Development of Sensitivity to Audiovisual Temporal Asynchrony During Midchildhood

Natalya Kaganovich
Purdue University

Temporal proximity is one of the key factors determining whether events in different modalities are integrated into a unified percept. Sensitivity to audiovisual temporal asynchrony has been studied in adults in great detail. However, how such sensitivity matures during childhood is poorly understood. We examined perception of audiovisual temporal asynchrony in 7- to 8-year-olds, 10- to 11-year-olds, and adults by using a simultaneity judgment task (SJT). Additionally, we evaluated whether nonverbal intelligence, verbal ability, attention skills, or age influenced children’s performance. On each trial, participants saw an explosion-shaped figure and heard a 2-kHz pure tone. These occurred at the following stimulus onset asynchronies (SOAs): 0, 100, 200, 300, 400, and 500 ms. In half of all trials, the visual stimulus appeared first (VA condition), and in the other half, the auditory stimulus appeared first (AV condition). Both groups of children were significantly more likely than adults to perceive asynchronous events as synchronous at all SOAs exceeding 100 ms, in both VA and AV conditions. Furthermore, only adults exhibited a significant shortening of reaction time (RT) at long SOAs compared to medium SOAs. Sensitivities to the VA and AV temporal asynchronies showed different developmental trajectories, with 10- to 11-year-olds outperforming 7- to 8-year-olds at the 300- to 500-ms SOAs, but only in the AV condition. Lastly, age was the only predictor of children’s performance on the SJT. These results provide an important baseline against which children with developmental disorders associated with impaired audiovisual temporal function—such as autism, specific language impairment, and dyslexia—may be compared.

Keywords: audiovisual temporal processing in children, development of audiovisual integration skills in children, cognitive development

This article was published Online First November 16, 2015.
Natalya Kaganovich, Department of Speech, Language, and Hearing Sciences and Department of Psychological Sciences, Purdue University.

This research was supported in part by grants P30DC010745 and R03DC013151 from the National Institute on Deafness and Other Communicative Disorders, National Institutes of Health. The content is solely the responsibility of the author and does not necessarily represent the official view of the National Institute on Deafness and Other Communicative Disorders or the National Institutes of Health. Thanks to Jennifer Schumaker, James Hengenius, Kevin Barlow, and Caryn Herring for their assistance with different stages of this project and to children and their families for participation.

Correspondence concerning this article should be addressed to Natalya Kaganovich, Department of Speech, Language, and Hearing Sciences, Purdue University, Lyles-Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907-2038. E-mail: kaganovi@purdue.edu
cantly larger separations than adults between auditory and visual modalities (both for speech and nonspeech stimuli) in order to notice the asynchrony (e.g., Lewkowicz, 1996, 2010; for a review, see also Lewkowicz, 2012), suggesting that the temporal binding window narrows with development. Such narrowing, however, may not start in earnest until approximately 6 years of age, based on the study by Lewkowicz and Flom (2014), who reported that 6-year-olds, but not 4- or 5-year-olds, were able to detect the 366-ms stimulus onset asynchrony (SOA) in speech stimuli when the visual modality lagged the auditory one.

The temporal binding window continues to narrow during school years, although the nature of such narrowing and its timing are still somewhat uncertain. Hillock and colleagues were among the first to examine sensitivity to audiovisual asynchrony in school-age children. In their 2011 study, Hillock, Powers, and Wallace (2011) compared the ability to detect audiovisual and visuo-auditory temporal asynchrony in 10- to 11-year-old children and in adults. They used simple auditory and visual stimuli (a pure tone and a ring flash) and presented them at SOAs that ranged from 50 to 450 ms. Children were significantly worse than adults at detecting 150- to 350-ms SOAs between the stimuli but only when the auditory stimulus preceded the visual one. In a follow-up study (Hillock-Dunn & Wallace, 2012), the authors used essentially the same stimuli (with a slightly different set of SOAs), but instead of testing just one group of children, they tested two—those who were 6–11 years of age and those who were 12–17 years of age. A group of 18- to 23-year-old adults was also included. They reported that in this case both groups of children were more likely to perceive stimuli as being synchronous at moderate-to-large SOAs regardless of the order of modalities. However, the two groups of children did not differ from each other. In other words, it appeared that there was no statistically significant improvement in children’s sensitivity to audiovisual asynchrony between approximately 9.5 and 14.7 years of age (which were the mean ages for each of the children’s groups). The two studies by Hillock-Dunn and colleagues were among the first to report on the surprisingly prolonged developmental course of audiovisual temporal processing and, as such, deserve recognition; yet, their results do not paint a coherent picture. According to the 2011 study, perception of visual-first (VA) asynchronies is mature by 10–11 years of age. Yet, according to the 2012 study, perception of VA asynchronies remains immature all throughout the teen years. In an earlier study, my laboratory also examined sensitivity to audiovisual temporal asynchrony in 7- to 11-year-old typically developing (TD) children, same-age children with a history of specific language impairment (H-SLI), and adults (Kaganovich, Schumaker, Leonard, Gustafson, & Macias, 2014). Similarly to the studies by Hillock and colleagues, we used simple stimuli (an explosion-shaped flash and a pure tone) presented at 100–500-ms SOAs in 100-ms intervals. We found not only that H-SLI children were significantly less likely to detect asynchrony than their TD peers at longest SOAs, but also that TD children were significantly less likely to detect asynchrony than adults at most SOAs exceeding 100 ms, regardless of the order of modalities.

