LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Children With SLI Can Exhibit Reduced Attention to a Talker's Mouth

Ferran Pons,[a] Monica Sanz-Torrent,[a] Laura Ferinu,[b]
Joan Birulés,[a] and Llorenç Andreu[b]

[a]Universitat de Barcelona and [b]Universitat Oberta de Catalunya

It has been demonstrated that children with specific language impairment (SLI) show difficulties not only with auditory but also with audiovisual speech perception. The goal of this study was to assess whether children with SLI might show reduced attention to the talker's mouth compared to their typically developing (TD) peers. An additional aim was to determine whether the pattern of attention to a talking face would be related to a specific subtype of SLI. We used an eye-tracker methodology and presented a video of a talker speaking the children's native language. Results revealed that children with SLI paid significantly less attention to the mouth than the TD children. More specifically, it was also observed that children with a phonological-syntactic deficit looked less to the mouth as compared to the children with a lexical-syntactic deficit.

Keywords specific language impairment (SLI); children; audiovisual speech; eyes-mouth

Introduction

When interacting with another person, we are exposed to both auditory and visual speech information. However, instead of experiencing these two distinct inputs separately, adults perceive them as a unified entity (Alsius, Navarra, Campbell, & Soto-Faraco, 2005). The auditory and visual speech cues that are typically accessible in the mouth region of a talker consist of correlated patterns of dynamic audible and visible information. Furthermore, studies have

This research was supported by the Ministerio de Economia y Competitividad (PSI2014-55105-P and EDU2016-75368-P) and by AGAUR Generalitat de Catalunya (2014SGR1413).

Correspondence concerning this article should be addressed to Ferran Pons, Department of Cognition, Development and Educational Psychology, Universitat de Barcelona, 171 Pg. de la Vall d'Hebrón, 08035 Barcelona, Spain. E-mail: ferran.pons@ub.edu

DOI: 10.1111/lang.12276

shown that when these two types of information are perceived together there is an increase on salience and comprehension compared to perceiving the auditory information only (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009; King & Calvert, 2001; Reisberg, McLean, & Goldfield, 1987; Rosenblum, 2008; Sumby & Pollack, 1954; Summerfield, 1979; Yehia, Rubin, & Vatikiotis-Bateson, 1998). Research has shown that from the first months of life infants are sensitive to audiovisual speech correspondences (Kuhl & Meltzoff, 1982; Lewkowicz, 2010; Lewkowicz & Ghazanfar, 2006; Lewkowicz & Pons, 2013; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009). Also, recent studies indicate that by the second half of the first year of life infants shift their attention from the eyes to the mouth of a talker. Just as they are beginning to produce speech (i.e., canonical babbling) at 8 to 10 months of age infants are fixating on the mouth, which allows them to access the source of audiovisual speech (Hillairet de Boisferon, Hansen-Tift, Minar, & Lewkowicz, 2017; Lewkowicz & Hansen-Tift, 2012). Moreover, bilingual infants have been found to shift their attention earlier to the mouth, in order to support their dual-language acquisition processes (Pons, Bosch, & Lewkowicz, 2015). Interestingly, it has also been demonstrated that the amount of looking time to the mouth during the first year of life predicts later productive vocabulary (Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan, 2015). Specifically, Young, Merin, Rogers, and Ozonoff (2009) showed that young infants who paid more attention to the mouth of their caregiver had higher expressive vocabulary at 2 years of age, compared to the ones who fixated more on the eyes. Taking the results of all these studies into account, it seems that mouth-looking supports the detection of visual speech cues, which in turn helps language acquisition. Similarly, in adults, the allocation of attention to the mouth appears under increased speech uncertainty. For example, Barenholtz, Mavica, and Lewkowicz (2016) reported that in a speech-processing task adults showed increased attention to the mouth of a talker speaking a nonnative language. Similarly, adults fixated more on the mouth area when auditory noise was added to the speech signal (Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998). Finally, under regular conditions, the speaker's mouth preference can be observed again in older ages, namely, older adults fixated more on the mouth regions of the face than younger adults (Thompson & Malloy, 2004).

