JSLHR

Research Article

Visual Speech Perception in Children With Language Learning Impairments

Victoria C. P. Knowland,[a] Sam Evans,[b] Caroline Snell,[a] and Stuart Rosen[b]

Purpose: The purpose of the study was to assess the ability of children with developmental language learning impairments (LLIs) to use visual speech cues from the talking face.
Method: In this cross-sectional study, 41 typically developing children (mean age: 8 years 0 months, range: 4 years 5 months to 11 years 10 months) and 27 children with diagnosed LLI (mean age: 8 years 10 months, range: 5 years 2 months to 11 years 6 months) completed a silent speechreading task and a speech-in-noise task with and without visual support from the talking face. The speech-in-noise task involved the identification of a target word in a carrier sentence with a single competing speaker as a masker.
Results: Children in the LLI group showed a deficit in speechreading when compared with their typically developing peers. Beyond the single-word level, this deficit became more apparent in older children. On the speech-in-noise task, a substantial benefit of visual cues was found regardless of age or group membership, although the LLI group showed an overall developmental delay in speech perception.
Conclusion: Although children with LLI were less accurate than their peers on the speechreading and speech-in noise-tasks, both groups were able to make equivalent use of visual cues to boost performance accuracy when listening in noise.

Children with developmental language learning impairments show a primary deficit in the acquisition and use of oral language that cannot be explained with recourse to sensory impairments, reduced opportunity to learn, or low nonverbal IQ. These children are variously referred to in the literature as having specific language impairment or language learning difficulties. Here we use the term *language learning impairment* (LLI; Tallal & Benasich, 2002) to reflect a primary deficit in the language domain while acknowledging the considerable difficulties recorded across multiple areas of cognitive development (e.g., Donlan, Cowan, Newton, & Lloyd, 2007; Henry, Messer, & Nash, 2012; Hill, 2001).

There is substantial heterogeneity within the population of children with LLI and, as yet, little understanding of the developmental pathways causally related to atypical language behavior. One fruitful, though controversial, area of research has been that of auditory processing. As a group, children with LLI show poor performance in comparison to their typically developing (TD) peers on multiple auditory processing tasks using both speech and nonspeech stimuli (Corriveau, Pasquini, & Goswami, 2007; McArthur & Bishop, 2004; Rosen, Adlard, & van der Lely, 2009; Wright et al., 1997; and see Rosen, 2003, for a review). Within the speech domain, children with LLI and especially those with mixed expressive and receptive difficulties (Stark & Heinz, 1996a) show greater variability in the placement of phonetic category boundaries (Burlingame, Sussman, Gillam, & Hay, 2005; Sussman, 1993) and are less accurate at identifying steady-state vowels than TD controls (Stark & Heinz, 1996b). Deficits are seen in speech tasks even when working memory demands, known to be an area of weakness in children with LLI (Henry et al., 2012; Marton, 2008), are kept low.

Although speech perception difficulties are evident in children with LLI when tested in optimal listening conditions, the effects of noise seem to exacerbate group differences. This pattern has been observed for syllable identification (Ziegler, Pech-Georgel, George, & Lorenzi, 2011; Ziegler, Pech-Georgel, George, Alario, & Lorenzi, 2005) and discrimination between minimal pairs (Vance & Martindale, 2012). In some cases, group effects have only been found under conditions of noise and not when given a clear auditory signal, for example, during phonetic categorization (Robertson, Joanisse, Desroches, & Ng, 2009) and sentence perception (Bradlow, Kraus, & Hayes, 2003). Furthermore, the perception of speech in noise has been found to predict later receptive language

---

[a]City University London, United Kingdom
[b]University College London, United Kingdom
Correspondence to Victoria C. P. Knowland: victoria.knowland.1@city.ac.uk

Editor: Jody Kreiman
Associate Editor: Megha Sundara
Received September 29, 2014
Revision received May 31, 2015
Accepted July 30, 2015
DOI: 10.1044/2015_JSLHR-S-14-0269

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

scores for children with impairments (Robertson et al., 2009; Vance & Martindale, 2012; Ziegler et al., 2011).

Behavioral studies of deficits in speech perception in children with LLI have been supported with electrophysiological data showing categorization difficulties that may be specific to speech stimuli (Uwer, Albrecht, & von Suchodoletz, 2002) or extend to stimuli with high spectral complexity more generally (McArthur & Bishop, 2005). The processing of speech in noise has been suggested to be particularly challenging for children with deficits in the language domain due to disruptions in the timing of cortical (Warrier, Johnson, Hayes, Nicol, & Kraus, 2004; Wible, Nicol, & Kraus, 2002) and subcortical (Anderson, Skoe, Chandrasekaran, & Kraus, 2010) responses that may result from inaccurate neural phase locking to the onset of auditory stimuli (Cunningham, Nicol, Zecker, Bradlow, & Kraus, 2001). In addition, children with LLI have a deficit in selective attention when presented with speech in noise (Stevens, Sanders, & Neville, 2006), possibly indicating atypical sensory gain in the auditory domain.

As yet, the relationship between possible processing deficits and the broader profile of language atypicalities seen in children with LLI is not clear. Auditory processing does not often account for a large degree of variance in language performance (see Rosen, 2003), and group effects are often driven by a subset of individuals in both behavioral (e.g., Rosen et al., 2009) and electrophysiological (McArthur & Bishop, 2005) studies. Despite this, problems with the auditory perception of speech-relevant stimuli may well be etiologically significant for at least a subgroup of children with unexplained language learning problems.

Audiovisual Speech in Childhood

Speech perception is a multimodal process with visual cues from the talking face correlating tightly with auditory cues (Chandraesekaran, Trubanov, Stilittano, Caplier, & Ghazanfar, 2009) and contributing substantially to adult speech perception, including but not limited to when listening under difficult conditions (Grant & Seitz, 2000; Reisberg, McLean, & Goldfield, 1987; Remez, 2005; Sumby & Pollack, 1954; and see Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007). Sensitivity to the multimodality of speech develops early in infancy. From just 2 months of age, infants respond to congruence between auditory and visual speech tokens (Baier, Idsardi, & Lidz, 2007; Kuhl & Meltzoff, 1982; Patterson & Werker, 2003). By 5 months old, infants are sensitive to the McGurk effect: a cross-modal illusion in which an incongruent visual speech token modulates the perceived identity of an auditory token. For example, an auditory /bɑ/ paired with a visual |gɑ| often results in the reported perception of /dɑ/ or /ðɑ/ (McGurk & MacDonald, 1976). Early sensitivity to this illusion has been shown both behaviorally (Burnham & Dodd, 2004; Rosenblum, Schmuckler, & Johnson, 1997) and with electrophysiological recording (Kushnerenko, Teinonen, Volein, & Csibra, 2008) and is taken to indicate cross-modal integration. At 4–8 months, infants show a heightened pattern of eye gaze to the mouth of speakers (Lewkowicz & Hansen-Tift, 2012) and at 6 months may use visual cues to help establish phonemic categories (Teinonen, Aslin, Alku, & Csibra, 2008). There is evidence that infants show early sensitivity to cross-modal matches and over the first year of life are vigilant to visual cues, possibly indicative of an active role in the development of speech perception. It is important to note, however, that the cross-modal integration of speech stimuli may not be as consistent or mandatory as it is for adults (Desjardins & Werker, 2004).

