formed a minimal pair differing in onset (e.g., bear and pear, goat and coat). They then heard the name of one of those images in the carrier sentence “I said the word [target] again”. Participants were instructed to select which image matched the sentence. Thirty-six trials were presented, comparing each pair of target phonemes six times. Both children with NH and children with HL were able to discriminate all phonemes. Children with NH from [8] had a mean score of 100% correct ($SD = 0.9\%$, range = 97%-100%), while children with HL scored 98% correct on average ($SD = 4\%$, range = 86%-100%).

2.2.2. Simple reaction time task

Participants then completed a simple reaction time task, training them to make button-press responses as quickly as possible. This was intended to mitigate issues such as those found in [6], where 5-10-year-old children were unable to provide sufficiently consistent RTs in a phoneme monitoring task. A white fixation cross was presented on a black screen at irregular intervals and participants were instructed to press a button on the response pad as quickly as possible whenever they saw the cross. This rapid button-press in response to a stimulus was the same response as required in the phoneme monitoring task.

2.2.3. Phoneme monitoring task

In the phoneme monitoring task, eight blocks, plus a block of practice trials, were presented in total: four in AV modality (one per target phoneme) and four in AO (again, one per target phoneme). For half the participants, AV blocks were presented first, followed by AO blocks, while the other half received the reverse order. Each block was structured as follows. An introductory sentence first informed participants which phoneme they were required to monitor for the duration of that block. Introductory sentences took the form “Listen for the

[phoneme], as in [sample word] and [sample word].” Then, eight trials were presented in random order, consisting of six test sentences and two catch sentences. Participants were instructed to make a button-press whenever they heard the target phoneme specified in the introductory sentence, but to make no response if the target phoneme was absent. Within each block, all sentences were presented in the same modality (either AV or AO), including the introductory sentence. At the end of each block, a short video showing the speaker pulling a silly face was shown to maintain participants’ attention.

All three tasks were presented in E-Prime 3 on an Alienware P69F gaming laptop computer with a Cedrus RB-840 response pad. Audio was presented via a Genelec 8020C external speaker.

3. Analysis and results

3.1. HL vs. NH comparison

The dataset for RT analysis was assembled by compiling data from the children with HL with the data from the children with NH taken from [8]. All catch trials and incorrect responses were excluded. Responses with unfeasibly short (less than 100 ms) or long (greater than 3000 ms) RTs were then removed [1]. A generalised linear mixed-effects model was fit to the data using the *lme4* package [21] in R [22], with inverse Gaussian family, identity link function, and untransformed RT as the dependent variable [23]. The fixed factors were Modality (AV vs. AO), Group (NH vs. HL) and place of articulation (PoA; bilabial vs. velar). The PoA factor, contrasting bilabial and velar target phonemes, was included to determine whether any effect of modality would extend to both visually-salient (bilabial) and non-visually-salient (velar) targets [8]. The model used maximal random effects, i.e., random intercepts for Participant and Item and random slopes for Modality and PoA by Participant and Modality and Group by Item, and the nAGQ argument was set to zero to facilitate model convergence. The model syntax used was: *glmer*($RT \sim \text{Modality} * \text{Group} * \text{PoA} + (1 + \text{Modality} + \text{PoA} / \text{Participant}) + (1 + \text{Modality} + \text{Group} / \text{Item})$, *data* = *RTData*, *family* = *inverse.gaussian*(*link* = “identity”), *nAGQ* = 0)). Outlying values were removed after model fitting by excluding observations that had standardised residuals more than 2.5 standard deviations from the mean [24]. The model was then re-fit and tested for statistical significance.