To date, only a few studies have explored the significance and use of the visual speech cues located at the mouth in children with speech disorders. It is still unclear whether these children could use the mouth's visual information in order to compensate for their auditory speech difficulties. First, Desjardins, Rogers, and Werker (1997) showed that children who had made developmental speech errors had poorer performance in lip-reading tasks than their control group. However, when assessing the perception of the classic McGurk illusion (McGurk & MacDonald, 1976), Dodd, Mcintosh, Erdener, and Burnham (2008) reported no differences between children with speech disorders and their matched controls. Lastly, Boliek, Keintz, Norrix, and Obrzut (2010) did observe a diminished McGurk effect in children with learning disabilities (LD).

Focusing on specific language impairment (SLI), recent studies indicate that the difficulties this population shows in perceiving speech might not only be due to auditory processing deficits but also reflect a failure in the processing and integration of the audiovisual information. For example, Norrix, Plante, Vance, and Boliek (2007) found that in a classic McGurk task, children with SLI were less biased by the visual information than were their controls. It was concluded that, besides the auditory difficulties children with SLI present, their audiovisual speech processing is also hampered (see also Meronen, Tiippana, Westerholm, & Ahonen, 2013). In the same line, it was observed that children with SLI are not as good as their typically developing (TD) peers in detecting audiovisual asynchrony in speech (Pons, Andreu, Sanz-Torrent, Buil-Legaz, & Lewkowicz, 2013) and nonspeech (Kaganovich, Schumaker, Leonard, Gustafson, & Macias, 2014). More recently, Knowland, Evans, Snell, and Rosen (2016) showed that children with SLI are able to use the visual cues to assist auditory speech perception, although they had lower accuracy than their TD peers on the speechreading and speech-in-noise tasks. Similarly, Heikkilä, Lonka, Kuitunen, and Tiippana (2014) found that children with SLI had poorer speechreading skills than their TD age-matched peers. In addition, Kaganovich, Schumaker, and Rowland (2016) have reported that children with SLI are less sensitive to auditory-articulatory correspondences, suggesting that they have poorly defined correspondences between speech sounds and observable articulatory movements.

The Present Research

The aim of the current study was to establish if children with SLI might show reduced attention to the talker's mouth in a free viewing task compared to their TD peers. An additional aim of this study was to determine whether the pattern of attention to a talking face would be related to a specific subtype of SLI. Although many studies have defined different SLI subtypes[1] (Conti-Ramsden, Crutchley, & Botting, 1997; Rapin & Allen, 1987; van Daal, Verhoeven, & van Balkom, 2004), only the phonological-syntactic and the lexical-syntactic deficits are currently accepted by the American Psychiatric Association (2013; Aguado et al., 2015). Based on previous findings showing that audiovisual speech perception difficulties in children with SLI were related to phonological abilities (e.g., McGurk effect), we predicted that the use of visual cues from a talking face would be different in the phonological-syntactic SLI (PhoSLI) group than in the lexical-syntactic SLI (LeSLI) group. Specifically, we expected the PhoSLI group would not focus on the mouth as much as the LeSLI group.

Method

Participants

Two groups of children participated in this study. The first group consisted of 18 children (13 boys, 5 girls) with SLI, with an age range of 5;01 to 9;07 (mean age: 7;54). The second group consisted of 18 children age-matched with the children with SLI (12 boys, 6 girls), ranging from 5;01–9;05 (mean age: 7;59). Written informed consent was obtained from the parents before the study was run. All participants were Spanish native speakers and had normal or corrected-to-normal vision.