Research with children has consistently shown that they are less influenced than adults are by visual speech cues (Massaro, 1984), although the pattern of observed change over development seems to depend on the nature of the task adopted. In some studies, children aged 5 to 6 years have, surprisingly, been largely uninfluenced by visual speech. For example, in a sentence perception task, this age group failed to use visual cues to obtain a release from informational masking (Wightman, Kistler, & Brungart, 2006), nor did they show sensitivity to the McGurk effect (Tremblay et al., 2007). In both these studies, a gradual increase in sensitivity was observed into the adolescent years. For single-word perception, Ross et al. (2011) found a benefit of visual cues for 5-year-olds with this benefit again increasing through adolescence. These changes over childhood are driven in part by experience with speech production (Desjardins, Rogers, & Werker, 1997) as well as perceptual experience with language-specific cross-modal matches, resulting in perceptual narrowing (Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009). Gradual behavioral maturation fits with what is known about the late development of superior temporal regions (Gogtay et al., 2004), crucial for both speech perception and cross-modal integration, and studies of neurophysiological responses to audiovisual speech in TD children (Dick, Solodkin, & Small, 2010; Knowland, Mercure, Karmiloff-Smith, Dick, & Thomas, 2014).

Audiovisual Speech in Children With LLI

To date, four studies have considered the ability of children with developmental LLI to use visual speech cues. Pons, Llorenc, Sanz-Torrent, Buil-Legaz, and Lewkowicz (2013) used eye-tracking to examine sensitivity to asynchrony across auditory and visual speech sources in Spanish-speaking children with LLI plus age- and language-matched controls. They found that children with LLI were more tolerant to audiovisual asynchrony, given an auditory lead, compared with the control groups, suggesting atypical integration. Three studies have adopted the McGurk illusion to study integration, finding that children with LLI are less likely to perceive the illusion than their TD peers, being instead more likely to report the auditory stimulus (Leybaert et al., 2014; Meronen, Tiippana, Westerholm, & Ahonen, 2013; Norrix, Plante, Vance, & Boliek, 2007). Children with LLI were also less able to perform silent speechreading (Meronen et al., 2013). It is interesting that McGurk responses in the LLI group were less influenced by the addition of auditory noise than typical control groups for whom noise resulted in more visually influenced responses (Meronen et al., 2013).

If children with LLI are less likely or less able to integrate across modalities, this may either be due to a difficulty with cross-modal integration more broadly or with the extraction of phonetic information in any modality. This is in contrast with children with peripheral hearing loss or auditory neuropathy, who are known to use visual speech cues to good effect (e.g., Bergeson, Pisoni, & Davis, 2003; Ramirez & Mann, 2005).

The Current Study

Previous studies have established that children with LLI are less likely than their peers to show integration of auditory and visual speech cues under illusory circumstances. The aim of the current study was to establish if children with LLI are able to benefit from visual speech cues during a nonillusory paradigm, which is therefore more representative of naturally occurring conditions.

Children completed a silent speechreading task and an adaptive test of speech perception in noise with and without visual support. It was hypothesized that children with LLI would show a developmental lag on both speech perception tasks. As previous work suggests that children with LLI are less likely than TD children to integrate speech cues across modalities, it was further hypothesized that for the speech-in-noise task children with LLI would benefit less from visual cues than their TD peers and that this benefit would be related to speechreading ability as has been previously found (Wightman et al., 2006).

Method

Participants

In total, 41 TD children were recruited to the study (mean age: 8 years 0 months, range: 4 years 5 months to 11 years 10 months), 23 of whom were boys and 18 of whom were girls, and 28 children with a diagnosis of LLI. One of these children was unable to complete the study, leaving 27 in the language-impaired group (mean age: 8 years 10 months, range: 5 years 2 months to 11 years 6 months), 21 of whom were boys and six of whom were girls. Children were recruited from schools in London and Birmingham, with the majority of the children with LLI being recruited from specialist language units. All children with LLI had unexplained oral language difficulties and were diagnosed by educational psychologists or speech and language therapists; the vast majority were reported to have both receptive and expressive difficulties with some having additional diagnoses of dyslexia and/or speech production difficulties. All children in this group were receiving intervention from speech and language therapy services at the time of the study. Children with reported social communication impairments or sensory deficits were excluded from the study. Parents gave informed, written permission for their children to participate, and children provided oral consent prior to the commencement of testing. The study was granted ethical approval by the City University London Division of Language and Communication Science Proportionate Review Committee.

Task Stimuli and Procedure

The two main experimental procedures carried out were the Test of Child Speechreading (ToCS; Kyle, Campbell, Mohammed, Coleman, & MacSweeney, 2013) and a test of speech-in-noise perception (SpiN) with and without visual support. In addition, children completed the following standardized assessments: the British Picture Vocabulary Scale–Second Edition (BPVS-II; Dunn, Dunn, Whetton, & Burley, 1997); the Test of Reception of Grammar–Version 2 (TROG-2; Bishop, 2003); the Castles and Coltheart Test 2, which assesses accuracy of reading regular, irregular, and nonwords (CC-2; Castles et al., 2009); the Recalling Sentences subtest from the Clinical Evaluation of Language Fundamentals–Fourth Edition UK (CELF-4; Semel, Wiig, & Secord, 2006); the Nonword Repetition (NWR) subscale from the Comprehensive Test of Phonological Processing (CTOPP; Rashotte, Torgesen, & Wagner, 1999); and the Pattern Construction subscale from the British Ability Scales–Second Edition (BAS-II; Elliott, Smith, & McCulloch, 1997) to give an indication of nonverbal IQ. Due to limitations in standardization, the 4-year-olds in the study did not complete the Recalling Sentences scale or the CC-2, and the 5-year-olds did not complete the CC-2.

ToCS

This task was developed to assess speechreading ability in typically hearing and deaf children (Kyle et al., 2013). ToCS consists of silent videos of a man or woman speaking in three conditions, which are always completed in the same order. The first condition, single words, consists of a single word at a time, such as *apple,* followed by four pictorial response options. The second condition, sentences, consists of short sentences, such as *She picks an apple from the tree,* again followed by four pictorial response options. In the final condition, short stories, a longer story is presented, such as *Ben was going to school. On the way, he stopped at a shop and bought an apple*. After each stimulus video in this condition, the child is asked two questions with each written above four pictorial response options. For the example given, the two questions were *Where was Ben going?* and *What did he buy?* For full details of the development of this task, see Kyle et al. (2013). All children in the current study completed the first two conditions, but some of the youngest (three children in the LLI group and seven in the TD group) did not complete the third. For children with weak reading skills, the questions presented during the short stories condition were read aloud by the experimenter.

SpiN

Full sentences were used as both target and distractor stimuli with a single target key word being presented in the context of a carrier sentence. Each child completed two conditions of this task, run as two separate blocks. The first condition was auditory only (AO), in which both target and distractor stimuli were auditory, and the second was audiovisual (AV), in which the target sentences were audiovisual (videos) and the distractor stimuli were auditory. Target sentences were of the form “Show the dog where the X is,” where X denotes a single-syllable, highly imageable noun (e.g., key, bear) with mean age of acquisition 4.2 years (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). Sentences were recorded with a digital video camera at 25 frames per second with auditory recording using an external microphone. A native English female speaker with training in phonetics provided the target sentences. The speaker’s head and neck were visible in the presented videos (see Figure 1). Similar distractor sentences were recorded by a native English male speaker and had the form “Show the cat where the X is.” These sentences were recorded as audio only and were used as competing speech tokens during the task. Nouns in the distractor sentences were similar to the target nouns in that they were single syllable (e.g., moon, clock), had a low mean age of acquisition (4.0 years; Kuperman et al., 2012), and were highly imageable. All sentences lasted approximately 2 s and had equal average volume (root-mean-square amplitude). The onset of the target sentences was on average 0.08 s after the onset of the distractor sentence, with the onset of the target key word occurring on average 1.52 s after the start of the sentence and 0.10 s behind the onset of the key word in the distractor sentence. Distractor sentences were sampled at random without replacement. A single competing speaker was selected as the distractor here as natural speech has been found to be the most effective masker of a speech signal (Rosen, Souza, Ekelund, & Majeed, 2013), and further work from this lab suggests that children with LLI find the condition of a single competing speaker particularly challenging. The aim here was to use a distractor that the children would find demanding to see if visual speech would support auditory processing when the auditory system was stressed. In addition, competing speech is most relevant to the classroom setting.

