Chapter 4

Study 3: Temporally Incongruent Audiovisual Stimuli

4.1 Overview

Just as spatial co-location provides a strong influence on the perception of a unified multisensory event, temporal synchrony is another factor that frequently facilitates integration. The ability to be able properly integrate the inputs from multiple sensory systems requires a certain latitude in the arrival of the pertinent information. The stimulus input to the brain—from arrival at the different sensory receptors to the execution of a motor response—may have different propagation rates due to environmental constraints (speed of light versus speed of sound) as well as differing processing rates depending on the levels cortical processing required. As such, it makes sense that the perception of synchronicity not strictly require exact, physical coincidence, but that inputs be *near enough.* Multisensory neurons in the superior colliculus of cats have been found to have a ”window of opportunity” where the synchronous or near-synchronous onset of auditory and visual stimuli produced an enhanced firing rate, falling off monotonically with temporal disparities beyond this range, and even producing a depressed firing rate for very large differences (Meredith et al., 1987). On a behavioral level, the relative timing of multimodal onsets have been found to produce a variety of asymmetrical results. The stream-bounce illusion can be more strongly influenced as a bounce when the sound is presented 150 ms before or coincident with intersection, but more weakly if presented 150 ms post-intersection, and sufficiently large delays between auditory onset and intersection strongly influence the perception in the reverse direction (Sekuler et al., 1997). Subjects watching a video of hammering or a person speaking English words, with the audio track leading or lagging the video track showed asymmetrical and differential tolerances for asynchrony; shorter discrepancies were noticed for non-language trials (75 ms leading or 175 ms lagging) than for language trials (130 ms leading and 250 ms lagging) (Dixon & Spitz, 1980). What is somewhat surprising is the magnitude of temporal disparity that can be tolerated while still producing behavioral effects.

Infants between the age of two and ten months of age were presented with temporally congruent and incongruent (150, 300, 450, and 600 ms) auditory and visual stimuli at ±25°. We found significantly increased response latencies for all temporal disparity conditions, as well as clear differences due to the modality of the leading target and to a lesser extent the age of the subjects. When the visual stimulus came on first, infants in all age groups were relatively insensitive to the magnitude of the temporal disparity but when the auditory stimulus came on first, there was a proportional increase in response latency with increasing temporal disparity, even for the longest delays (600 ms). We found strong indications that this difference is due largely to the relative capabilites of the orienting response to the leading stimulus modality. Similar profiles were found in adults. There were also indications of developmental changes in the magnitude of percentage increase in response latency from the synchronous condition, peaking around 6–8 months of age.

4.2 Experimental Design and Methods

4.2.1 Participants

Participation and recruitment was the same as in Study 1, with a few differences. Four age-groups were defined, a priori: 2–4, 4–6, 6–8, and 8–10 months of age and

Table 4.1: Total number of subjects, their mean age and gender ratio, and total percentage of valid trials (see *Methods*) for each age group and experiment. A subset of infants in each age group completed only one of the two experiments (last row).

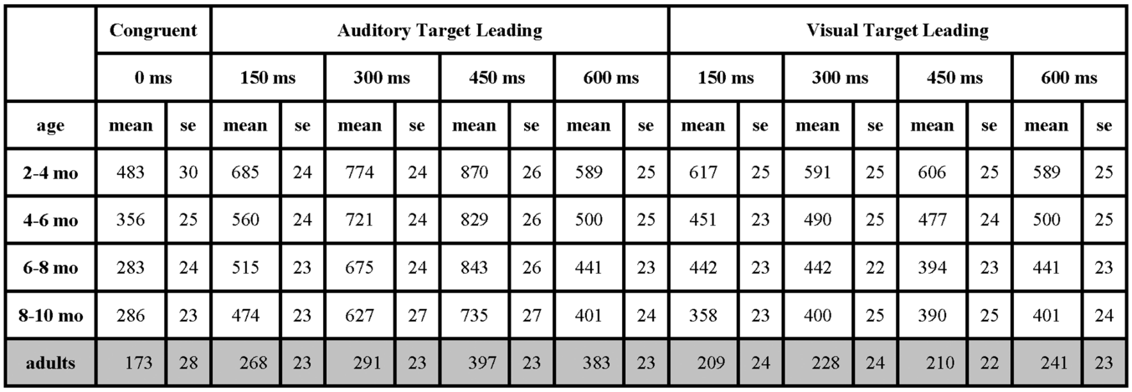
Subject Info \ Age Group		**2-4 mo**	**4-6 mo**	**6-8 mo**	**8-10 mo**	**adults**
Experiment 2	**N**	15	15	15	15	5
	Mean Age	3.04 ± 0.6	5.07 ± 0.4	6.64 ± 0.6	8.86 ± 0.4	25 ± 7 yrs
	Male/Female	7/8	8/7	7/8	9/6	2/4
	% good trials	55.7%	63.0%	65.7%	62.0%	87.5%
Experiment 4	**N**	15	15	15	15	5
	Mean Age (months)	3.15 ± 0.6	5.11 ± 0.5	7.03 ± 0.5	9.08 ± 0.6	25 ± 7 yrs
	Male/Female	7/8	7/8	8/7	5/10	2/4
	% good trials	61.7%	61.2%	65.2%	59.5%	76.5%
# of subjects to complete both exps versus only one exp		1/14	1/14	2/13	0/15	5/5

during each visit, infants were tested with temporally congruent (Experiment 2) and temporally incongruent (Experiment 4) audiovisual stimuli. From the pool of valid candidates, 15 infants were randomly selected for each age group within each experiment; not all infants completed both. Based on these criteria, a total of 53 individual infants (26 male, 27 female), ranging between the ages of 2.07 and 9.89 months, participated in this study, with some infants participating at more than one age group. Repeat participation was as follows: only one age group (21 infants), two age groups (10 infants), three age groups (13 infants), and all four age groups (9 infants). Five adults (1 male, 4 females) also participated in this study (Table 4.1).