A significant effect of Modality was found ($\beta = -26.35$, $SE = 7.11$, $p < .001$): Overall, participants responded 26 ms faster to AV than AO stimuli. This was qualified by a significant interaction between Modality and PoA ($\beta = 12.10$, $SE = 5.12$, $p = .02$; Figure 1). Post-hoc pairwise comparisons on this interaction using the *emmeans* package [25] revealed that there was a significant difference in RT between the two modalities only when the target phoneme was bilabial ($p < .001$), but not when the target was velar ($p = .10$). Nevertheless, estimated marginal mean RTs to AV stimuli were faster than to AO stimuli for both PoAs (bilabial: 77 ms faster; velar: 29 ms faster). No significant effect or interaction of Group was found (all $p \geq .11$; Figure 2).

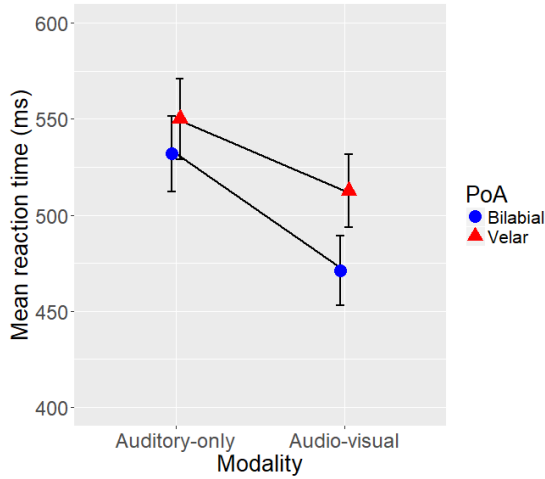


Figure 1: Mean RTs across AO and AV modalities for bilabial (blue circles) and velar (red triangles) PoAs, collapsed across HL and NH groups. Error bars show one standard error of the mean.

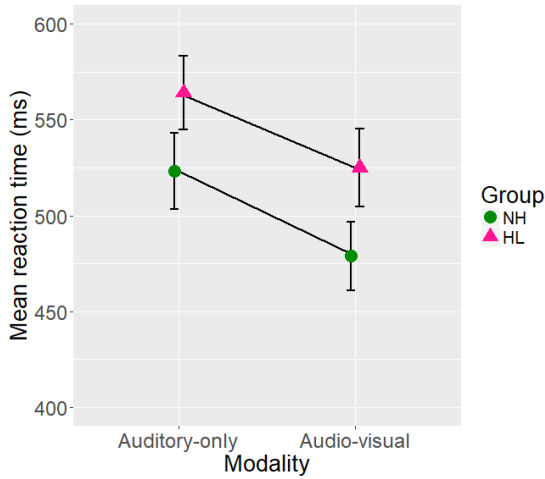


Figure 2: Mean RTs across AO and AV modalities for children with NH (green circles) and HL (pink triangles), collapsed across PoAs. Error bars show one standard error of the mean.

3.2. CI vs. HA comparison

A secondary analysis was conducted on responses provided by the HL group only, comparing the performance of CI and HA users. Data preparation, model fit, and outlier exclusion followed the same procedure as previously. In this case the fixed factors in the model were Modality (AV vs. AO) and Device (CI vs. HA). As previously, the random effects were maximal, including random intercepts for Participant and Item and random slopes for Modality by Participant and Modality and Device by Item, and $nAGQ$ was set to zero. The model syntax was: `glmer(RT ~ Modality * Device + (1 + Modality / Participant) + (1 + Modality + Device / Item), data = RTDataHLonly, family = inverse.gaussian(link = "identity"), nAGQ = 0)`.

This model also revealed a main effect of Modality ($\beta = -30.40$, $SE = 13.33$, $p = .02$): Participants with HL responded

on average 30 ms faster to AV than to AO stimuli. An interaction between Modality and Device approached significance ($\beta = -19.59$, $SE = 11.16$, $p = .08$; Figure 3), suggesting that HA users may show a greater AV benefit than CI users. It is possible that this effect did not reach significance due to limited statistical power. Ongoing data collection will increase the power available and may reveal a significant effect in future. No other comparisons were significant (all $p \geq .32$).