In summary, previous studies on the development of sensitivity to audiovisual temporal asynchrony during preschool and school years point unequivocally to a protracted developmental course of this ability. However, previous reports disagree to some extent on whether age-related improvement in audiovisual temporal function proceeds on the same timeline for auditory-visual and visual-auditory SOAs and on whether this function improves significantly during school years. In part, the discrepancy may result from significant individual variability in sensitivity to temporal asynchrony, especially in children of various ages. As an example, in the earlier study from my laboratory described above, adults varied greatly in the ability to detect audiovisual temporal asynchrony, but only when SOAs were relatively short (100–300 ms). At longer SOAs (400–500 ms), adults’ responses were very consistent, with the absolute majority detecting asynchrony on 94–100% of trials. In contrast, children tended to have large individual variability even at long SOAs, with trials on which children reported perceiving asynchrony varying between 42% and 100%. The degree of individual variability in children suggests that age-related improvements in temporal function may be obscured when children have a broad range of ages included in the same group.

In order to examine in greater detail how sensitivity to audiovisual temporal information matures during school years, we tested two groups of children with narrow age ranges—7- to 8-year-olds and 10- to 11-year-olds—as well as adults on a simultaneity judgment task (SJT). We used a flash of an explosion-shaped figure and a pure tone as stimuli and presented them either synchronously or with SOAs that ranged from 100 to 500 ms in 100-ms steps. The use of simple stimuli allowed us to turn the paradigm into a game and to easily compare our findings with earlier reports, most of which used stimuli similar to ours. We measured the number of synchronous perceptions at each SOA and expected that as the separation between modalities increased, the number of synchronous reports would decrease. We also recorded response time (RT) as the average amount of time from the onset of the first stimulus in each pair (or from the onset of the audiovisually synchronous stimulus) needed to make a simultaneity judgment. Studies of the temporal binding window rarely, if ever, report RT, focusing primarily on percent of synchronous perceptions. However, the RT measure adds important complementary information about individuals’ sensitivity to audiovisual temporal SOAs. More specifically, RT data provide insight into the ease or certainty with which a decision is made. We predicted that synchronous trials and trials with the longest SOAs would be associated with relatively short RTs, while trials with intermediate SOAs (where the uncertainty about the synchronicity of audiovisual information remains high) would be associated with longer RTs.

Lastly, the factors that underlie improvement in audiovisual temporal function during childhood are not yet clear. One possibility is that greater sensitivity to audiovisual temporal asynchrony may be related to stronger attention, language, or nonverbal intelligence skills because such skills may help children better cope with the demands of the experimental task. In order to test this hypothesis, we evaluated children’s attention and language aptitude as well as nonverbal intelligence and explored the relationship between these functions and sensitivity to audiovisual temporal asynchrony in a multiple regression analysis.
Method

Participants

Fifteen 7- to 8-year-olds (mean age 8 years, 2 months; range 7 years, 4 months to 8 years, 9 months; 8 female), 16 10- to 11-year-olds (mean age 11 years, 2 months; range 10–12 years; 10 female), and 16 adults (mean age 23, range 20–40; 5 female) participated in the study. All gave their written consent or assent to participate in the experiment. Additionally, at least one parent of each child gave a written consent to enroll their child in the study. The study was approved by the Institutional Review Board of Purdue University, and all study procedures conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki) (1964). None of the children showed any signs of autism (Childhood Autism Rating Scale, 2nd ed.; CARS-2; Schopler, Van Bourgondien, Wellman, & Love, 2010) or had a diagnosis of attention-deficit/hyperactivity disorder (ADHD) (based on parental report). Mothers of 7- to 8-year-old children had on average more years of education than did mothers of 10- to 11-year-old children, group: $F(1, 30) = 4.275, p = .048$; however, this difference was only 2.1 years. A difference in father’s years of education was not significant, $F(1, 28) = 3.593, p = .069$ (data for father’s years of education were not available for one child in each group). According to the Laterality Index of the Edinburgh Handedness Questionnaire, one child in the 7- to 8-year-old group was left-handed; all other participants were right-handed (Cohen, 2008; Oldfield, 1971). All participants were free of neurological disorders, passed a hearing screening at a level of 20 dB HL at 500, 1000, 2000, 3000, and 4000 Hz, and reported having normal or corrected-to-normal vision.