4.2.2 Apparatus and Stimuli

The experimental apparatus and stimuli used were the same as in Study 1. The temporally congruent experiment (Experiment 2) had both the auditory and visual components presented synchronously at the same location, at ±25°. For the temporally asynchronous experiment (Experiment 4), the auditory and visual components

Table 4.2: The group mean was calculated for each temporal disparity and leading target modality for each age group.



	Congruent		Auditory Target Leading								Visual Target Leading							
	0 ms		150 ms		300 ms		450 ms		600 ms		150 ms		300 ms		450 ms		600 ms	
age	**mean**	**se**	**mean**	**se**	**mean**	**se**	**mean**	**se**	**mean**	**se**	**mean**	**se**	**mean**	**se**	**mean**	**se**	**mean**	**se**
2-4 mo	483	30	685	24	774	24	870	26	589	25	617	25	591	25	606	25	589	25
4-6 mo	356	25	560	24	721	24	829	26	500	25	451	23	490	25	477	24	500	25
6-8 mo	283	24	515	23	675	24	843	26	441	23	442	23	442	22	394	23	441	23
8-10 mo	286	23	474	23	627	27	735	27	401	24	358	23	400	25	390	25	401	24
adults	173	28	268	23	291	23	397	23	383	23	209	24	228	24	210	22	241	23

were presented at the same location (±25°) but offset by ±150 ms, ±300 ms, or ±450 ms, producing 16 target conditions.

4.2.3 Procedure

Procedures for trial presentation were the same as in Study 1. Response latencies were always calculated from the onset of the leading target stimulus to the onset of the orienting response. Five blocks of trials were presented in Experiment 2 (temporally synchronous), for a total of ten trials, and six blocks of trials were presented in Experiment 4 (temporally asynchronous)—with a short break after three blocks—for a total of 96 trials. The order of target presentation was randomized within each block. All good trials, for both Experiment 2 and Experiment 4, were pooled within each age group for purposes of statistical analysis (Table 4.2).

4.3 Results

We began our analysis of the response latency data with a 3-way univariate ANOVA (2 x 4 x 4), with the modality of the leading target (A_{lead} or V_{lead}) and absolute temporal disparity between target onsets (150, 300, 450, and 600 ms) as within-subject factors, and age group (2–4, 4–6, 6–8, and 8–10 months) as between-subject factors, using a value of $p = 0.03$ for the significance threshold. We found significant main effects for modality, $F_{(1,3338)} = 932.9, p < 0.001$), absolute temporal disparity, $F_{(3,3338)} = 72.3, p < 0.001$), and age, $F_{(3,3338)} = 70.2, p < 0.001$). There were also significant disparity x modality, $F_{(3,3338)} = 64.5, p < 0.001$), and modality x age, $F_{(3,3338)} = 5.6, p = 0.001$) interactions. In adults, a 2-way univariate ANOVA (2 x 4) was also performed, with modality and absolute temporal disparity as factors. We found significant main effects for absolute temporal disparity, $F_{(3,359)} = 4.7, p = 0.003$), and modality, $F_{(1,359)} = 47.5, p < 0.001$), as well as the 2-way modality x disparity interaction, $F_{(3,359)} = 3.6, p = 0.013$).

4.3.1 Main Effect of Leading Target Modality

Pooled across all disparity conditions and infant age groups, the main effect for modality was due to the grand mean response latency for auditory-leading trials (RT_{AV}) being approximately half again as long as visual-leading (RT_{VA}) ($RT_{AV} = 739 \pm 6$ ms; $RT_{VA} = 474 \pm 6$ ms). Adults also had a shorter response latency for RT_{VA} (222 ± 12 ms) than RT_{AV} (345 ± 12 ms).

4.3.2 Main Effect of Age

A repeated planned contrast analysis was then performed, comparing the response latency between each age group and the next oldest. These comparisons indicated that there was a significant difference in latency between all four age groups ($p \leq 0.006$), which is a slight deviation from the previous two studies where the two older age groups had not been significantly different (Figure 4.1a). A paired contrast between the oldest infant age group and adults was also significant ($p < 0.001$).

4.3.3 Main Effect of Absolute Temporal Disparity

In order to examine the main effect for absolute temporal disparity, multiple *post-hoc* comparisons were performed using the Games-Howell procedure, due to our unequal sample size with non-homogenous variances (homogeneity of variances tested using