Each trial began with a stimulus presentation. In the case of AV trials (see Figure 1), this consisted of video and audio presentation with a simultaneous auditory distractor. In the case of AO trials, target and distractor stimuli were presented simultaneously while the computer screen remained blank. After stimulus presentation, the response screen was immediately initiated. This consisted of an image of a dog next to four pictures, requiring a four-alternative, forced-choice response from the participant. The four pictures depicted the target plus three distractors, each of which was a minimal pair with the target. For example, the target *chair* was presented with the response options *chair, hair, pear,* and *bear*. The SpiN task was presented on Dell Latitude laptops using MATLAB and PsychToolbox software and through on-ear Panasonic HTX7E-A earphones. Overall intensity of target and distractor combined was approximately 70 dB SPL, as measured at the ear pads using a sound-level meter, remaining consistent regardless of signal to noise ratio (SNR). Before the first trial, the task was explained to participants and three practice trials were completed at an SNR of 20 dB.

The SpiN task used a transformed up-down adaptive staircase (Levitt, 1971) to find the SNR required for

Figure 1. Example audiovisual trial with response screen in the test of speech-in-noise perception task. Stock images are Copyright © Dreamstime.com. Printed with permission.

performance at 70.7% accuracy. In order to find the speech reception threshold (SRT) efficiently only one correct trial was required to reduce the SNR until the first reversal, after which the test reverted to a two-down, one-up procedure to converge on the appropriate SRT. The starting SNR was 20 dB with initial step size 10 dB to the first reversal, 8 dB to the second, 6 dB to the third, and 4 dB thereafter; reversals were counted after the final step size was reached, and from this point, the task ran until six reversals had been made at 4 dB steps or until 32 trials had been completed, whichever occurred sooner.

Session Procedure

Participants completed two sessions lasting around 45 min each. Sessions were run in quiet rooms in the children's schools, such as areas used for special educational needs provision, or private offices. Tasks were always completed in the same order: in the first session, NWR subscale from the CTOPP, CC-2, Pattern Construction subscale from the BAS-II, ToCS, and Recalling Sentences subtest from the CELF-4; in the second session, the first condition of the SpiN task, TROG-2, the second condition of the SpiN task, and BPVS-II. The order in which SpiN conditions were completed was counterbalanced across participants by alternation. At the end of the second session, each child also completed an expressive vocabulary check to make sure that he or she knew all the target words presented during the SpiN task. Accuracy on this vocabulary check was 98.6% (*SD* = 2.1) for the TD group and 96.0% (*SD* = 3.0) for the LLI group. Children were given breaks as judged necessary and were rewarded with a small toy and a certificate.

Results

Groups

Children were recruited into the LLI group if they had a diagnosis of LLI and to the TD group if their teacher had no concerns over their development. However, group membership did not remain stable after a criterion was applied for inclusion in the LLI group, defined as having *z* scores at least as low as −1 on more than one of the standardized language tasks. Three children recruited with a diagnosis of LLI did not meet criterion for impairment, and one child recruited into the TD group did. This resulted in 43 children being included in the TD group (mean age: 8 years 1 month), 23 of whom were boys and 20 of whom were girls, and 25 children in the LLI group (mean age: 8 years 10 months), 21 of whom were boys and four of whom were girls. As the four children who switched group membership may have shown slightly different profiles compared with their fellow group members, all analyses were rerun excluding these four participants. There was no effect on the pattern of results reported. Table 1 gives *z* scores for each standardized task for both groups and the results of fully factorial univariate analyses of covariance run on each test; main effects of Group (TD and LLI) and interactions between Group and Chronological Age (CA; in months) are given. As expected, either a main effect of Group or an interaction between Group and CA is seen for each standardized test of language but not for the Pattern Construction subscale from the BAS-II, supporting criterion for LLI group membership.

Whenever analyses are performed with age (chronological or language) as a variable, age in months has been rescaled such that 0 is the age of the youngest child in the LLI group. This transformation allows for easier interpretation of group effects as groups are compared at the youngest age at which both are represented.

ToCS

The ToCS is scored out of 15 for the single-words and sentences conditions and out of 10 for the short stories condition. Table 2 gives percentage correct scores for each level (along with mean speech reception thresholds for each condition of the SpiN task). The short stories condition was not analyzed further as 10 participants did not attempt this level, and of the remaining 58, 17 scored below chance (<25%), and 52 failed to score significantly above chance (>50%).

To analyze the ToCS data, a linear mixed-effects model analysis was run with Group (TD/LLI), Condition (Words/Sentences), and CA as predictors. The analysis revealed a significant three-way interaction between Condition, Group, and CA, $F(1, 128) = 5.496$, $p = .021$. The effects of CA and Group were then considered for each condition separately. For the single-words condition, the nonsignificant interaction between CA and Group ($p = .585$) was removed, leaving significant main effects of CA, $F(1, 65) = 21.053$, $p < .001$, $f^2 = 0.222$, and Group, $F(1, 65) = 10.739$, $p = .002$, $f^2 = 0.076$. For the sentences condition, an interaction between CA and Group, $F(1, 64) = 9.453$, $p = .003$, emerged, but no main effect of Group ($p = .525$). This interaction was explained by performance improving over age on the sentences condition for the TD group, $R^2 = .500$, $F(1, 41) = 40.60$, $p < .001$, but not the LLI group ($p = .250$).

By including CA as a predictor, the data suggest that the LLI group showed a developmental delay in comparison with their TD peers on the single-words condition of the task from the youngest age measured. On the sentences condition, the LLI group showed a typical onset but an atypically slow rate of development. These patterns are seen in Figure 2, in which, for the single-words condition, performance improves for each group in parallel with increasing age, and for the sentences condition, the gap between the groups progressively widens. To further explore this effect, we repeated the linear mixed-effects analysis using Language Age (LA) as a covariate. BPVS-II age-equivalent score in months (rescaled to 0) was selected as the LA predictor as receptive skills are the most likely language domain to be influenced by speech perception abilities.

In the mixed-effects model, a three-way interaction between Condition, Group, and LA emerged, $F(1, 128) = 7.641$,

Table 1. *Z* scores for each standardized test for both groups and the results of fully factorial univariate analyses of covariance run on each test.

Measure	Group	μ	σ	Analysis of variance	Degrees of freedom	*F*	*p*	η_p^2
NWR[a]	TD	0.000	1.000	Group	1,64	12.038	.001*	.158
	LLI	−2.488	1.043	Age × Group	1,64	2.990	.089	.045
BPVS-II	TD	0.891	0.589	Group	1,64	2.498	.119	.038
	LLI	−0.954	0.599	Age × Group	1,64	28.205	< .001*	.306
TROG-2	TD	0.448	0.779	Group	1,64	19.955	< .001*	.238
	LLI	−2.026	0.791	Age × Group	1,64	2.260	.138	.034
Recalling Sentences	TD	0.519	1.203	Group	1,59	16.111	< .001*	.214
	LLI	−2.054	0.528	Age × Group	1,59	0.846	.361	.014
Reading								
Regular	TD	0.89	0.99	Group	1,52	9.168	.004*	.150
	LLI	−1.66	0.78	Age × Group	1,52	1.621	.209	.030
Irregular	TD	0.677	0.882	Group	1,52	12.198	.001*	.190
	LLI	−1.730	0.721	Age × Group	1,52	0.910	.344	.017
Nonwords	TD	0.592	1.197	Group	1,52	13.891	< .001*	.211
	LLI	−1.516	0.819	Age × Group	1,52	0.193	.662	.004
Average	TD	0.719	0.915	Group	1,52	13.958	< .001*	.212
	LLI	−1.635	0.715	Age × Group	1,52	0.397	.531	.008
Pattern Construction	TD	0.549	0.929	Group	1,64	1.318	.255	.020
	LLI	−0.891	0.671	Age × Group	1,64	2.926	.092	.044

Note. NWR = Nonword Repetition subscale from the Comprehensive Test of Phonological Processing; TD = typically developing; LLI = language learning impairment; BPVS-II = British Picture Vocabulary Scale–Second Edition; TROG-2 = Test of Reception of Grammar–Version 2; Recalling Sentences = Recalling Sentences subtest from the Clinical Evaluation of Language Fundamentals–Fourth Edition UK; Pattern Construction = Pattern Construction subscale from the British Ability Scales–Second Edition. Main effects of Group (TD and LLI) and interactions between Group and Age (in months from the youngest participant in the LLI group) are given.