Stimuli and Experimental Design

We used a flash of light shaped as a cartoon explosion (see Figure 1) and a pure 2-kHz tone, both 200 ms in duration, as stimuli. They were presented either synchronously or at the following SOAs: 100, 200, 300, 400, and 500 ms. In half of all audiovisual trials, the sound preceded the flash of light (AV trials), while in the other half, the flash of light preceded the sound (VA trials). Additionally, auditory only (A) and visual only (V) trials were also included in order to examine potential differences in responding to each modality at different ages. The experiment consisted of 10 blocks, with five instances of each of 13 types of trial (5 AV SOAs, 5 VA SOAs, synchronous presentation, A, and V) presented in a random order. This arrangement yielded 50 responses for each trial type. In order to avoid visual aftereffects, the explosion image was presented at slightly different locations in the center of the screen on each consecutive trial. Auditory stimuli were presented at 60 dB SPL via a sound bar located directly under the computer monitor. All participants sat approximately 4 feet from the monitor inside a dimly lit sound-attenuating booth. Presentation of trials was controlled by the Presentation software (www.neurobs.com). Responses were recorded within a 2200-ms response window time-locked to the appearance of the first stimulus in each pair. The response window was followed by an intertrial interval varying randomly among four values: 350, 700, 1050, and 1400 ms. Hand-to-response-button mapping was counterbalanced across participants.

The SJT used in this study was identical to that described in an earlier study from my laboratory (Kaganovich et al., 2014). Prior to the start of the session, all participants viewed a video with instructions and practiced the task until it was clear. Instructions were kept identical for children and adults. At the beginning of the experiment, participants saw a dragon in the middle of the screen and a boy and a girl holding futuristic-looking weapons at the top-left and top-right corners of the screen (see Figure 1). Participants were told that the boy and the girl live on the planet Cabula, where dragons raid apple orchards and eat all the apples. The only way to keep dragons away is to use special weapons—one weapon shoots lights while the other shoots sounds. When the light and the sound “hit” dragons at exactly the same time, dragons run away. However, if either the sound or the light is delayed, even by the tiniest amount, or if one of the children forgets to shoot altogether (and hence there is either only the flash of light or only the sound present), dragons do not get scared away. After each trial, participants were asked to press one button on a response pad (RB-530, Cedrus Corporation) if the dragon was scared away (i.e., if the sound and the flash of light were perceived as synchronous) and another button if the dragon was not scared away (i.e., if the sound and the flash of light were perceived as being asynchronous and if only one of the two stimuli was present). The images of the girl and the boy with futuristic weapons were present during the instruction video and on the first screen of each block prior to the onset of test trials. To avoid a shift of attention away from the dragon, their images were not present while participants performed the task. Participants were encouraged to respond as fast as possible while also maintaining good accuracy.
Measures of Attention, Language, and Nonverbal Intelligence

All children were administered the following battery of tests. Nonverbal intelligence was assessed with the Test of Non-Verbal Intelligence (4th ed.; TONI-4; Brown, Sherbenou, & Johnsen, 2010). Attention was evaluated with the Parent Rating Scale of the Conners’ Rating Scales—Revised (Conners, 1997), which provides an estimate of risk for developing ADHD based on a parental evaluation of the current behaviors of the child. Linguistic ability was measured with the help of the Clinical Evaluation of Language Fundamentals (4th ed.; CELF-4; Semel, Wiig, & Secord, 2003). Each child was tested on four separate subtests of CELF-4. Taken together, these subtests provide a Core Language Score (CLS), which is a general measure of language skills. The four subtests varied depending on children’s age. All children were administered the Concepts and Following Directions (CFD), Recalling Sentences (RS), and Formulated Sentences (FS) subtests. In addition, 7- to 8-year-olds also completed the Word Structure subtest while 9- to 11-year-olds completed the Word Classes-2 subtest.