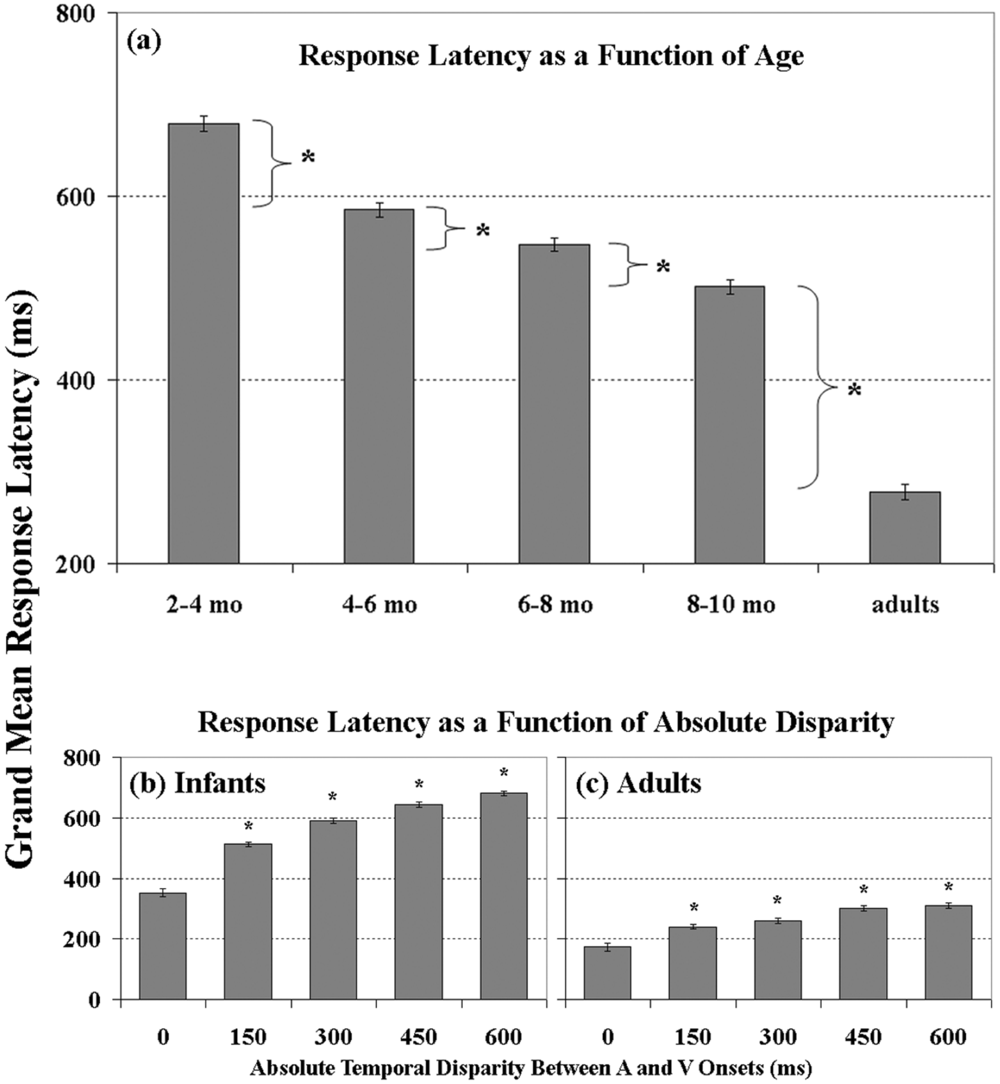


Figure 4.1: Response latencies as a function of (a) age and the absolute temporal disparity between the onsets of the auditory and visual stimuli in (b) infants and (c) adults. Error bars represent the standard error of the mean.

Levenes Test of Equality of Variances, $p < 0.001$), the results of which were a significant difference ($p < 0.03$) between all disparity conditions in infants (Figure 4.1b). In adults, *post-hoc* comparisons between disparity conditions found a significant effect only between the shortest and the largest (150 and 600 ms) disparities ($p = 0.026$). All subjects, infants and adults, showed a significant difference ($p < 0.001$) for all disparities from the temporally congruent condition (Figure 4.1b,c).

4.3.4 Temporal Disparity and Modality Interactions

Breaking the data up by age group, and looking more closely at the effects of temporal disparity as a function of the modality of the leading target, it was revealed that the differences between non-zero disparity conditions was due largely to A_{lead} trials (Figure 4.2). One-way univariate ANOVAs (5) for disparity (0, 150, 300, 450, and 600 ms) found significant main effects for all infants, as well as for adults ($p < 0.001$). *Post-hoc* multiple comparisons between all disparity and non-disparity conditions found three different patterns for 2–4 months, 4–10 months, and adults. In the youngest infants, the significant differences in response latencies fell into three disparity ranges: (0 ms $< 150 - 300$ ms $< 450 - 600$ ms). For all the older infants, the differences were more finely tuned into four ranges: (0 ms $<$ 150 ms $<$ 300 ms $< 450 - 600$ ms). Adults had a slightly modified range than either groups of infants: (0 ms $< 150 - 300$ ms $\leq 450 < 600$ ms). Conversely, response latencies for V_{lead} trials, when plotted independently, had a flat profile for all non-zero disparities (Figure 4.3). A 1-way ANOVA (5) for disparity found significant main effects for all infants, as well as for adults ($p \leq 0.003$) but this time the *post-hoc* multiple comparisons revealed a much different case from A_{lead}; all non-zero disparity latencies were significantly longer than the synchronous case, but not different between each other: (0 ms $< 150 - 600$ ms).