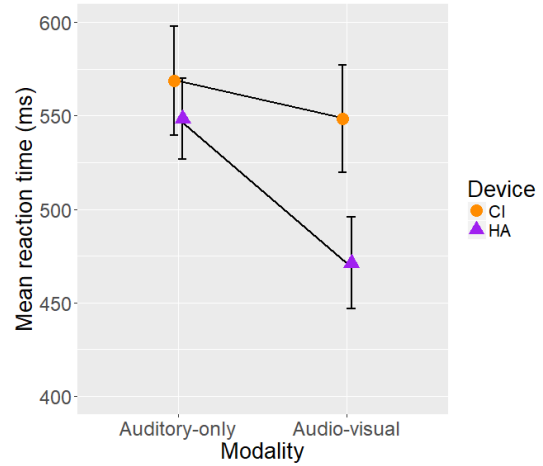


Figure 3: Mean RTs across AO and AV modalities for children with HL who use CIs (orange circles) and HAs (purple triangles). Error bars show one standard error of the mean.

4. Discussion

This study used a phoneme monitoring task with AV and AO stimuli to address three questions. First, do children with HL show an AV benefit for processing speed? Second, how does this AV benefit compare to the AV benefit for speed shown by children with NH? And third, do children with HL who use CIs and HAs differ in the size of their AV benefit?

Our results indicate that children with HL do indeed show an AV benefit for processing speed: They detected phonemes in sentences faster when these were presented audiovisually rather than in auditory modality alone. It seems that the difficulties with phonological awareness typically experienced by children with HL [11] do not prevent them from receiving this AV benefit. However, when PoA was taken into account, the AV benefit was only statistically significant when the target phoneme was bilabial, and therefore visually salient, but not when it was velar. This discrepancy between bilabial and velar PoAs is unsurprising because, when an oral stop consonant is produced, the (audible) release burst is preceded by a (silent) period of closure. In the case of a bilabial stop, the closure period is visible (lip closure) and thus available earlier than the auditory cue (release burst). This is not the case for velar stops, which do not have a salient visual cue. Therefore, if participants make use of visual cues during the task, a cue to the presence of an upcoming bilabial stop is present earlier in AV modality than AO, allowing participants to respond faster in this condition. One could therefore contend that the relative timing of visual and auditory cues may be solely responsible for the AV benefit observed, rather than any change in processing speed. However, we take this result as evidence suggesting children with HL can indeed use visual cues to speed up their

processing, regardless of the source of such benefit (processing changes or cue timing advantages). Additionally, there is a (non-significant) trend towards an AV benefit for velar phonemes as well. As mentioned previously, data collection for this study is ongoing, and it is possible that we may see a significant AV benefit for velar target phonemes as well once statistical power is increased. This would provide even stronger evidence that the presence of visual cues is able to improve processing speed for these children, as any AV benefit for velar target phonemes must then be due to overall improvements in processing speed, rather than to a temporal discrepancy between auditory and visual cues.

Regarding our second research question, we found no evidence to suggest that children with HL and NH differ in the magnitude of their AV benefits for processing speed: No significant effect of or interaction with Group was found. Contrary to our expectations, our results are thus in line with those of [13] rather than [12]: Children with HL do not seem to show any extra AV benefit over that shown by children with NH, despite their poorer access to auditory information and greater room for improvement in processing speed. Interestingly, children with HL were not significantly slower to respond than children with NH overall, challenging the notion that children with HL suffer from slow speech processing compared to their NH peers [1]. However, we believe this may at least in part be attributed to the relative familiarity of children with HL with language testing and assessment. These children are likely to have completed many more language assessments and auditory training tasks than their peers with NH, and so may have had an advantage due to their familiarity with the testing environment. Additionally, the stimuli used in our study were designed to make phoneme monitoring as easy as possible, with consistent target positioning and facilitative prosody. This contrasts with the more difficult stimuli used in [1]. Differences in processing speed between children with HL and NH may have been increased in [1] by the use of more challenging material, which may have disproportionately affected the performance of the children with HL.