Statistical Analyses

We conducted several analyses in order to compare groups’ performance on the SJT. First, we compared children’s and adults’ response times and rates of asynchrony perception at each SOA. These data were analyzed in repeated-measures analyses of variance (ANOVs) with group (7- to 8-year-olds, 10- to 11-year-olds, and adults) as a between-groups variable and SOA (0, 100, 200, 300, 400, and 500 ms) as a within-group variable. Because previous research showed asymmetry in sensitivity to auditory-visual as compared to visual-auditory SOAs (Bushara et al., 2001; Grant et al., 2004; Lewkowicz, 1996; Zemtsov, & Wallace, 2012), more specifically, we used the glmfit function in MATLAB (MATHWORKS, Inc., Natick, MA) in order to fit two sigmoid functions to the number of synchronous responses as a function of SOA. The width of the left (based on the VA SOAs) and the right (based on the AV SOAs) side of each function was then measured in order to fit two sigmoid functions to the number of synchronous responses as a function of SOA.

In all statistical analyses, significant main effects with more than two levels were evaluated with a Bonferroni post hoc test. In such cases, the reported p value indicates the significance of the Bonferroni test, rather than the adjusted alpha level. When omnibus analysis produced a significant interaction, it was further analyzed with step-down ANOVAs or t tests, with factors specific to any given interaction. Mauchly’s test of sphericity was used to check for the violation of sphericity assumption in all repeated-measures tests that included factors with more than two levels. When the assumption of sphericity was violated, we used the Greenhouse-Geisser adjusted p values to determine significance. Accordingly, in all such cases, adjusted degrees of freedom and the epsilon value (\(\epsilon\)) are reported. Effect sizes, indexed by the partial eta squared statistic (\(\eta^2_p\)), are reported for all significant repeated-measures ANOVA results.

Lastly, we conducted a multiple regression analysis with children’s age, the ADHD index, the TONI-4 score, and the CLS score of CELF-4 as predictors and the width of the TBW as a response variable in order to gauge the degree to which these cognitive, language, and attention skills contributed to children’s performance on the SJT. Standardized scores from all of the above tests were used. A stepwise regression was performed to arrive at the final model. Separate regression analyses were conducted on the left (VA) and right (AV) portions of the TBW.

Results

The results of standardized tests of attention, nonverbal intelligence, and language skills are summarized in Table 1. As can be seen, both groups of children were quite comparable, except for the CFD subtest of CELF-4, on which, somewhat surprisingly, the younger group performed better than did the older one.

Perception of Asynchrony at Different SOAs

Percentage of trials at each SOA on which each group perceived synchrony is shown in Figure 2. Additionally, individual data for each of three groups are shown in Figure 3. Further analyses of visual and AV conditions revealed similar results, namely, a significant effect of SOA (VA: F(2.626,115.529) = 351.1, p < .001, \(\eta^2_p = 0.889\), \(\epsilon = 0.525\); AV: F(2,814,123.8) = 459.9, p < .001, \(\eta^2_p = 0.913\), \(\epsilon = 0.563\), group (VA: F(2,44) = 13.363, p < .001, \(\epsilon = 0.378\); AV: F(2, 44) = 31.315, p < .001, \(\epsilon = 0.587\), and a Group \times SOA interaction (VA: F(10, 220) = 12.961, p < .001, \(\epsilon = 0.371\); AV: F(10, 220) = 15.133, p < .001, \(\epsilon = 0.408\)). However, follow-up tests showed that the pattern of group effects at specific SOAs was different for the VA and AV conditions. More specifically, for the VA condition, the effect of group was significant at 200-, 300-, 400-, and 500-ms SOAs (for all SOAs, F(2,46) = 6.466–22.127, p < .001–0.003), with a similar trend for the 100-ms SOA, F(2, 46) = 3.078, p = .056. The Bonferroni test for multiple comparisons showed that adults were better able to detect temporal asynchrony than either group of children at 200- to 500-ms SOAs (p = < .001–0.039) and were marginally better than 7- to 8-year-olds at the 100-ms SOA (p = .057). Importantly, how-
ever, the two groups of children did not differ from each other at any VA SOA (p = .109–1.0). For the AV condition, the effect of group was significant for the same range of SOAs as in the VA analysis: 200, 300, 400, and 500 ms (for all SOAs, F(2, 46) = 18.583–42.124, p < .001). The Bonferroni test for multiple comparisons revealed that adults, again, were more sensitive to temporal asynchrony at the 200- to 500-ms SOAs (p = .001 to <.0001) than either group of children. However, in contrast to the VA analysis, the older group of children was also better able to detect temporal asynchrony than their younger counterparts at 300-, 400-, and 500-ms SOAs (p = .001–0.018) (see Figure 2).