[a] NWR *z* scores are derived from sample scores out of 63 syllables correct as the published scores are only available from 5 years 0 months of age.

*Significance at p <.01.

$p = .007$, as we saw with CA. For the single-words condition, the nonsignificant interaction between LA and Group ($p = .077$) was removed, leaving a main effect of LA, $F(1, 65) = 16.566$, $p < .001$, $f^2 = 0.348$, but not Group ($p = .724$). For the sentences condition, a significant interaction between Group and LA emerged, $F(1, 64) = 4.744$, $p = .033$, with performance in the TD group improving with increasing LA, $R^2 = .490$, $F(1, 41) = 39.540$, $p < .001$; a pattern not seen for the LLI group ($p = .697$). This analysis suggests that performance on the simpler single-words condition is comparable across groups when compared by language ability, given that no main effect of Group emerged. However, on the sentences condition, the TD group but not the LLI group showed improvement as receptive vocabulary increased, suggesting that the developmental trajectory is still atypical in the LLI group even when compared with peers of comparable language ability. Figure 2 illustrates these relationships and also highlights a limitation here in that variance in BPVS-II scores is reduced in the LLI group (*SD* = 13.9 months) compared with the TD group (*SD* = 36.7 months), restricting the value of LA as a predictor.

Table 2. Percentage correct scores for each Test of Child Speechreading (ToCS) condition (single words/sentences/stories), and speech reception threshold for each condition (auditory only [AO]/audiovisual [AV]) of the test of speech-in-noise perception (SpiN) task.

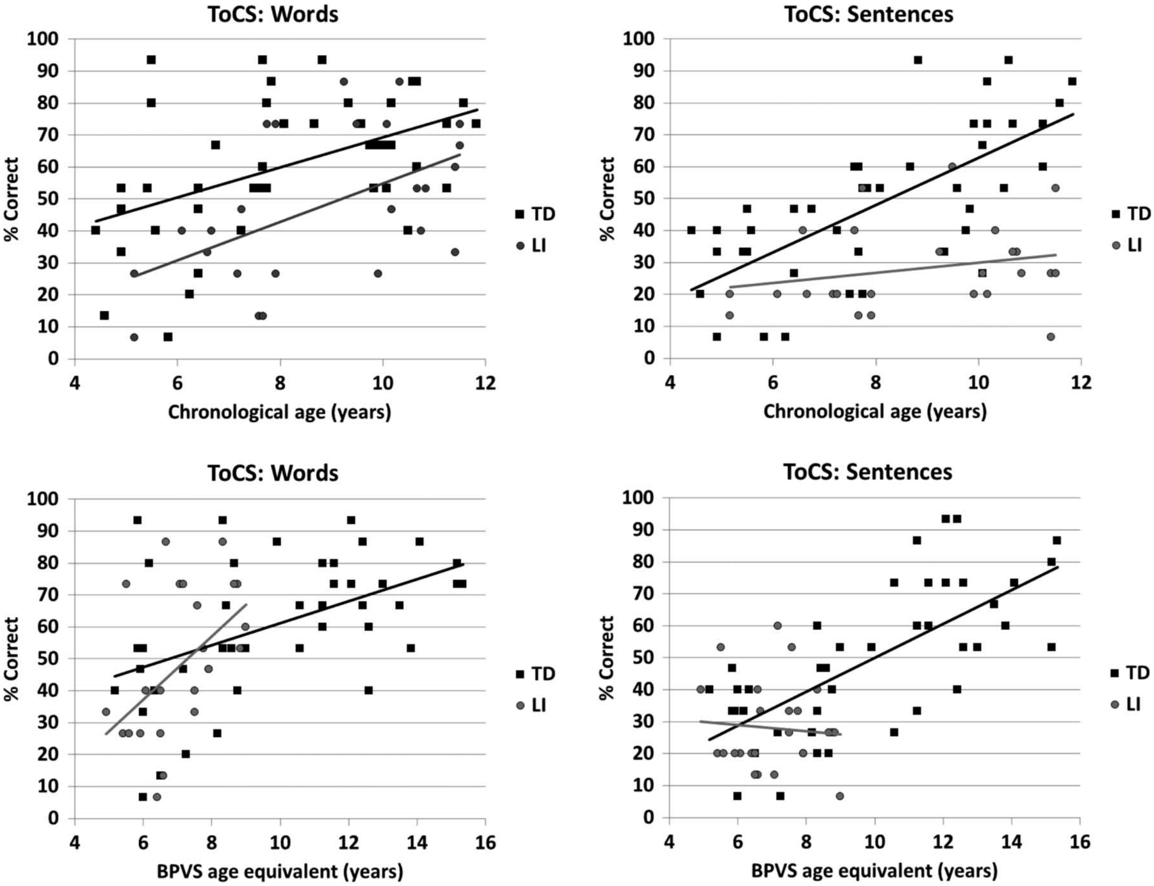
	TD				LLI			
Task	μ	σ	Range	*n*	μ	σ	Range	*n*
ToCS								
Single words	60.3	21.5	6.7:93.3	43	47.7	23.5	6.7:86.7	25
Sentences	48.7	23.2	6.7:93.3	43	28.0	13.6	6.7:60	25
Stories	37.8	17.0	0.0:80.0	36	27.3	13.5	10:50	22
SpiN (dB)								
AO	−7.1	7.2	−24.0:4.7	40	−1.6	7.8	−14.0:17.3	20
AV	−14.6	9.0	−35.3:−1.3	40	−7.8	9.6	−27.0:6.0	20

Note. TD = typically developing; LLI = language learning impairment.

Speech in Noise

SRTs were calculated by averaging across reversals made at steps of 4 dB on each condition (AO and AV) of the SpiN task, such that lower values indicate better performance on the task (see Table 2 for group averages). The difference between this threshold estimate and an estimate made on the basis of interpolation of the point at which performance reached 70.7% from fitted psychometric functions was taken for each participant. A substantial difference between these two measures indicates that factors other than SNR are affecting performance, resulting in a flattened psychometric function. For eight participants (five from the LLI group and three from the TD group), this difference exceeded 5 dB in one or both conditions, and they were consequently excluded from further analysis.

Figure 2. Test of Child Speechreading (ToCS) scores for each of the typically developing and language learning impairment groups in the single-words and sentences conditions plotted against Chronological Age in years and Language Age represented by age-equivalent British Picture Vocabulary Scale–Second Edition (BPVS-II) scores (in years). Regression lines are shown.