4.3.5 Sensory System Maturity

Since the durations of each component of the stimulus used in this study were not independent of each other—the first stimulus was presented, than after an appropriate

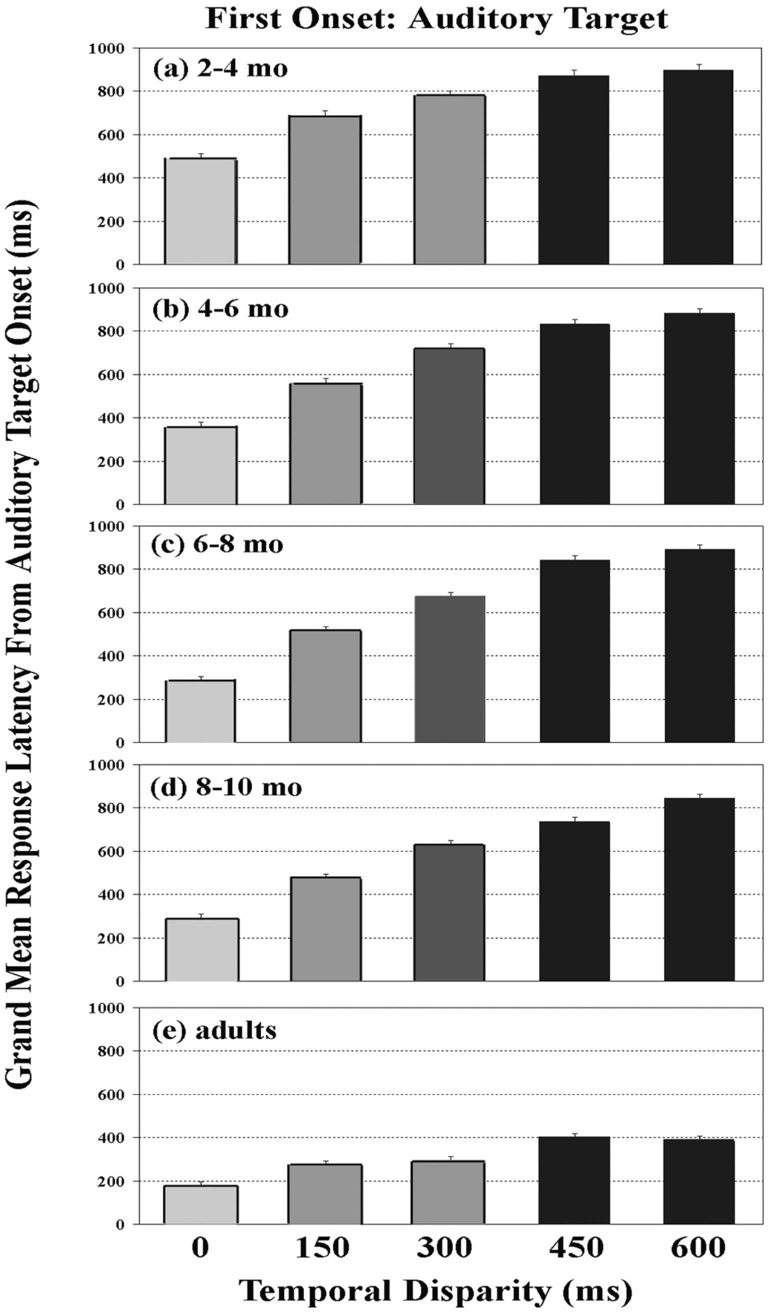


Figure 4.2: Response latencies as a function of the temporal disparity between visual and auditory targets, when the auditory target was on first, for (a) 2–4, (b) 4–6, (c) 6–8, (d) 8–10 month olds and (e) adults. Error bars represent the standard error of the mean.