Groups also differed significantly in responses to unimodal stimuli: auditory, F(2, 46) = 7.385, p = .002; visual, F(2, 46) = 7.995, p = .001, with adults being significantly more accurate in response to both V and A stimuli than the 7- to 8-year-old children (both ps = .001) and marginally more accurate in response to the V stimuli than the 10- to 11-year-old children (p = .064). The two groups of children did not differ from each other (auditory, p = .218; visual, p = .34).

### Response Time at Different SOAs

Participants’ RT data are shown in Figure 4. The omnibus analysis of RT showed a significant effect of SOA (VA: F(2,401,105.65) = 75.666, p < .001, ηp² = 0.632, ε = 0.48; AV: F(2,95,129.797) = 54.108, p < .001, ηp² = 0.552, ε = .59), group (VA: F(2, 44) = 24.19, p < .001, ηp² = 0.524; AV: F(2, 44) = 24.337, p < .001, ηp² = 0.525), and a Group × SOA interaction (VA: F(10, 220) = 8.084, p < .001, ηp² = 0.269; AV: F(10, 220) = 9.121, p < .001, ηp² = 0.293). In regard to the effect of group, follow-up tests revealed that adults’ responses were significantly faster than those of either group of children (adults vs. 7- to 8-year-olds, p < .001; adults vs. 10- to 11-year-olds, p < .001), which held true at all SOAs (all ps < .001). At the same time, the two groups of children did not differ from each other at any SOA in either the VA or the AV condition (all ps = 0.187–1). The most striking difference between children and adults was the relationship between the SOA and RT: In adults, the RT first increased for short SOAs compared to synchrony and then decreased as the SOA became longer. In children, on the other hand, there was no noticeable shortening of RT at longest SOAs (see Figure 4).

In order to better describe this difference, we compared the RT to the 200-ms SOA (which, at least in absolute terms, required the longest time to respond to in adults), the RT to true synchrony, and the RT to the longest SOA of 500 ms in each group. In adults, the RT to the 200-ms SOA was significantly longer than that to both the synchronous trials and the 500-ms SOA trials (SOA, VA: F(1,831,27.465) = 28.752, p < .001, ηp² = 0.657, ε = 0.366; VA200 vs. SYNC, p < .001; VA200 vs. VA500, p = .042; SOA, AV: F(2,518,37.775) = 12.205, p < .001, ηp² = 0.449, ε = 0.504; AV200 vs. SYNC, p < .001, AV200 vs. AV500, p = .043). In 7- to 8-year-old children, the RT to the 200-ms SOA trials was longer than to synchronous trials but shorter than to the 500-ms SOA trials (SOA, VA: F(2,465,34.513) = 33.629, p < .001, ηp² = 0.706, ε = 0.493; AV200 vs. SYNC, p = .003; VA200 vs. VA500, p = .01; SOA, AV: F(5, 70) = 31.592, p < .001, ηp² = 0.693;
AV200 vs. SYNC, \( p < .001 \), \( \eta^2_p = 0.681 \), \( \varepsilon = 0.476 \); AV200 vs. AV500, \( p = 1 \).

Comparing Sensitivity to AV Versus VA SOAs

The order of modalities had a significant effect on the overall number of synchronous perceptions, \( F(1, 44) = 10.342, p = .002 \), \( \eta^2_p = 0.19 \). This effect interacted with SOA, \( F(2.808, 123.545) = 5.353, p = .002 \), \( \eta^2_p = 0.09 \). Follow-up paired-samples t tests revealed that VA stimuli elicited a significantly higher number of synchronous perceptions at the 200-, 300-, and 400-ms SOAs, \( t(46) = 2.305–3.276, p = .026–0.002 \), two-tailed. There was no Group × Order interaction, \( F(2, 44) = 1.088, p = .346 \), \( \eta^2_p = 0.047 \).

Analysis of response time as a function of the order of modalities yielded a significant Order × SOA × Group interaction, \( F(8, 176) = 2.169, p = .042 \), \( \eta^2_p = 0.09 \). Further analyses showed that adults took longer to respond to VA compared to AV stimuli at the 200-, 300-, and 400-ms SOAs, \( t(15) = 2.516–2.813, p = .013–0.024 \), two-tailed, with a trend in the same direction at the 500-ms SOA, \( t(15) = 1.862, p = .082 \), two-tailed. In contrast, neither group of children showed asymmetry in response time to AV versus VA stimuli at any SOA (7- to 8-year-olds, \( t(14) = −0.453–1.547, p = .144–0.658 \), two-tailed; 10- to 11-year-olds, \( t(15) = −0.729–0.028, p = .477–0.996 \), two-tailed).