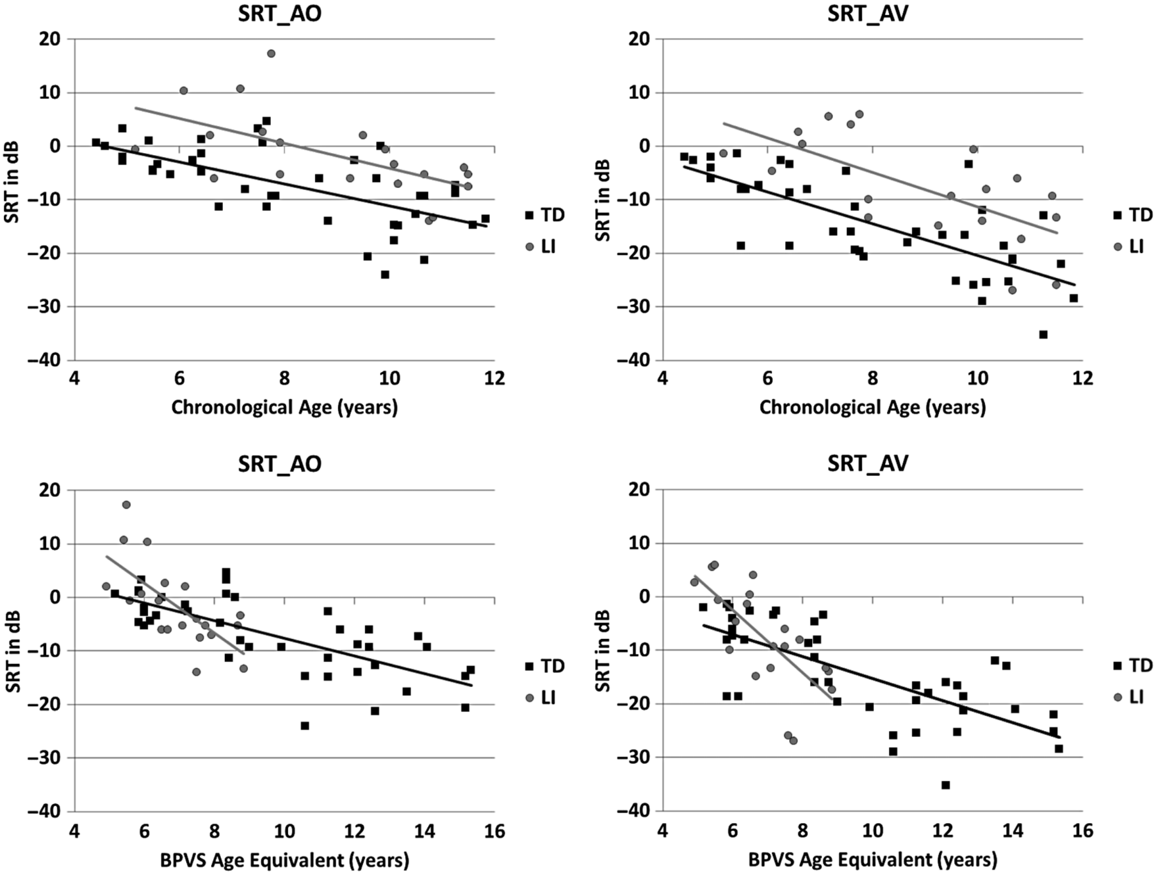


Inspection of the adaptive tracks for these participants indicated lapsing of attention either throughout the task or toward the end of one or both conditions. Excluding these participants resulted in SRT scores showing normal distribution for both the AO condition (TD zSkew = −1.26, zKurtosis = −0.53; LLI zSkew = 1.57, zKurtosis = 0.80) and the AV condition (TD zSkew = −0.48, zKurtosis = −1.17; LLI zSkew = −0.75, zKurtosis = −0.38). Figure 3 shows the relationship between SRTs for each condition and CA/LA (receptive vocabulary) for each group.

SRTs were subjected to a linear mixed-effects model analysis, with Condition (AO/AV), Group (TD/LLI), and CA in months added as predictors. Nonsignificant interactions were systematically eliminated until a main effects model was left, with CA, $F(1, 116) = 97.212$, $p < .001$, $f^2 = 0.374$; Group, $F(1, 116) = 47.441$, $p < .001$, $f^2 = 0.104$; and Condition, $F(1, 116) = 39.374$, $p < .001$, $f^2 = 0.164$, all predicting performance. The failure to find any interactions here indicates that, first, although the LLI group performed more poorly than their peers, the difference between groups did not change over chronological age; second, that both groups were equally likely to make use of visual cues when available; and third, that the benefit of visual cues did not change over age.

Again this analysis was rerun with BPVS-II age-equivalent scores in months (rescaled to 0) to give a measure of LA. A significant interaction between Group and LA emerged, $F(1, 115) = 15.429$, $p < .001$, with each group subsequently analyzed separately. For the TD group, a nonsignificant interaction between LA and Condition ($p = .115$) was removed, leaving main effects of both LA, $F(1, 77) = 76.449$, $p < .001$, $f^2 = 0.692$, and Condition, $F(1, 77) = 33.244$, $p < .001$, $f^2 = 0.217$. For the LLI group, again a nonsignificant interaction between LA and Condition ($p = .433$) was removed, leaving main effects of LA, $F(1, 37) = 34.621$, $p < .001$, $f^2 = 0.745$, and Condition, $F(1, 37) = 9.464$, $p = .004$, $f^2 = 0.133$. The groups showed very similar patterns here, and *R* values did not differ significantly between the groups for the relationship between LA and either AO scores ($p = .889$) or AV scores ($p = 1.000$). This suggests that any group differences actually emerged from the compression of the CA range into the LA range for the LLI group (BPVS-II age-equivalent score *SD* = 13.9 months for the LLI group and *SD* = 37.3 months for the TD group).

Figure 3. Mean speech reception threshold (SRT; in dB) for each participant in the auditory only (AO) and audiovisual (AV) conditions of the test of speech-in-noise perception task plotted against Chronological Age and Language Age given by age-equivalent scores on the British Picture Vocabulary Scale–Second Edition (BPVS-II) in years. Regression lines are shown. Better performance on this task is indicated by a lower score.



See Table 3 for correlation coefficients and regression equations for the relationships between CA/LA and each condition of the ToCS and SpiN tasks.

Intertask Relationships

In order to run correlations between the tasks, to assess if the benefit afforded by visual cues was related to speechreading ability, z scores were computed from the TD group data to norm for the effects of age. Table 4 gives the resulting one-tailed, intertask correlations for the whole sample as well as for each group separately. The benefit of visual cues on the SpiN task was calculated by subtracting scores on the AO condition from scores on the AV condition (AV − AO), such that a more negative score indicates a greater visual advantage.

Cross-task correlations were variable for the sample as a whole. Significant negative relationships emerged between the AV condition of the SpiN task and each level of ToCS, and for the AO condition, a significant negative correlation only emerged with the sentences condition of ToCS. The benefit of visual cues on the SpiN task negatively correlated with the single-words condition of ToCS but not the sentences condition. Each of the correlations emerging from the whole sample remained for the TD group alone, but only the relationship between the single-words condition of ToCS and the benefit of visual cues on the SpiN task remained significant for the LLI group. However, these group differences largely resulted from the different sample sizes given that the Fisher r-z transformation revealed only the correlation between the sentences condition of ToCS and the AO condition of SpiN to be significantly lower in the LLI group (p = .020).

Discussion

The aim of this study was to assess the ability of children with developmental LLI to use visual cues from the talking face to support auditory speech perception. Children completed a silent speechreading task (the ToCS) and a speech-in-noise (SpiN) task with and without visual support from the talking face. On the basis of previous literature, it was hypothesized that children with LLI would be less able than their TD peers to speechread and less likely to use visual cues to support auditory speech perception.

Table 3. Correlation coefficients and regression equations for the relationships between Chronological Age (CA)/Language Age (LA) and each condition of the Test of Child Speechreading (ToCS) and test of speech-in-noise perception (SpiN) tasks.