The children with SLI were selected according to standard criteria for diagnosing SLI (Leonard, 1998; Stark & Tallal, 1981). In order to assess their nonverbal intelligence and level of language development, children were evaluated with the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 2004 [Spanish version]), the Spanish version of the Peabody Picture Vocabulary Test III (PPVT-III; Dunn, Dunn & Arribas, 2006), and the *Test de comprensión de estructuras gramaticales* (CEG; Comprehension of syntactic structures; Mendoza, Carballo, Muñoz, & Fresneda, 2005). All the children chosen for this study had been diagnosed with SLI by a speech therapist and they were receiving language therapy. Based on the test scores together with the diagnosis given by the speech therapist, the SLI sample was divided into two groups: (1) the LeSLI deficit group ($n = 8$), characterized by impaired syntax, comprehension difficulties, lexical access, and semantic deficits, and (2) the PhoSLI deficit group ($n = 7$), characterized by a significant alteration of the form of language (phonology and syntax), which may not affect language comprehension. Three participants remained unclassified.

The control group was equivalent in age to the SLI group, $F(1, 34) = 0.12$, $p = .73$. Children who had had speech therapy or psychological treatment were not selected. The control group was selected by their teachers, who also confirmed that they had normal academic performance and language development for their age. All of the children chosen for this study belonged to state schools in Catalonia. The language assessments analyzed by the Mann-Whitney test indicated a significant lower performance by the children with SLI in the

Table 1 Means (standard deviations) for group age, cognitive measures, and language performance

		SLI subtypes		
	SLI	PhoSLI	LeSLI	Age controls
Age (years)	7.54 (1.58)	8.08 (1.77)	7.11 (1.55)	7.59 (1.51)
Peabody (percentile)	23.44[a] (23.01)	33.85[a] (25.10)	11.62[a,b] (11.97)	68.41 (13.71)
CEG (percentile)	14.33[a] (16.55)	17.85 (21.32)	8.50[a] (10.63)	37.53 (25.24)
K-BIT Vocabulary	92.89[a] (12.97)	98.00 (17.15)	86.50[a] (8.26)	107.59 (11.20)
K-BIT Matrices	93.11 (14.84)	94.71 (13.48)	90.50 (18.03)	101.88 (7.98)
MLUw	4.95[a] (1.85)	5.25[a] (2.04)	4.58[a] (2.06)	8.22 (2.24)

Notes. MLU-w = mean length of utterance in words; SLI = specific language impairment. [a]SLI group differs from age controls at $p < .05$; [b]PhoSLI differs from LeSLI at $p < .05$.

PPVT-III ($U = 20.50$, $p < 0.01$), in the CEG ($U = 64$, $p < 0.01$), and in the K-BIT Voc ($U = 56.50$, $p < 0.01$), in comparison to the control group (see Table 1).

Procedure

The experiment was run in the school of the children. They were tested individually, seated approximately 22 inches in front of the Tobii T120 eye-tracker integrated into a 17-inch TFT monitor at a sampling rate of 120 Hz. The Tobii eye tracker's nine-point calibration routine was used. After the calibration routine was completed, a 45-second video of a female talker reciting a prepared monologue in the children's native language was presented. Children were not given any specific instructions other than that they had to attend to the monitor screen while a woman was explaining a story. While children watched the video, their selective attention was measured by monitoring point of gaze (POG) with the eye tracker. Children were only included if they looked at the face for at least 75% of the time (34 seconds). In the current study all the children met this criteria.

To analyze the relative degree to which they attended to the eyes versus the mouth of the talker, the same areas of interest (AOI) used in Pons et al. (2015) were chosen. One was the eye AOI, which consisted of the area around the talker's eyes, and the other was the mouth AOI, which consisted of the area around the talker's mouth. We measured the amount of time children spent gazing at each AOI.