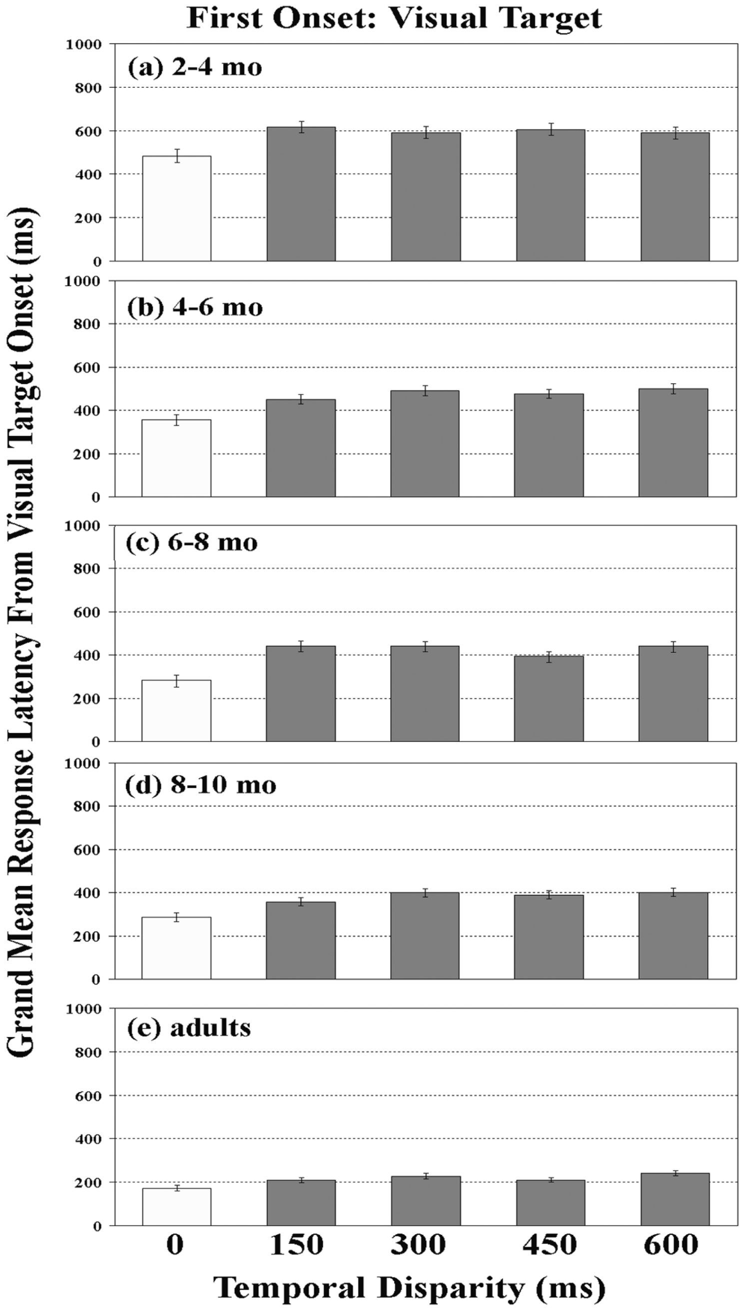


Figure 4.3: Response latencies as a function of the temporal disparity between visual and auditory targets, when the visual target was on first, for (a) 2–4, (b) 4–6, (c) 6–8, (d) 8–10 month olds and (e) adults. Error bars represent the standard error of the mean.

delay (150, 300, 450, or 600 ms), the second stimulus was turned on and both kept on until baby oriented—a better way to think about the stimulus presentation is less that of independent auditory and visual targets with distinct onsets, but rather as an audiovisual target that is preceeded by either an auditory-only or a visual-only target by variable temporal delays. From this reference point, an infant's response to the asynchronous event (the period of time including both auditory and visual onsets) could be dictated by one of several possible scenarios.

On one extreme, the onsets of the auditory and visual stimuli could be sufficiently far enough separated in time that they are no longer associated as a single event, but separate, sequential events, where the secondary target provides neither help nor hindrance in the orienting response. In this case, the response would be indistinguishable from the mean response latency to a unimodal target (unimodal dominance). At the other end of the spectrum, the disparity could be sufficiently short that the infant is incapable of distinguishing the incongruent from the congruent condition and the mean response latency would be comparable to that of a synchronous audiovisual target (bimodal dominance). In between these two extremes would have the baby start a response to the leading stimulus but have the onset of the lagging stimulus disrupt the response in some way, either to speed up or slow down the response. The most simplistic model one could assume, a "no interference" model, would be that an infant's performance is neither facilitated nor inhibited by a temporally disparate target, but rather flips from a response driven by the unimodal, leading stimulus to a response driven by the bimodal stimulus (defined by the onset of the lagging component) as the magnitude of the disparity drops below some threshold. An inhibited (or facilitated) response, however, would result in response latencies that were slower (or faster) than could be expected by the "no interference" model.

4.3.5.1 Visual Target Leading

The apparent insensitivity of the V_{lead} latencies at all infant groups, irregardless of the temporal disparity, suggests that when a visual target is presented first, subjects respond to it rather than to the total temporally disparate event. An analysis of the number of total V_{lead} trials with a reaction time after the visual onset but before the auditory onset (mid-event response) found a fair percentage of responses in all infant groups were occuring before the second stimulus was even presented (Figure 4.4b). This is consistent with the results found from Study 1; mean response latencies were found for visual-only targets at 25° with values that fell within the range of temporal disparities for the older three infant groups (4–10 months) and adults (Figure 2.2). However, the onset of the second stimulus does seem to have some effect. Percent differences calculated between the temporally disparate and leading component's unimodal (Study 1) mean latencies—normalized by the temporally congruent audiovisual mean latency—revealed values that were longer than would be expected (most age groups) if the babies were responding to just the leading stimulus without interference from the lagging stimulus (Figure 4.5). Distributions of the reaction times for all V_{lead} trials across all temporal disparity conditions (not shown) have only one peak at all age groups, and not two as might be expected if this slower response were only due to a subset of trials where the response was after the second (auditory) stimulus onset. So, while the response in infants seems to be instigated by the leading visual component, the onset of a delayed auditory component—by even large temporal disparities—seems able to slow the mean response latency one would get for a synchronous audiovisual event, and in the oldest infants and adults, is slower than even an independent unimodal event.

4.3.5.2 Auditory Target Leading

The completely different profile for conditions where the auditory target came on first suggests something else is going on. Unlike the V_{lead} condition, for A_{lead} trials the mean response latencies increase as a function of increasing temporal disparity. It's therefore unlikely the response is being dominated by the auditory stimulus, as is likely the case for the visual-leading trials. Far fewer of the reaction times to A_{lead} trials occurred before the onset of the second stimulus, only a maximum of 20% for the largest disparities and oldest infants (Figure 4.4a). Two possible situations could