	TD		LLI	
Factors	R^2	y	R^2	y
CA				
ToCS single words	0.23	46.67 + 4.68 × CA	0.27	25.94 + 5.98 × CA
ToCS sentences	0.50	27.07 + 7.4 × CA	0.06	22.23 + 1.58 × CA
SpiN AO	0.42	−1.18 + −0.17 × CA	0.35	7.08 + −0.19 × CA
SpiN AV	0.56	−6.01 + −2.97 × CA	0.45	4.25 + −3.22 × CA
LA				
ToCS single words	0.24	43.62 + 3.46 × LA	0.24	26.4 + 9.92 × LA
ToCS sentences	0.49	23.01 + 5.32 × LA	0.01	30.06 + −0.96 × LA
SpiN AO	0.51	0.69 + −1.64 × LA	0.48	7.67 + −4.66 × LA
SpiN AV	0.50	−4.86 + −2.05 × LA	0.50	3.87 + −5.85 × LA

Note. TD = typically developing; LLI = language learning impairment; AO = auditory only; AV = audiovisual.

Visual Speech Performance

Children with LLI have previously shown speechreading deficits at the level of consonant-vowel-consonant syllables (Leybaert et al., 2014; Meronen et al., 2013). Consistent with this, on the single-words condition of the ToCS task in the current study, the LLI group showed a pattern of delayed performance in comparison with their TD peers but a normal rate of development. On the sentences condition, however, the LLI group failed to show progression as a function of CA, and this atypical pattern persisted when this group was compared with language-matched peers. However, as noted, unexpectedly low performance in the LLI group on the receptive vocabulary test (BPVS-II) and subsequent low variability limits the reliability of conclusions drawn from comparisons of the groups on the basis of language ability. What is clear is that the LLI group found speechreading whole sentences disproportionately more challenging than single words, and the group difference over development is striking. Sentence-level performance is likely to be limited in the LLI group by working memory (Archibald & Gathercole, 2006; Leonard et al., 2007; Marton & Schwartz, 2003; Vugs, Cuperus, Hendriks, & Verhoeven, 2013) given that during this task children are required to hold in memory any phonological information they were able to extract from the talking face while assessing four pictorial response options. Working memory capacity has previously been suggested as an indirect predictor of sentence-based speechreading (Ronnberg, Samuelsson, & Lyxell, 1998).

Table 4. Correlations between each of the conditions for the Test of Child Speechreading (ToCS) and test of speech-in-noise perception (SpiN; speech reception threshold) tasks, including $A/V_{benefit}$.

	ToCS			
	zWords		zSentences	
SpiN	*r*	*p*	*r*	*p*
zAO				
TD (*n* = 40)	.012	.471	−.295	.033*
LLI (*n* = 20)	.160	.251	.358	.061
All (*N* = 60)	−.125	.170	−.337	.004*
zAV				
TD (*n* = 40)	−.279	.040*	−.313	.025*
LLI (*n* = 20)	−.336	.074	.111	.320
All (*N* = 60)	−.430	< .001**	−.434	< .001**
$zA/V_{benefit}$				
TD (*n* = 40)	−.273	.044*	−.040	.403
LLI (*n* = 20)	−.451	.023*	−.193	.208
All (*N* = 60)	−.368	.002**	−.164	.105

Note. *Z* scores for each measure are derived from TD group data, norming for age. AO = auditory only; TD = typically developing; LLI = language learning impairment; AV = audiovisual.

*Significance at $p < .05$. **Significance after Bonferroni correction for multiple comparisons.

Performance on the SpiN task was significantly improved by the availability of the talking face irrespective of group membership with a moderate effect size ($f^2 = 0.164$) evident for the whole sample when comparing auditory and audiovisual conditions. The actual improvement in the SRT was also quite large, on average about 7 dB. Taking SRT as the dependent variable, children improved on this task in both conditions (AO and AV) with increasing age, although the LLI group performed poorly in comparison with the TD group with delayed development at the youngest point of measurement. The lack of interaction between Condition and CA suggests that the benefit of visual cues was equivalent across ages despite the fact that previous studies have found the benefit of visual cues to increase over development (Ross et al., 2011; Wightman et al., 2006). Failure to find an equivalent effect here may be due to the relative simplicity of the current task for younger children, revealing competence, or may be due to the role that visual speech cues are likely to play during different tasks. No interaction was evident between Group and CA, indicating that the LLI group progressed over development in line with their peers on this task. It is crucial that there was also no interaction between Group and Condition, indicating that both groups benefitted equally from visual cues. So even though the LLI group showed a deficit in speechreading skills, they were able to use visual cues to support auditory processing.

For both groups, the benefit of visual cues (the difference between conditions on the SpiN task) correlated with speechreading ability at the single-word level to the same degree, but neither group showed a correlation with speechreading at the sentence level. Although there are processes common to both speechreading and audiovisual speech perception, a key distinction is that, for speechreading, children rely on visual information to access amodal phonological representations, and in the case of audiovisual speech perception, visual cues support and modulate auditory processing. This distinction is particularly important if auditory speech perception is atypical in children with LLI and especially if it is relevant to the etiology of the disorder (or at least for a subgroup of children). The suggestion from the current study is that children with LLI may be able to boost weak auditory processing in the speech domain by using visual cues to speech. If this could be shown at the earliest point in childhood that a deficit in auditory speech processing is seen, then the use of visual cues could be strengthened clinically to support the development of speech perception over childhood. Speechreading at the sentence level may dissociate from the benefit of visual cues as it requires skills over and above those that support speechreading at the single-word level, such as greater reliance on working memory.

The Audiovisual Advantage

An important question is what support visual cues might provide in the context of the SpiN task adopted here. Visual speech cues have been proposed to play a number of roles in speech perception in adults, including providing complementary information about the content of the speech signal (Grant, Walden, & Seitz, 1998) or predicting the content of the upcoming auditory signal (van Wassenhove, Grant, & Poeppel, 2005), directing spatial attention, and directing attention in time to peaks in acoustic energy (Kim & Davis, 2004). These roles are not mutually exclusive, and it is possible that the nature of the visual benefit changes over development. For example, there is evidence that at 6 months infants use visual cues to help establish phonemic categories (Teinonen et al., 2008). It is also likely that the function of visual cues is task-specific.

In the case of informational masking of auditory targets, as seen during the SpiN task here, multiple aspects of visual speech may prove valuable, including but not limited to content. Visual cues are more beneficial to sentence perception when the masker consists of competing speakers as compared with steady-state noise (Helfer & Freyman, 2004; Wu, Cao, Wu, & Li, 2013), especially when the speakers are perceived to come from the same location (Helfer & Freyman, 2004). Helfer and Freyman (2004) suggest this supports the notion that visual cues are useful in segregating the target speaker in the context of informational masking. This role of visual speech to disambiguate acoustic background and foreground has been highlighted before (Helfer & Freyman, 2004; Wightman et al., 2006; Xu & Barker, 2007), and may be independent of visemic content so long as audiovisual synchrony is intact.

One informative way of thinking about visual cues supporting release from informational masking is in terms of the notion that objects are the unit of attention. This notion is supported through behavioral (Shinn-Cunningham, 2008) and neurophysiological (Ding & Simon, 2012) studies that suggest directing attention toward an object requires first object formation then object selection. In the current SpiN task, then, visual cues may assist in auditory object selection. Selection of the target object during AO trials relies on top-down attention and the ability to use perceptual features of the target and distractor signals to select and process one differentially. The addition of visual cues that correlate temporally with auditory cues (Chandraesekaran et al., 2009) provides information about the temporal structure of the target auditory signal, enabling listeners to override bottom-up salience given lower SNRs. If this interpretation is correct, then children with LLI and young TD children were able to use visual cues here to assist in auditory object selection.