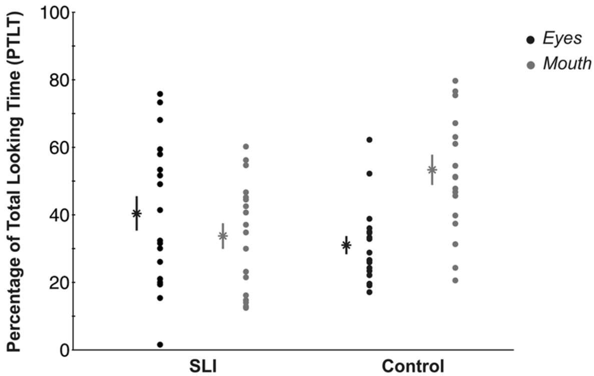

Figure 1 Proportion of total looking time at the speaker's eyes and mouth for the children with specific language impairment and age-matched control group. Circles represent each child's score; asterisk with error bars represent mean score and standard error of the mean for each group.

Results

To first determine whether overall attention varied across groups, we analyzed the total amount of looking at the face with a one-way analysis of variance (ANOVA) with group (SLI, TD) as a between-participants factor. The group effect was not significant, $F(1, 34) = .579, p = .452$, indicating that the pattern of shifting attention was not attributable to differences in overall attention. Subsequently, the proportion of total looking time (PTLT) directed at each AOI was calculated, respectively, by dividing the total amount of time children looked at each AOI by the time they spent looking at any part of the face (see Figure 1). The PTLT scores were analyzed by way of a mixed ANOVA with AOI (eyes, mouth) as a within-participants factor and group (SLI, TD) as a between-participants factor. Results yielded a significant main effect of AOI, $F(1, 34) = 7.434$, $p = .01$, $\eta^2 = .179$, indicating that there was an overall preference for the mouth. Also, a significant AOI $\times$ Group interaction was found, $F(1, 34) = 9.883$, $p = .003$, $\eta^2 = .225$, indicating that looking at the two areas of the face differed as a function of group. Planned comparison tests of the interaction showed that the TD group looked longer at the mouth than the eyes, $F(1, 17) = 16.626$, $p < .001$, $\eta^2 = .494$, but the SLI group looked equally at the eyes and mouth, $F(1, 17) = .090, p = .767$.

Additionally, the behavior of the specific subtypes of children with SLI was compared in order to see whether the specific deficit would modulate the use of the visual speech cues located at the mouth. Nonparametric statistical

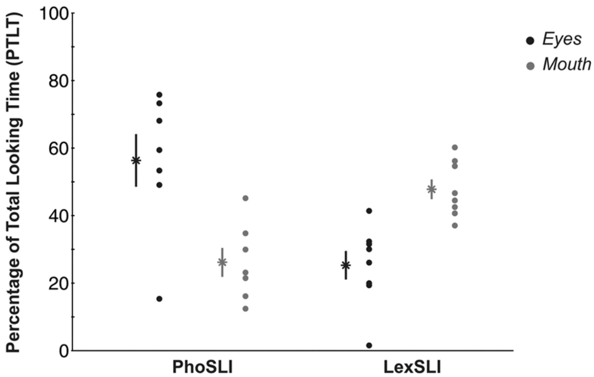

Figure 2 Proportion of total looking time at the speaker's eyes and mouth for the two subgroups of children with specific language impairment: lexical-syntactic (LeSLI) and phonological-syntactic (PhoSLI) subgroups. Circles represent each child's score; asterisk with error bars represent mean score and standard error of the mean for each group.

tests were chosen due to the small size of the subgroups. A nonparametric Mann-Whitney U test revealed a significant difference between the PhoSLI and the LeSLI subgroups in the PTLT to the mouth ($U = 4$, $p = .005$) and to the eyes ($U = 7$, $p = .015$). Specifically, the LeSLI subgroup was more likely to look to the mouth ($M = 0.48$) than the PhoSLI subgroup ($M = 0.26$). Moreover, the PhoSLI subgroup also showed a higher proportion of looks to the eyes ($M = 0.56$) than the LeSLI subgroup ($M = 0.25$). Finally, in order to compare the PTLT scores to the eyes with the PTLT scores to the mouth for each subtype separately, a within-subjects comparison for each group using a Wilcoxon Signed-ranks test was used. Results revealed a significant difference between PTLT to the eyes and PTLT to the mouth in both the PhoSLI group ($Z = -2.197$, $p = .028$) and in the LeSLI group ($Z = -2.375$, $p = .018$) (see Figure 2).