Our final research question asked whether children with HL who use different types of hearing devices differ in the magnitude of AV benefit for processing speed. We observed a trend in this direction that approached significance: Children who used HAs appeared to benefit more from the addition of visual cues than children who used CIs. We attribute this to HA users' greater experience with circumstances where the use of visual cues may be essential for speech understanding (i.e., at times when their HAs are not in use). We cannot draw a strong conclusion from this result as the effect does not reach significance, however as more data is collected in future, statistical power will be increased, allowing us to explore this interaction more fully. Furthermore, we cannot be certain of which characteristics of CI and HA users are responsible for any difference between these groups. As mentioned previously, CI and HA users differ systematically on several factors aside from their device type. Further research is required to conclude whether one of these additional factors may be responsible.

These findings have practical implications for those who interact with children with HL. Our results suggest that in circumstances where speech processing speed is of importance, such as in the classroom, ensuring that visual speech cues are available will benefit children with HL. This may involve seating children with HL close to the teacher, and ensuring that the teacher faces the class as much as possible while they speak. There is also potential for the development of interventions for

children with HL to improve their processing speed based on our findings.

It is important to note that the population of children with HL is highly heterogeneous, so the effects found in this study may be diluted by extreme variability in performance. In addition, AV benefits in this study were measured using a phonological task, which is unlikely to be representative of the types of task children with HL must perform when processing speech in everyday life. In future, AV benefits for processing speed may be examined using a comprehension-based task, which is likely more representative of real language use.

5. Conclusions

Using a phoneme monitoring task, this study demonstrates that 7-11-year-old children with HL can take advantage of visual speech cues to improve their speed of processing. However, we find no indication that children with HL receive an AV benefit that differs from that of children with NH. Furthermore, there is a suggestion that children who use CIs and HAs may differ in the amount of processing speed benefit they derive from visual cues. These findings have implications for facilitating language processing among children with HL.

6. Acknowledgments

We thank Isabel O'Keeffe and Tracy Hopkins for assistance with participant recruitment. This project was supported by the Australian Research Council (ARC FL130100014; ARC CE110001021) and Macquarie University (MQ RIS83034101).

7. References

- [1] C. M. Holt, K. Demuth, and I. Yuen, "The use of prosodic cues in sentence processing by prelingually deaf users of cochlear implants", *Ear and Hearing*, vol. 37, pp. e256-e262, Jul. 2016.
- [2] D. Pisoni, W. Kronenberger, A. Roman, and A. Geers, "Article 7: Measures of digit span and verbal rehearsal speed in deaf children following more than 10 years of cochlear implantation," *Ear and Hearing*, vol. 32, pp. 60s-74s, Feb. 2011.
- [3] K. T. Mealings, K. Demuth, J. M. Buchholz, and H. Dillon, "The effect of different open plan and enclosed classroom acoustic conditions on speech perception in Kindergarten children," *The Journal of the Acoustical Society of America*, vol. 138, pp. 2458-2469, Oct. 2015.
- [4] C. C. Crandell and J. J. Smaldino, "Classroom acoustics for children with normal hearing and with hearing impairment," *Language, Speech, and Hearing Services in Schools*, vol. 31, pp. 362-370, Oct. 2000.
- [5] K. Kuppler, M. Lewis, and A. K. Evans, "A review of unilateral hearing loss and academic performance: Is it time to reassess traditional dogmata?" *International Journal of Pediatric Otorhinolaryngology*, vol. 77, pp. 617-622, May 2013.
- [6] M. Fort, E. Spinelli, C. Savariaux, and S. Kandel, "Audiovisual vowel monitoring and the word superiority effect in children," *International Journal of Behavioral Development*, vol. 36, pp. 457-467, Nov. 2012.
- [7] D. W. Maidment, H. J. Kang, H. J. Stewart, and S. Amitay, "Audiovisual integration in children listening to spectrally degraded speech," *Journal of Speech, Language, and Hearing Research*, vol. 58, pp. 61-68, Feb. 2015.
- [8] R. Holt, L. Bruggeman, and K. Demuth, "Visual speech cues speed processing and reduce effort for children listening in quiet and noise," under review.
- [9] K. I. Kirk, M. J. Hay-McCutcheon, R. F. Holt, S. Gao, R. Qi, and B. L. Gerlain, "Audiovisual spoken word recognition by children with cochlear implants," *Audiological Medicine*, vol. 5, pp. 250-261, Jan. 2007.