Note that this interpretation is not necessarily contradictory to the finding that children with LLI are less likely than their TD peers to report the McGurk illusion (Leybaert et al., 2014; Meronen et al., 2013; Norrix et al., 2007). Massaro, Thompson, Barron, and Laren (1986) suggested that younger children, who are less capable lip readers, may have less complete information about the visual source of the cross-modal speech signal. A less complete, or less reliable, source will be weighted to a lesser degree during integration (Burr & Alais, 2006; Schwartz, 2010) but may well carry sufficient temporal information to support the selection of the task-relevant auditory object. In future work, this dissociation in task performance needs to be demonstrated in the same sample, then may be further explored by establishing the developmental trajectories of both unisensory speech perception and the extent to which cross-modal cue integration can be described as optimal in each group. Optimal weighting of cross-modal cues during integration has been demonstrated in adults during audiovisual speech perception (Alais & Burr, 2004) and has been shown to emerge across visual and haptic domains (Gori, Del Viva, Sandini, & Burr, 2008) by 8–10 years in children. If individuals with LLI either show weak processing of one or other speech source or suboptimal integration across sources, this may indicate new routes by which to support audiovisual speech perception in this group of children.

Conclusions and Limitations

Previous literature suggests that children with LLI are less likely to report responses indicating the integration of auditory and visual speech signals in illusory contexts. Despite this, the current study demonstrates that this population can use visual speech cues to support auditory perception given a relatively more naturalistic paradigm. It is suggested that in the context of informational masking, visual cues may assist the segregation and preferential processing of the target auditory speech signal. It should be noted, though, that the children with LLI also showed atypical development of speechreading skills, indicating that although beneficial, access to and/or use of visual cues is not equivalent across groups. This raises the possibility that improving speechreading skills could assist children with LLI in their speech perception and language comprehension abilities, either by strengthening amodal phonological representations or by encouraging more robust support for auditory processing. Given that the youngest children in the current study were also able to benefit from visual speech cues, this study emphasizes the importance of better understanding visual speech processing in relation to the development of speech perception and language comprehension over the preschool and early school years. Over this period, children's vocabularies are developing rapidly, and the typical listening conditions change dramatically as children enter the classroom where they are expected to listen to a single speaker in a background of other voices.

The primary limitation of the current data set is the cross-sectional nature of the sample, meaning that the exacerbation of the speechreading deficit over age in the LLI group may be an effect of individual differences (arising from different selection biases across age, for example) rather than developmental change. This is an important consideration in the case of language disorders as profiles vary substantially over developmental time in the LLI population (e.g., Conti-Ramsden & Botting, 1999; Conti-Ramsden, Botting, Simkin, & Knox, 2001), such that children recruited at different ages could show different profiles to children followed to the same age over time longitudinally.

In terms of the tasks adopted here, the requirement to respond to a single target word embedded in a predictable carrier sentence during the SpiN task may have resulted in falsely low SRTs for some of the older TD children who may have been able to adequately speechread targets at low SNRs. The excellent performance of a few members of the TD group point to this task limitation. Future work charting the benefit and limitations of visual speech cues for TD children would be a valuable endeavor to promote better understanding of when and how visual cues support both everyday speech perception in childhood and ongoing changes in perceptual abilities.

Acknowledgments

This work was supported by a fellowship from City University London. We would like to thank Lydia Davis and Ream Katbeh for their assistance with data collection and the schools, parents, and children who were kind enough to participate.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14,* 257–262.

Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010). Neural timing is linked to speech perception in noise. *The Journal of Neuroscience, 30,* 4922–4926.

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders, 41,* 675–693.

Baier, R., Idsardi, W. J., & Lidz, J. (2007, September). *Two-month-olds are sensitive to lip rounding in dynamic and static speech events.* Paper presented at Auditory-Visual Speech Processing 2007 (AVSP2007), Hilvarenbeek, the Netherlands.

Bergeson, T. R., Pisoni, D. B., & Davis, R. A. O. (2003). A longitudinal study of audiovisual speech perception by children who have cochlear implants. *The Volta Review, 103,* 347–370.

Bishop, D. V. M. (2003). *Test for Reception of Grammar–Version 2.* London, United Kingdom: Harcourt Assessment.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research, 46,* 80–97. doi:10.1044/1092-4388(2003/007)

Burlingame, E., Sussman, H. M., Gillam, R. B., & Hay, J. F. (2005). An investigation of speech perception in children with specific language impairment on a continuum of formant transition duration. *Journal of Speech, Language, and Hearing Research, 48,* 805–816.

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology, 45,* 204–220.

Burr, D., & Alais, D. (2006). Combining visual and auditory information. *Progress in Brain Research, 155,* 243–258.

Castles, A., Coltheart, M., Larsen, L., Jones, P., Saunders, S., & McArthur, G. (2009). Assessing the basic components of reading: A revision of the Castles and Coltheart test with new norms. *Australian Journal of Learning Difficulties, 14,* 67–88. doi:10.1080/19404150902783435

Chandraesekaran, C., Trubanov, A., Stilittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology, 5,* e1000436. doi:10.1371/journal.pcbi.1000436

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42,* 1195–1204.

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: Outcomes at 11 years of age. *International Journal of Language & Communication Disorders, 36,* 207–219.

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: A new look at an old hypothesis. *Journal of Speech, Language, and Hearing Research, 50,* 647–666. doi:1092-4388/07/5003-0647

Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech-in-noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology, 112,* 758–767.

Desjardins, R. N., Rogers, J., & Werker, J. (1997). An exploration of why pre-schoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology, 66,* 85–110.

Desjardins, R. N., & Werker, J. F. (2004). Is the integration of heard and seen speech mandatory for infants? *Developmental Psychobiology, 45,* 187–203. doi:10.1002/dev.20033

Dick, A. S., Solodkin, A., & Small, S. (2010). Neural development of networks for audiovisual speech comprehension. *Brain and Language, 114,* 101–114.

Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 11854–11859.

Donlan, C., Cowan, R., Newton, E. J., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition, 103,* 23–33.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale–Second Edition*. Windsor, United Kingdom: NFER-Nelson.

Elliott, C. D., Smith, P., & McCulloch, K. (1997). *British Ability Scales: Second Edition* (BAS II). London, United Kingdom: NFER-Nelson.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., . . . Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 8174–8179.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology, 18,* 694–698.

Grant, K. W., & Seitz, P.-F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America, 108,* 1197–1208.

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. *The Journal of the Acoustical Society of America, 103,* 2677–2690.

Helfer, K. S., & Freyman, R. L. (2004). The role of visual cues in reducing energetic and informational masking. *The Journal of the Acoustical Society of America, 117,* 842–849.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *The Journal of Child Psychology and Psychiatry, 53,* 37–45. doi:10.1111/j.1469-7610.2011.02430.x

Hill, E. (2001). Non-specific nature of specific language impairment: A review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders, 36,* 149–171. doi:10.1080/13682820010019874

Kim, J., & Davis, C. (2004). Investigating the audiovisual speech detection advantage. *Speech Communication, 44,* 19–30.

Knowland, V. C. P., Mercure, E., Karmiloff-Smith, A., Dick, F., & Thomas, M. S. C. (2014). Audio-visual speech perception: A developmental ERP investigation. *Developmental Science, 17,* 110–124.

Kuhl, P., & Meltzoff, A. (1982, December 10). The bimodal perception of speech in infancy. *Science, 218,* 1138–1141.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition for 30,000 English words. *Behavioural Research Methods, 44,* 978–990. doi:10.3758/s13428-012-0210-4

Kushnerenko, E., Teinonen, T., Volein, A., & Csibra, G. (2008). Electrophysiological evidence of illusory audiovisual speech percept in human infants. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 11442–11445.

Kyle, F., Campbell, R., Mohammed, T., Coleman, M., & MacSweeney, M. (2013). Speechreading development in deaf and hearing children: Introducing the test of child speechreading. *Journal of Speech, Language, and Hearing Research, 56,* 416–426. doi:10.1044/1092-4388(2012/12-0039)

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50,* 408–428.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(Suppl. 2), 467–477.