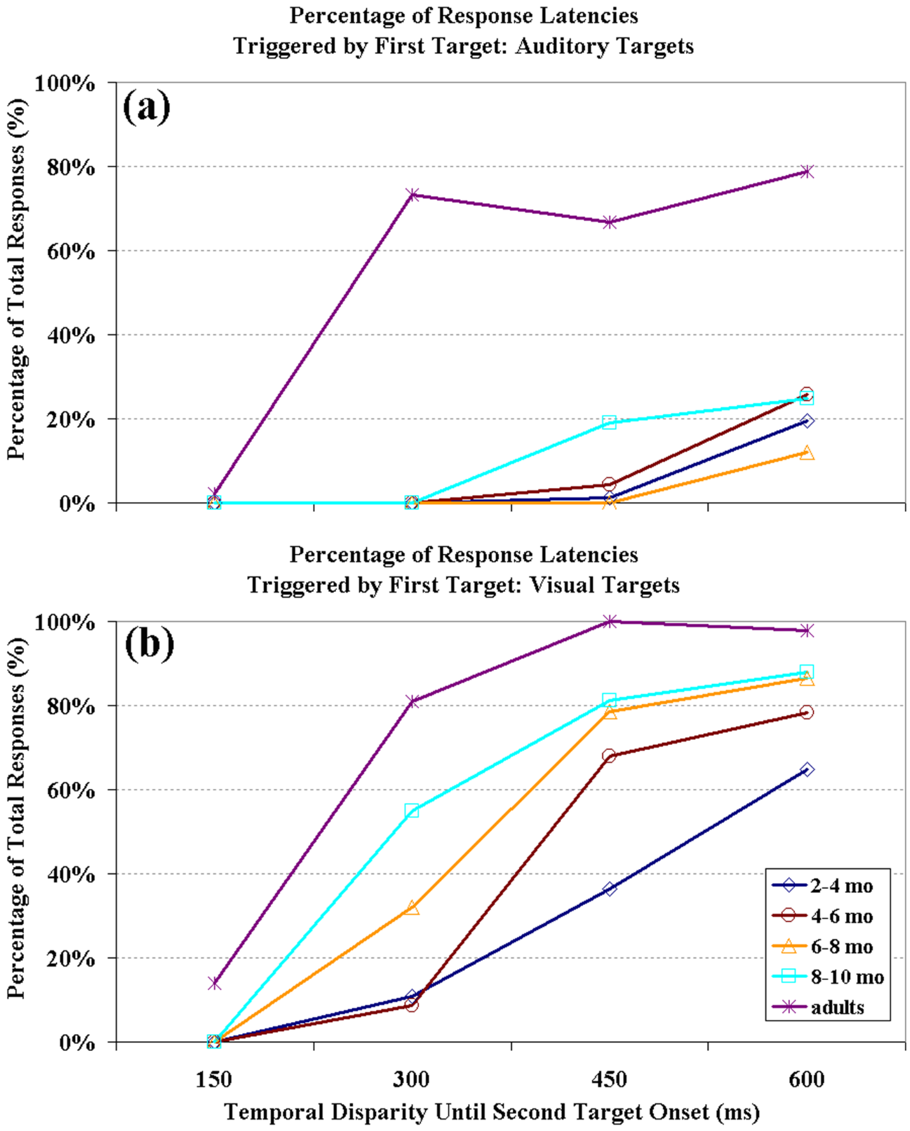


Figure 4.4: The percentage of trials that occurred before the onset of the second target with increasing temporal disparity for (a) A_{lead} and (b) V_{lead} conditions in infants and adults.