Size of the TBW

The TBWs for each group at 75% and 50% of the function’s maximum are shown in Figure 5. Analysis of the TBW at the 75% of the maximum yielded a significant effect of group only over the

AV200 vs. SYNC, \( p = .001 \); AV200 vs. AV500, \( p = .017 \). In other words, their RT continued to increase for longer SOAs. Finally, 10- to 11-year-old children fell in between the other two groups: their RT to the 200-ms SOA trials was longer than to synchronous trials but did not differ from that to the 500-ms SOA trials (SOA, VA: \( F(2.196,32.944) = 28.549, p < .001 \), \( \eta^2_p = 0.656 \), \( \varepsilon = 0.439 \); VA200 vs. SYNC, \( p < .001 \); VA200 vs. VA500, \( p = 1 \); SOA, AV: \( F(2.381,35.711) = 32.027, p < .001 \), \( \eta^2_p = 0.681 \), \( \varepsilon = 0.476 \); AV200 vs. SYNC, \( p < .001 \); AV200 vs. AV500, \( p = 1 \)).

Figure 3. Individual differences in sensitivity to audiovisual asynchrony. Percentage of synchronous responses provided by each participant in each group is shown for all SOAs. Note a significant reduction in individual variability of responses at the two longest SOAs (400–500 ms) in adults compared to either group of children. Also of note is a high degree of individual variability in adults at the 200- and 300-ms SOAs. SOA = stimulus onset asynchrony; VA = visual stimulus appears first; AV = auditory stimulus appears first. See the online article for the color version of this figure.

Figure 4. Response time at different SOAs. Time for making a simultaneity judgment at each SOA is shown for the three groups of participants. Error bars represent standard errors of the mean. Note a qualitatively different pattern of RT distribution over different SOAs in children and in adults, with adults’ RTs peaking at 200-ms SOA and then shortening at 500-ms SOA and with children’s RTs either continuously increasing (the 7- to 8-year-old group) or leveling off after the 200-ms SOA (the 10- to 11-year-old group). SOA = stimulus onset asynchrony; RT = response time; VA = visual stimulus appears first; AV = auditory stimulus appears first. See the online article for the color version of this figure.
right (AV) side of the TBW—right: $F(2, 41) = 3.576, p = .037$; left: $F(2, 41) = 1.142, p = .33$—with the 7- to 8-year-old group having a significantly larger right TBW than adults (7- 8-year-olds vs. adult, $p = .039$). No other pairwise comparison was significant (7- to 8-year-olds vs. 10- to 11-year-olds, $p = .2$; 10- to 11-year-olds vs. adult, $p = 1$).

Analysis of the TBW at 50% of the maximum showed a different pattern of group differences. Namely, groups differed significantly over both the left and the right sides of the TBW, left, $F(2, 45) = 12.883, p < .001$; right, $F(2, 45) = 33.006, p < .001$. Pairwise comparisons revealed that adults’ left and right TBW were significantly smaller than those of either group of children (left, 7- to 8-year-olds vs. adults, $p < .001$; 10- to 11-year-olds vs. adults, $p = .007$; right, 7- to 8-year-olds vs. adults, $p < .001$, 10- to 11-year-olds vs. adults, $p < .001$); however, the two groups of children differed only over the right side of the TBW (right, 7- to 8-year-olds vs. 10- to 11-year-olds, $p = .009$; left, 7- to 8-year-olds vs. 10- to 11-year-olds, $p = .214$), with a larger TBW in 7- to 8-year-olds.

**Multiple Regressions**

Only one relationship was found to be significant enough to be included in the model—namely, age significantly predicted the size of the right TBW when the latter was measured at 50% of the function’s maximum. More specifically, the size of the right TBW became increasingly smaller with age, $F(1, 29) = 10.761, p = .003$; $B = -17.54$, $SE = 5.347$, $\beta = -.527$. On the basis of $R^2$ (0.278), age accounted for approximately 28% of variance in the size of the right TBW (see Figure 6).

**Discussion**

We examined audiovisual temporal function in two groups of children with narrow age ranges (7- to 8-year-olds and 10- to 11-year-olds) and in adults by using an SJT. Several major differences between children and adults and between the two groups of children have emerged.