Lewkowicz, D. J., & Hansen-Tift, A. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 1431–1436. doi:10.1073/pnas.1114783109

Leybaert, J., Macchi, L., Huyse, A., Champoux, F., Bayard, C., Colin, C., & Berthommier, F. (2014). Atypical audiovisual speech perception and McGurk effects in children with specific language impairment. *Frontiers in Psychology, 5,* 422. doi:10.3389/fpsyg.2014.00422

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Diorders, 43,* 181–200. doi:10.1080/16066350701340719

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46,* 1138–1153.

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

Massaro, D. W., Thompson, L. A., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology, 41,* 93–113.

McArthur, G. M., & Bishop, D. V. M. (2004). Frequency discrimination deficits in people with specific language impairment: Reliability, validity, and linguistic correlates. *Journal of Speech, Language, and Hearing Research, 47,* 527–541.

McArthur, G. M., & Bishop, D. V. M. (2005). Speech and nonspeech processing in people with specific language impairment: A behavioural and electrophysiological study. *Brain and Language, 94,* 260–273.

McGurk, H., & MacDonald, J. (1976, December 23). Hearing lips and seeing voices. *Nature, 264,* 746–748.

Meronen, A., Tiippana, K., Westerholm, J., & Ahonen, T. (2013). Audiovisual speech perception in children with developmental language disorder in degraded listening conditions. *Journal of Speech, Language, and Hearing Research, 56,* 211–221. doi:10.1044/1092-4388(2012/11-0270)

Norrix, L. W., Plante, E., Vance, R., & Boliek, C. A. (2007). Auditory-visual integration for speech by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 50,* 1639–1651. doi:10.1044/1092-4388(2007/111)

Patterson, M. L., & Werker, J. F. (2003). Two-month-olds match phonetic information in lips and voice. *Developmental Science, 6,* 191–196.

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 10598–10602. doi:10.1073pnas.0904134106

Pons, F., Llorenc, A., Sanz-Torrent, M., Buil-Legaz, L., & Lewkowicz, D. J. (2013). Perception of audiovisual synchrony in Spanish-speaking children with and without specific language impairment. *Journal of Child Language, 40,* 687–700. doi:10.1017/S0305000912000189

Ramirez, J., & Mann, V. (2005). Using auditory speech to probe the basis of noise-impaired consonant-vowel perception in dyslexia and auditory neuropathy. *The Journal of the Acoustical Society of America, 118,* 1122–1133. doi:10.1121/1.1940509

Rashotte, C., Torgesen, J., & Wagner, R. (1999). *The Comprehensive Test of Phonological Processing*. London, United Kingdom: Pearson Assessments.

Reisberg, D., McLean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 97–113). Hillsdale, NJ: Erlbaum.

Remez, R. E. (2005). Three puzzles of multimodal speech perception. In E. Vatikiotis-Bateson, G. Bailly, & P. Perrier (Eds.), *Audiovisual speech* (pp. 12–19). Cambridge, MA: MIT Press.

Robertson, E. K., Joanisse, M. F., Desroches, A. S., & Ng, S. (2009). Categorical speech perception deficits distinguish language and reading impairments in children. *Developmental Science, 12,* 753–767. DOI:10.1111/j.1467-7687.2009.00806.x

Ronnberg, J., Samuelsson, S., & Lyxell, B. (1998). Conceptual constraints in sentence-based lip-reading. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing eye II: The psychology of speechreading and auditory-visual speech* (pp. 143–154). East Sussex, United Kingdom: Psychology Press.

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics, 31,* 509–527. doi:10.1016/S0095-4470(03)00046-9

Rosen, S., Adlard, A., & van der Lely, H. K. (2009). Backward and simultaneous masking in children with grammatical specific language impairment: No simple link between auditory and language ability. *Journal of Speech, Language, and Hearing Research, 52,* 396–411. doi:10.1044/1092-4388(2009/08-0114)

Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A. (2013). Listening to speech in a background of other talkers: Effects of talker number and noise vocoding. *The Journal of the Acoustical Society of America, 133,* 2431–2443. doi:10.1121/1.4794379

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics, 59,* 347–357.

Ross, L. A., Molholm, S., Blanco, D., Gomez-Ramirez, M., Saint-Amour, D., & Foxe, J. J. (2011). The development of multisensory speech perception continues into the late childhood years. *European Journal of Neuroscience, 33,* 2329–2337.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex, 17,* 1147–1153.

Schwartz, J.-L. (2010). A reanalysis of McGurk data suggests that audiovisual fusion in speech perception is subject-dependent. *The Journal of the Acoustical Society of America, 127,* 1584–1594. doi:10.1121/1.3293001

Semel, E., Wiig, E. H., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals–Fourth Edition UK*. London, United Kingdom: Pearson Assessments.

Shinn-Cunningham, B. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences, 12,* 182–186.

Stark, R. E., & Heinz, J. M. (1996a). Perception of stop consonants in children with expressive and receptive-expressive language impairments. *Journal of Speech and Hearing Research, 39,* 676–686. doi:10.1044/jshr.3904.676

Stark, R. E., & Heinz, J. M. (1996b). Vowel perception in children with and without language impairment. *Journal of Speech and Hearing Research, 39,* 860–869. doi:10.1044/jshr.3904.860

Stevens, C., Sanders, L., & Neville, H. (2006). Neurophysiological evidence for selective auditory attention deficits in children with specific language impairment. *Brain Research, 1111,* 143–152. doi:10.1016/j.brainres.2006.06.114

Sumby, W., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America, 26,* 212–215.

Sussman, J. (1993). Perception of formant transition cues to place of articulation in children with language impairments. *Journal of Speech and Hearing Research, 36,* 1286–1299.

Tallal, P., & Benasich, A. (2002). Developmental language learning impairments. *Development and Psychopathology, 14,* 559–579. doi:10.1017.S0954579402003097

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition, 108,* 850–855. doi:10.1016/j.cognition.2008.05.009

Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lapore, F., & Theoret, H. (2007). Speech and non-speech audiovisual illusions: A developmental study. *PLoS ONE, 8,* 742.

Uwer, R., Albrecht, R., & von Suchodoletz, W. (2002). Automatic processing of tones and speech stimuli in children with specific language impairment. *Developmental Medicine & Child Neurology, 44,* 527–532. doi:10.1111/j.1469-8749.2002.tb00324.x

Vance, M., & Martindale, N. (2012). Assessing speech perception in children with language difficulties: Effects of background noise and phonetic contrast. *International Journal of Speech and Language Pathology, 14,* 48–58. doi:10.3109/17549507.2011.616602

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 1181–1186. doi:10.1073/pnas.0408949102

Vugs, B., Cuperus, J., Hendriks, M., & Verhoeven, L. (2013). Visuospatial working memory in specific language impairment: A meta-analysis. *Research in Developmental Disabilities, 34,* 2586–2597.

Warrier, C. M., Johnson, K. L., Hayes, E. A., Nicol, T., & Kraus, N. (2004). Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech-in-noise. *Experimental Brain Research, 157,* 431–441. doi:10.1007/s00221-004-1857-6

Wible, B., Nicol, T., & Kraus, N. (2002). Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. *Clinical Neurophysiology, 113,* 485–494.

Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory-visual integration. *The Journal of the Acoustical Society of America, 119,* 3940–3949.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997, May 8). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387,* 176–178.

Wu, C., Cao, S., Wu, X., & Li, L. (2013). Temporally pre-presented lipreading cues release speech from informational masking. *The Journal of the Acoustical Society of America, 133,* 281–285.

Xu, S., & Barker, J. (2008). Stream weight estimation for multistream audiovisual speech recognition in a multispeaker environment. *Speech Communication, 50,* 337–353.

Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F. X., & Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 14110–14115.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2011). Noise on, voicing off: Speech perception deficits in children with specific language impairment. *Journal of Experimental Child Psychology, 110,* 362–372. doi:10.1016/j.jecp.2011.05.001