- [10] R. F. Holt, K. I. Kirk, M. Hay-McCutcheon, "Assessing multimodal spoken word-in-sentence recognition in children with normal hearing and children with cochlear implants," *Journal of Speech, Language, and Hearing Research*, vol. 54, pp. 632-657, Apr. 2011.
- [11] M. P. Moeller, J. B. Tomblin, C. Yoshinaga-Itano, C. M. Connor, and S. Jerger, "Current state of knowledge: Language and literacy of children with hearing impairment," *Ear and Hearing*, vol. 28, pp. 740-753, Dec. 2007.
- [12] S. Moradi, B. Lidestam, and J. Rönnberg, "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*, vol. 20, pp. 1-15, Jun. 2016.
- [13] N. Tye-Murray, M. S. Sommers, and B. Spehar, "Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing," *Ear and Hearing*, vol. 28, pp. 656-668, Sep. 2007.
- [14] T. Foust and J. Hoffman, "Audiology 101: an introduction to audiology for nonaudiologists," in *The NCHAM E-Book: A Resource Guide for Early Hearing Detection and Intervention*. L. R. Schmeltz, Ed. 2019, pp. 5.1-5.16. Accessed on: May 9, 2019. [Online]. Available: <http://www.infantheating.org/ehdi-ebook/>
- [15] E. A. Walker *et al.*, "Predictors of hearing aid use time in children with mild-to-severe hearing loss," *Language, Speech, and Hearing Services in Schools*, vol. 44, pp. 73-88, Jan. 2013.
- [16] A. Cutler and D. Norris, "Monitoring sentence comprehension", in *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, M. F. Garret, W. E. Cooper, and E. C. T. Walker, Eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979, pp. 113-134.
- [17] R. Trieman, A. Salasoo, L. M. Slowiaczek, and D. B. Pisoni, "Effects of syllable structure on adults' phoneme monitoring performance", in *Research on Speech Perception Progress Report no. 8*, D. B. Pisoni, Ed. Bloomington, IN: Indiana University Speech Research Laboratory, 1982, pp. 63-82.
- [18] J. Sasisekaran and C. Weber-Fox, "Cross-sectional study of phoneme and rhyme monitoring abilities in children between 7 and 13 years," *Applied Psycholinguistics*, vol. 33, pp. 253-279, Apr. 2012.
- [19] W. J. Van Heuven, P. Mander, E. Keuleers, and M. Brysbaert, "SUBTLEX-UK: A new and improved word frequency database for British English," *The Quarterly Journal of Experimental Psychology*, vol. 67, pp. 1176-1190, Jun. 2014.
- [20] P. Boersma and D. Weenink, Amsterdam. 2019. *Praat: Doing phonetics by computer* [Online]. Available: <http://www.fon.hum.uva.nl/praat/>
- [21] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 69, pp. 1-33, 2015.
- [22] R Core Team, Vienna. 2018. *R: a language and environment for statistical computing* [Online]. Available: <https://cran.r-project.org/>
- [23] S. Lo and S. Andrews, "To transform or not to transform: using generalized linear mixed models to analyse reaction time data," *Frontiers in Psychology*, vol. 6, 1171, 2015.
- [24] R. H. Baayen and P. Milin, "Analyzing reaction times," *International Journal of Psychological Research*, vol. 3, pp. 12-28, 2010.
- [25] R. Lenth, H. Singmann, J. Love, P. Buerkner, and M. Herve, 2018. *Estimated marginal means, aka least-squares means* [Online]. Available: <https://cran.r-project.org/web/packages/emmeans/>