Discussion

The purpose of this study was to investigate if children with SLI might show reduced attention to a talker's mouth compared to their TD peers. Moreover, we aimed to analyze if the pattern of attention to the talker's eyes and mouth might be related to a specific subtype of SLI. To test these predictions, we presented a video of a female talker who spoke in the children's native language and recorded their eye gaze. First, the results reveal that TD children spent more time attending to the mouth compared to the eyes. As reported during infancy (Lewkowicz & Hansen-Tift, 2012), this indicates that looking at the mouth may still be advantageous for children in this speech task. Furthermore, as expected, the group of children with SLI did not show any preference for the mouth or the eyes: They looked equivalently to both. These findings align with previous results showing that children with SLI seem to have difficulties perceiving and integrating the audiovisual cues in speech (Kaganovich et al., 2016; Knowland et al., 2016; Pons et al., 2013). The previously reported difficulties of children with SLI in speechreading, auditory processing may bring about this inability to properly process audiovisual cues. The current results support again the idea that the typical speech-perception difficulties that children with SLI present are likely to extend beyond auditory processing deficits only (Kaganovich et al., 2014, 2016; Knowland et al., 2016; Meronen et al., 2013; Norrix et al., 2007).

The second aim of this study was to determine whether the pattern of attention to a talking face would be related to the specific subtypes of SLI. As predicted, children with SLI and phonological-syntactic deficits (PhoSLI) spent less time looking to the mouth than children who had lexical deficits (LeSLI). Considering the language capacities of these children, one possible explanation would be that the low production abilities (PhoSLI) are related to an inability to properly use and combine the visual-speech information (Knowland et al., 2016; Meronen et al., 2013; Norrix et al., 2007). Following this idea, children with lexical-syntactic deficit would show more looks to the mouth because they are able to extract the relevant information that may help them to disambiguate and/or imitate speech sounds. On the other hand, children with a phonological deficit would not be able to use or integrate the audiovisual speech information properly and therefore would disregard this visual cue (Barenholtz et al., 2016; Knowland et al., 2016; Lewkowicz & Hansen-Tift, 2012). Our data might suggest that children with SLI are using a different learning strategy to the one employed by TD children, which can be modulated by their specific linguistic deficit. These conclusions are important considering that the stimuli used in the current experiment consisted of naturalistic fluent speech, framed in a real context (provided by the narration), as opposed to presenting an isolated syllable or a congruent/incongruent paradigm (as in the McGurk paradigm).

In other words, our results indicate that children with SLI as a group look less to the mouth than the TD group, and this pattern seems to be led by the children who have a phonological-syntactic deficit. It is likely that paying attention to the mouth—and hence being able to better extract the speech sound information—is related to having better phonological skills; whether the former influences the latter or vice versa is yet to be studied. It is important to note that no comprehension measure of the video children watched was obtained. Therefore, the link between the total looking time spent on the mouth and the discourse comprehension is still missing. Further studies that measure comprehension and compare it to looking patterns to a talking face should be done in order to tackle this question.

Finally, an important practical aspect of this study concerns its clinical implications. Primarily, our results highlight again the importance of including audiovisual materials in speech therapy and stress the relevance of paying attention to the mouth. The training in the audiovisual speech integration reported in adults (Preminger & Ziegler, 2008; Woodhouse, Hickson, & Dodd, 2009) and hearing-impaired children (Lachs, Pisoni, & Iler Kirk, 2001) indicates the importance of using audiovisual approaches for speech and language intervention. It is crucial to understand that within the children with SLI there might be differences in processing audiovisual speech. Hence, considering the SLI subtypes can help adjust and individualize audiovisual speech therapies.