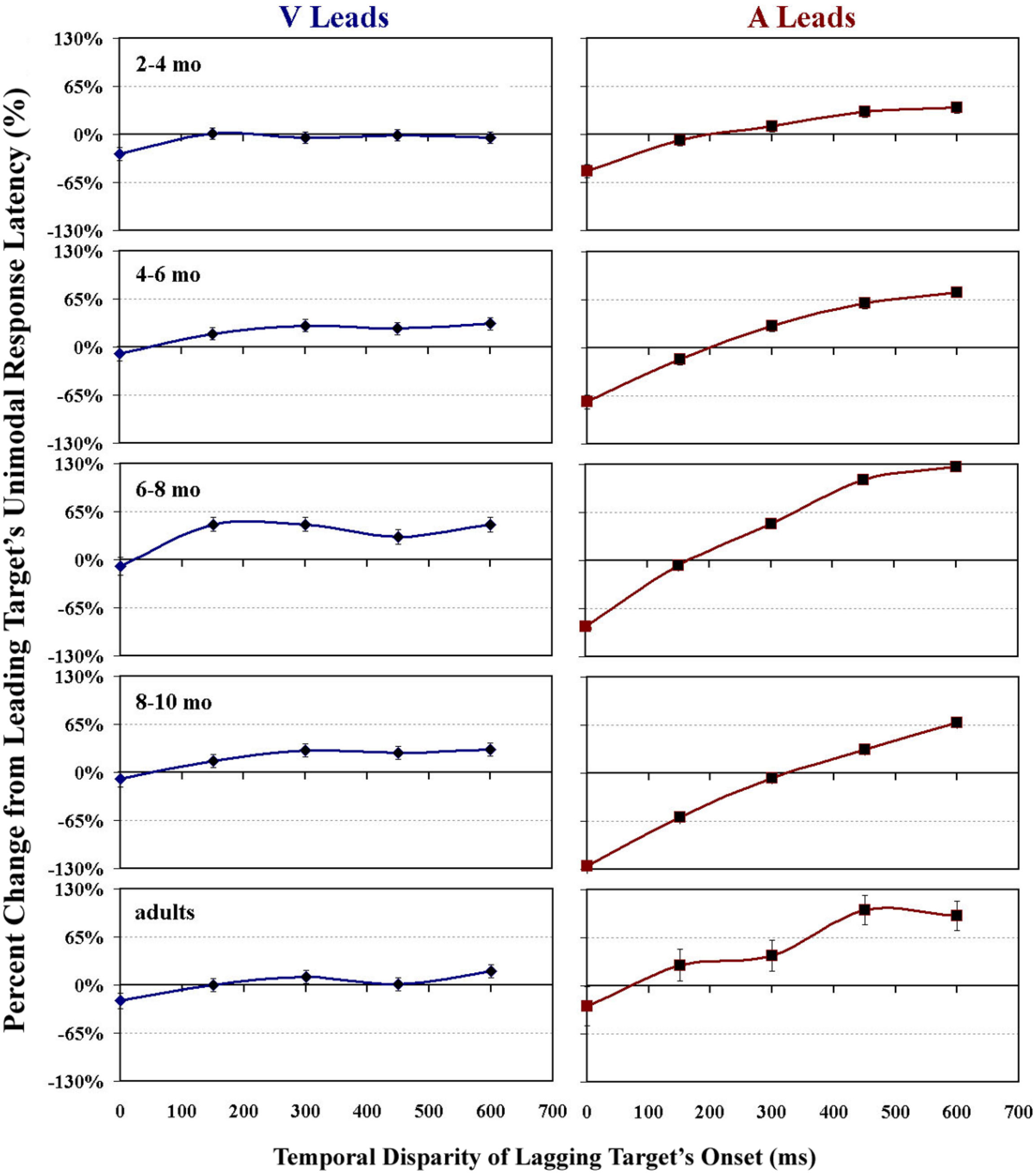


Figure 4.5: Percent change in temporally incongruent from unimodal (leading target, Study 1) mean response latencies, normalized by the temporally congruent mean latency at (a) and (b) 2–4 months, (c) and (d) 4–6 months, (e) and (f) 6–8 months, (g) and (h) 8–10 months, and (i) and (j) adults. Filled data symbols indicate the response latency was significantly different ($p < 0.001$) than the temporally congruent audiovisual condition. Error bars represent the combined standard errors.

be taking place: Infants are either initiating their orienting response to the onset of the leading auditory stimulus but are very slow to respond before the second onset overtakes, or they are choosing to wait until the second onset in anticipation of the more salient audiovisual event. Since all temporally disparate conditions ($\geq$ 150 ms) were tested in one experiment (Experiment 4) and all temporally congruent in another (Experiment 2), even with the relatively short duration of the experiment, the babies could have learned to expect the onset of an audiovisual target after an auditory-only. This latter theory is unlikely for two reasons. A 2-way univariate ANOVA (4 x 2) was performed, with disparity (150, 300, 450, 600 ms) and the trial presentation number (first half versus last half) for all A_{lead} conditions and age groups. If subjects were learning to expect a follow-up onset to an auditory stimulus and were delaying their response for it's onset, one would expect that trials in the last half of the experiments to be longer than the first half; they were not. We found no main effect or interaction for the trial presentation order at any age group ($p \geq 0.15$). What seems more likely explanation for the different profile between A_{lead} and V_{lead} trials can be found in the infants' relative performance to unimodal auditory and visual targets at 25° (Study 1). Where the response latencies for older infants (4–10 months) were in the range of 300–380 ms for a visual-only target—well within the range of half the temporal disparities in this study—the response latencies for auditory-only targets at the same eccentricity were much slower (Figure 2.2). An examination of the cumulative distribution functions for both unimodal conditions at 25° shows the low rise of the probability distribution of response latencies to an auditory-only target (Figure 4.6). In the youngest infants (2–4 months), by the time of the longest delay used in this study (600 ms), only a bit more than 50% of the reaction times had occurred. In the older infants, 300–450 ms had to have passed before there was even a 50% chance of a response having taken place to an auditory-only target. This weighs against the idea that the longer response latencies for the A_{lead} condition are due to infants waiting for the more salient, second (audiovisual) onset, but rather a limitation due to the immature development of their auditory spatial localization mechanism. Based upon these results, one could predict that if temporal disparity conditions were performed for this same range of ages but with targets more peripherally located—where unimodal latencies had been found to be closer in value—the degree of asymmetry between A_{lead} and V_{lead} conditions would be reduced. As in the V_{lead} condition, the temporal disparity for the onset of the second stimulus also adversely effected the mean response latency, and all disparities were as slow as or slower than would be expected if the babies were responding to a unimodal auditory target alone (Figure 4.5).

4.3.6 Temporal Disparity: Facilitation versus Inhibition

As in the previous study, we examined the facilitory and/or inhibitory interaction that temporal disparity might play on the mean response latency for a spatially and temporally congruent audiovisual target. The percent difference from the unimodal response latency of the leading target was calculated for all temporal disparities—normalized by the temporally congruent response latency—for each age group and plotted (Figure 4.5). In all ages and modality conditions, the result of temporal disparity was a significantly slower response latency for all disparity/modality conditions, even adults. As discussed above, the magnitude of response latency increase was greater when an auditory stimulus was presented first compared to a visual stimulus, most likely due to a greater immaturity in auditory spatial localization capabilities, though even adults—normal individuals aged between 21 and 30 years, who could be expected to be fully mature and free from degenerative affects due to aging—showed a greater change for A_{lead} than for V_{lead} conditions.

4.3.6.1 2–4 months

In the youngest infants, the V_{lead} mean response latencies were indistinguishable from a unimodal visual response latency (Figure 4.5a) and insensitive to large degrees of disparity, showing a generalized, slower response irregardless of disparity. Under A_{lead} conditions there was slight increase in latency with increasing temporal disparity (Figure 4.5b). Disparities of 150–300 ms slowed the response approximately 40% and