First, both groups of children lagged significantly behind adults in the ability to detect audiovisual temporal asynchrony at all SOAs exceeding 100 ms, in both VA and AV conditions. This finding extends an earlier report from my laboratory, which showed that 7- to 8-year-old children were less sensitive to audiovisual temporal asynchrony compared to adults. Because in our previous study the children’s group included a relatively broad range of ages, it was possible that the effect was carried mostly by younger children. The results of the current study show unequivocally that by 10–11 years of age, children are still much more likely than adults to perceive asynchronous events as synchronous, even at SOAs as large as 400–500 ms. This result is significantly more dramatic than that reported by Hillock-Dunn et al. (2011), who showed that 10- to 11-year-old children differed from adults only over the 150- to 350-ms SOAs and only in the VA condition. One potentially significant difference between our studies is the length of the stimuli. In Hillock-Dunn et

![Figure 5](https://example.com/figure5.png)  
*Figure 5.* Temporal binding windows (TBW) measured at 75% and 50% of the maximum. Group mean data for the left and right sides of the TBW window at 75% (the left panel) and 50% (right panel) of the maximum are shown. The size of the left TBW at 75% of the maximum was nearly identical for 7- to 8-year-old and 10- to 11-year-old children. As a result, their drop-down lines overlap. Yos = year olds; SOA = stimulus onset asynchrony; VA = visual stimulus appears first; AV = audio stimulus appears first. See the online article for the color version of this figure.
al.’s reports, the length of the ring flash and the sound tone was only 8–10 ms. As a result, the onsets of auditory and visual stimuli never overlapped. Our stimuli were 200 ms in length and, therefore, at the SOAs of 100 and 200 ms there was no separation between the offset of the first stimulus and the onset of the second one in each audiovisual pair. And furthermore, even at longest SOAs, the offset to onset separations were significantly smaller than those in Hillock-Dunn and colleagues’ studies. This might have made the simultaneity judgment more difficult. However, since many events in real life continue for more than 10 ms, children’s performance with longer stimuli is highly informative and suggests that the maturation of the audiovisual temporal function may be even more prolonged than previously reported.

Second, children’s ability to detect audiovisual temporal asynchrony undergoes a significant change during middle childhood years. This change is asymmetrical in nature, with 10- to 11-year-old children surpassing their 7- to 8-year-old counterparts at longest SOAs (300–500 ms), but only when the auditory stimulus precedes the visual one (i.e., in the AV condition). In other words, sensitivities to VA and AV asynchronies have different developmental trajectories, with AV SOAs eliciting more accurate discriminations by 10–11 years of age. However, the age at which sensitivity to VA SOAs begins to improve compared to early school years remains to be determined. This finding dovetails with a well-documented fact in adult literature showing that, overall, adults are more sensitive to AV asynchronies than to VA ones (Bushara et al., 2001; Dixon & Spitz, 1980; Grant et al., 2004; Lewkowicz, 1996; McGrath & Summerfield, 1985; van Wassenhove et al., 2001). On the basis of our results, we can conclude that such enhanced sensitivity to AV asynchronies emerges during middle childhood. It is worth noting that we have replicated the earlier finding of asymmetry in audiovisual temporal function, with overall fewer perceptions of synchrony reported for AV as compared to VA SOAs at medium (200–400 ms) SOAs. In adults, reduced sensitivity to VA temporal offsets was accompanied by increased RT.

We also compared sensitivity to audiovisual temporal asynchrony in children and adults by using a single metric—the TBW—which helps control for a potential bias in groups’ responses to individual modalities. When the width of the TBW was measured at 50% of the function’s maximum, this analysis yielded results that were completely parallel to those obtained from the analyses of individual SOAs. More specifically, in adults, both the left and the right sides of the TBW were significantly smaller than in either group of children, and in 10- to 11-year-olds the right (AV) side of the TBW was significantly smaller than in 7- to 8-year-olds.

Third, compared to children, adults were overall faster to make a simultaneity judgment, and, most importantly, had a qualitatively different relationship between RT and SOAs (see Figure 4). More specifically, in adults, the RT first increased with increasing SOA, peaking at the 200-ms offset, then shortened significantly at the 500-ms SOA. In children, on the other hand, the RT either continued to increase all through the longest SOA (7- to 8-year-old group) or levelled off at SOAs exceeding 200 ms (10- to 11-year-old group). While behavioral measures do not tell us whether SOAs that require the longest RT present perceptual or cognitive (i.e., decision making) challenge or both, these findings clearly show that different SOAs are associated with different degrees of certainty in regard to their audiovisual synchronicity, with the 200- to 300-ms SOAs requiring the longest time to respond to (at least in adults) and thus representing perhaps the most ambiguous stimulus. Interestingly, the same SOAs also tended to have the largest individual variability in adults (see Figure 3). This pattern of results suggests that children found making a simultaneity judgment effortful even when presented with the 400- to 500-ms SOAs. We take this finding to indicate that these large separations between modalities did not lead to a clear perception of asynchrony in children—a conclusion that is in agreement with the observed higher number of synchronous perceptions in these groups. The fact that in 10- to 11-year-olds, the RTs leveled off at longer SOAs (rather than continuing to increase as in the younger group of children), may signal that their pattern of RT distribution across various SOAs is beginning to approach that of adults, with RT to the longest SOAs being in the process of shortening (compared to the 200-ms SOA).