Final revised version accepted 7 November 2017

Note

1 Originally, Rapin and Allen (1983) proposed a classification of the SLI with six subtypes. The classification was based on an assessment of spontaneous and directed language, taking into account different linguistic levels (phonological, morphosyntactic, semantic-lexical, and pragmatic). The subtypes were phonologic-programming deficit, verbal dyspraxia, phonologic-syntatic deficit, verbal auditory agnosia, lexical-syntactic deficit, and semantic-pragmatic deficit.

References

Aguado, G., Coloma, C. J., Martínez, A. B., Mendoza, E., Montes, A., Navarro, R., et al. (2015). Documento de consenso elaborado por el comité de expertos en TEL sobre el diagnóstico del trastorno. *Rev. logop. foniatr. audiol. (Ed. impr.)*, 147–149.

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*, 839–843. https://doi.org/10.1016/j.cub.2005.03.046

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington, DC: American Psychiatric Pub.

Barenholtz, E., Mavica, L., & Lewkowicz, D. J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition*, *147*, 100–105. https://doi.org/10.1016/j.cognition.2015.11.013

Boliek, C., Keintz, C., Norrix, L., & Obrzut, J. (2010). Auditory-visual perception of speech in children with learning disabilities: The McGurk effect. *Canadian Journal of Speech-Language Pathology & Audiology*, *34*(2), 124–131.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*, e1000436. https://doi.org/10.1371/journal.pcbi.1000436

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research*, *40*, 765–777. https://doi.org/10.1044/jslhr.4004.765

Desjardins, R. N., Rogers, J., & Werker, J. F. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology*, *66*, 85–110. https://doi.org/10.1006/jecp.1997.2379

Dodd, B., Mcintosh, B., Erdener, D., & Burnham, D. (2008). Perception of the auditory-visual illusion in speech perception by children with phonological disorders. *Clinical Linguistics & Phonetics*, *22*, 69–82. https://doi.org/10.1080/02699200701660100

Dunn, L. M., Dunn, L. M., & Arribas, D. (2006). *PPVT-III. Peabody. Test de vocabulario en imágenes*. Madrid, Spain: TEA Ediciones.

Heikkilä, J., Lonka, E., Kuitunen, S., & Tiippana, K. Speechreading skills in Finnish-speaking adults, typically developing children and children with specific language impairment. In LSCD 2014: *Workshop on Late Stages in Speech and Communication Development* (p. 70).

Hillairet de Boisferon, A., Hansen-Tift, A., Minar, N. J., & Lewkowicz, D. J. (2017). Selective attention to a talker's mouth in infancy: role of audiovisual temporal synchrony and linguistic experience. *Developmental Science*, *20*(3), e12381. https://doi.org/10.1111/desc.12381

Kaganovich, N., Schumaker, J., Leonard, L. B., Gustafson, D., & Macias, D. (2014). Children with a history of SLI show reduced sensitivity to audiovisual temporal asynchrony: An ERP study. *Journal of Speech, Language, and Hearing Research*, *57*, 1480–1502. https://doi.org/10.1044/2014_JSLHR-L-13-0192

Kaganovich, N., Schumaker, J., & Rowland, C. (2016). Atypical audiovisual word processing in school-age children with a history of specific language impairment: An event-related potential study. *Journal of NeurodevelopMental Disorders*, *8*, 33. https://doi.org/10.1186/s11689-016-9168-3

Kaufman, A. S., & Kaufman, N. L. (2004). *KBIT: Kaufman Brief Intelligence Test (KBIT, Spanish version)*. Madrid, Spain: TEA Editions.

King, A. J., & Calvert, G. A. (2001). Multisensory integration: Perceptual grouping by eye and ear. *Current Biology*, *11*, R322-R325. https://doi.org/10.1016/S0960-9822(01)00175-0

Knowland, V. C., Evans, S., Snell, C., & Rosen, S. (2016). Visual speech perception in children with language learning impairments. *Journal of Speech, Language, and Hearing Research*, *59*, 1–14. https://doi.org/10.1044/2015_JSLHR-S-14-0269

Kuhl, P. K., & Meltzoff, A. N. (1982, December 10). The bimodal perception of speech in infancy. *Science*, 1138–1141. https://doi.org/10.1126/science.7146899

Lachs, L., Pisoni, D. B., & Kirk, K. I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear and hearing*, *22*(3), 236.

Leonard, L. (1998). *Specific language impairment*. Cambridge, MA: MIT Press.

Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental Psychology*, *46*, 66. https://doi.org/10.1037/a0015579

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences*, *103*, 6771–6774. https://doi.org/10.1073/pnas.0602027103

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

Lewkowicz, D. J., & Pons, F. (2013). Recognition of amodal language identity emerges in infancy. *International Journal of Behavioral Development*, *37*, 90–94. https://doi.org/10.1177/0165025412467582

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

Mendoza, E., Carballo, G., Muñoz, J., & Fresneda, M. D. (2005). *Test de comprensión de estructuras gramaticales (CEG)*. Lexicografía y enseñanza de la lengua Española, 151.

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. https://doi.org/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. https://doi.org/10.1044/1092-4388(2007/111)

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

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infant's selective attention to the mouth of a talking face. *Psychological Science*, *26*, 490–498. https://doi.org/10.1177/0956797614568320

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*, *106*, 10598–10602. https://doi.org/10.1073/pnas.0904134106

Preminger, J. E., & Ziegler, C. H. (2008). Can auditory and visual speech perception be trained within a group setting?. *American Journal of Audiology*, *17*(1), 80–97.

Rapin, I., & Allen, D. A. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading, and spelling* (pp. 155–184). New York: Academic Press.

Rapin, I., & Allen, D. A. (1987). Developmental dysphasia and autism in preschool children: Characteristics and subtypes. In J. Martin, P. Martin, P. Fletcher, P. Grunwell, & D. Hall (Eds.), *Proceedings of the First International Symposium on Specific Speech and Language Disorders in Children* (pp. 20–35). London: AFASIC.

Reisberg, D., Mclean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In Dodd B. & Campbell R. (Eds), *Hearing by eye: The psychology of lip-reading* (pp. 97–114). London: Erlbaum.

Rosenblum, L. D. (2008). Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science*, *17*, 405–409. https://doi.org/10.1111/j.1467-8721.2008.00615.x

Stark, R. E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of speech and hearing disorders*, *46*, 114–122. https://doi.org/10.1044/jshd.4602.114

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*, 212–215. https://doi.org/10.1121/1.1907309

Summerfield, A. Q. (1979). The use of visual information in phonetic perception. *Phonetica*, *36*, 314–331. https://doi.org/10.1159/000259969

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, *42*, 1173–1190. https://doi.org/10.1017/S0305000914000725

Thompson, L. A., & Malloy, D. (2004). Attention resources and visible speech encoding in older and younger adults. *Experimental Aging Research*, *30*, 241–252. http://doi.org/10.1080/03610730490447877

van Daal, J., Verhoeven, L., & van Balkom, H. (2004). Subtypes of severe speech and language impairments: Psychometric evidence from 4-year-old children in the Netherlands. *Journal of Speech, Language, and Hearing Research*, *47*, 1411–1423. http://doi.org/10.1044/1092-4388(2004/105)

Vatikiotis-Bateson, E., Eigsti, I. M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Attention, Perception, & Psychophysics*, *60*, 926–940. https://doi.org/10.3758/BF03211929

Woodhouse, L., Hickson, L., & Dodd, B. (2009). Review of visual speech perception by hearing and hearing - impaired people: Clinical implications. *International Journal of Language & Communication Disorders*, *44*(3), 253–270.

Yehia, H., Rubin, P., & Vatikiotis-Bateson, E. (1998). Quantitative association of vocal-tract and facial behavior. *Speech Communication*, *26*, 23–43. https://doi.org/10.1016/S0167-6393(98)00048-X

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6-months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, *12*, 789–814. https://doi.org/10.1111/j.1467-7687.2009.00833.x