Fourth, somewhat surprisingly, none of the standardized tests were significantly correlated with children’s performance on the SJT, with age being the only significant predictor. Furthermore, age predicted only the width of the right (AV SOAs) TBW, underlining the fact that sensitivity to AV SOAs undergoes a more profound change during school years than sensitivity to VA SOAs. This finding is in general agreement with the study by Hillock-Dunn and Wallace (2012), who reported that children’s performance on the SJT did not correlate with verbal or nonverbal IQ, reading ability, or socioeconomic status but did correlate with age. Therefore, while we know that sensitivity to multisensory temporal asynchrony gradually increases during middle childhood and is impaired in a number of developmental disorders (for a comprehensive review of this issue, see Wallace & Stevenson, 2014), such as autism (Foss-Feig et al., 2010; Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011; Stevenson et al., 2014), specific language impairment (SLI) (Grondin et al., 2007; Kaganovich et al., 2014), and dyslexia (Hairston, Burdette, Flowers, Wood, & Wallace, 2005), the functional significance of multisensory temporal function for
language and, more generally, for cognitive development is poorly understood and requires future study. One possibility for the lack of a relationship between children’s performance on the SJT and other cognitive skills in our data. Namely, given the amount of individual variability in sensitivity to audiovisual temporal asynchrony, even in adults, it is possible that poor audiovisual temporal function does not interfere with typical development until it reaches a certain threshold — a threshold that we would rarely see crossed in typically developing children but that could be exceeded by clinical populations. Some support for this hypothesis comes from an earlier study from my laboratory, which showed a significant correlation between language skills and the SJT performance in children with a history of Specific Language Impairment (H-SLI) (Kaganovich et al., 2014), with those children who did well on the SJT also having better language scores. This relationship did not hold for age-matched typically developing children, who, as a group, performed markedly better on the SJT compared to their H-SLI peers. In other words, as long as sensitivity to audiovisual temporal information stays within a certain range, however broad, it may be sufficient for proper cognitive development. If so, determining the boundaries of this acceptable variability range will be an important empirical question to address by future work.

In the current study, we used simple nonspeech stimuli in part so we could easily compare our findings with earlier developmental reports (which used simple stimuli as well) and in part because they allowed us to convert the SJT paradigm into a game suitable for testing school-age children. While we acknowledge that the use of more ecologically valid stimuli should be considered in future work, we believe that the use of simple stimuli does not diminish the significance of our findings. Recent work shows that measures of sensitivity to audiovisual asynchrony in simple and complex events are correlated (Stevenson & Wallace, 2013), and that cognitive functions that are affected by audiovisual temporal processing (such as perception of degraded speech) are correlated with audiovisual temporal function in both speech and nonspeech stimuli (Conrey & Pisoni, 2006).

In conclusion, the results of the current study provided several important insights into the development of audiovisual temporal function during midchildhood. More specifically, we showed that by 10–11 years of age, children are still significantly more likely than adults to perceive asynchronous audiovisual events as synchronous, even at SOAs as large as 400–500 ms. In agreement with this finding, only adults showed a significant shortening of RT at long SOAs compared to medium SOAs. We also found that audiovisual temporal function undergoes a significant change between approximately 8 and 11 years of age. This change is asymmetrical in nature, with sensitivity to AV asynchronies improving and sensitivity to VA asynchronies remaining constant. As mentioned above, a reduced sensitivity to audiovisual temporal asynchrony has been reported for a number of developmental disorders, suggesting that a timely maturation of audiovisual temporal processing may play a significant role in cognitive and linguistic development. Our results may serve as an important baseline against which children with developmental disorders associated with impaired audiovisual temporal function may be compared.

References

Keetels, M., & Vroomen, J. (2012). Perception of synchrony between the
senses. In M. M. Murray & M. T. Wallace (Eds.), The neural bases of multisensory processes (pp. 147–177). New York: CRC Press.