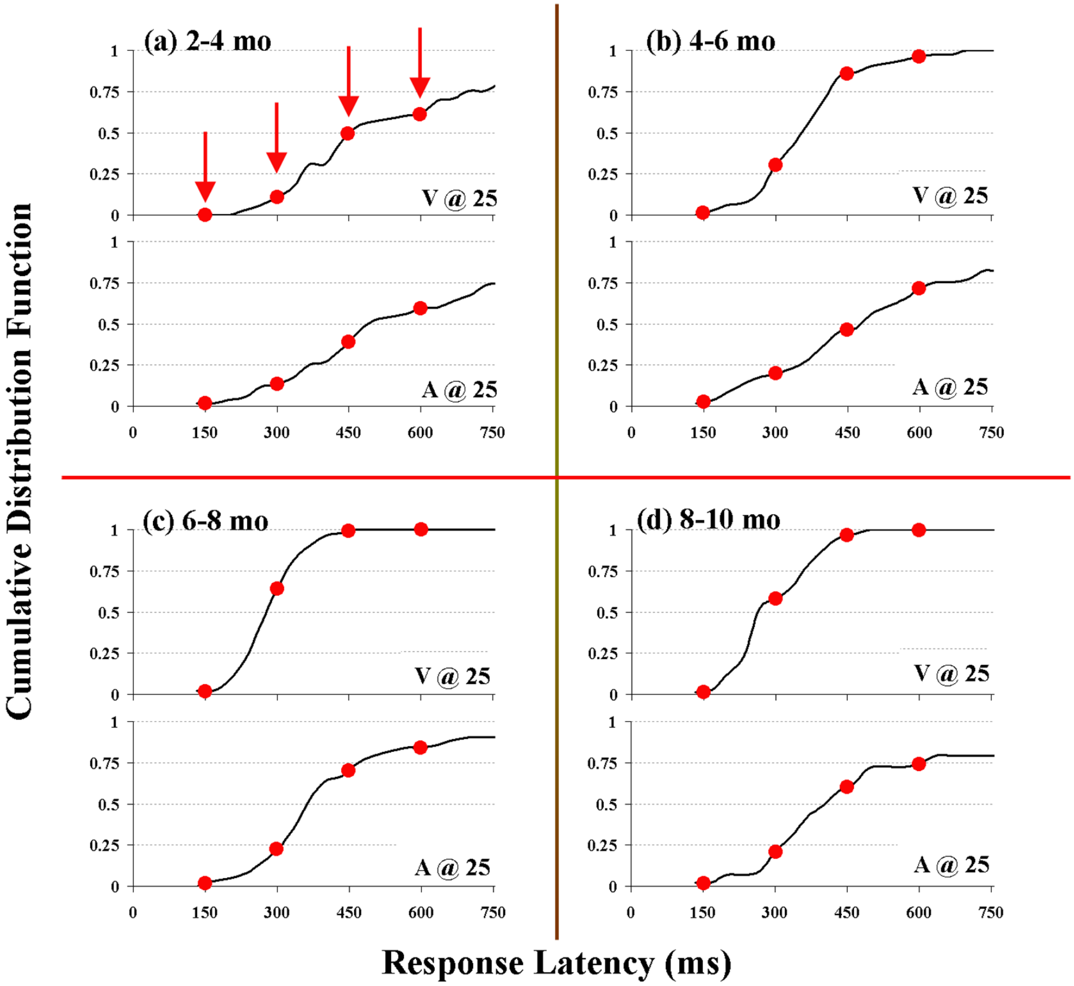


Figure 4.6: The cumulative distribution functions for visual-only and visual-only targets at 25° for (a) 2–4, (b) 4–6, (c) 6–8, and (d) 8–10 month olds. Dots correspond to probability of an onset occuring in response to the leading stimulus onset at the moment of the secondary stimulus onset for each of the four non-zero disparity conditions (150, 300, 450, and 600 ms).

disparities of 450–600 ms slowed the latencies an additional 40%.

4.3.6.2 4–6 months

In the 4–6 month olds, the relative percent change from congruent audiovisual (0 ms) latencies increased compared to younger infants, but with the same insensitivity to V_{lead} disparities and the linear increasing trend for A_{lead} disparities (Figure 4.5c, d). The degree to which the response was slowed above that for the corresponding unimodal latency suggests that in this age group, temporal disparity is providing a greater interference than before.

4.3.6.3 6–8 months

Continuing the pattern, 6–8 months showed response latencies that were nearly half again as long (40–50%) for all V_{lead} temporal disparities, with no significant change between disparity conditions (Figure 4.5e). For A_{lead} conditions, the congruent audiovisual response already being near 80% faster than mean auditory-only latency, the introduction of a temporal disparity dramatically increased response latencies (Figure 4.5f), over 200% for the largest disparity (600 ms).

4.3.6.4 8–10 months

In the oldest infant group, the pattern of ever increasing interference with age was interrupted, though the disparate response latencies for V_{lead} condition were still slower than even the mean unimodal response (Figure 4.5g). Unlike the completely flat response curve in the three youngest age groups for V_{lead} conditions, 8–10 month olds have the beginnings of a profile closer to that seen in adults, with the 150 ms disparity condition slightly faster than greater disparities, though not yet significantly less. Likewise, in the A_{lead} condition, the relative magnitude of interference—though still significant from audiovisual congruency for all disparities—is less than the previous age group, with the largest disparity causing a mean latency closer to 150% slower (Figure 4.5h). The percent difference for the 150 ms disparity condition was also the lowest that had been seen previously in any of the other infant groups.

4.3.6.5 Adults

Overall, infants tended to show the same general shape for both the V_{lead} and the A_{lead} percentage difference curves as adults (Figure 4.5i, j), with a few developmental differences. As in infants, the response latencies to a temporally disparate target were significantly longer then the synchronous condition, for all disparities tested. But in adults, the magnitude of difference was much closer to a visual-only response latency than in any infant group over 4 months.

4.4 Discussion

In all infant and adult age groups, temporal disparities of 150–600 ms between the onset of both stimuli produced response latencies that were significantly slower than a temporally congruent audiovisual stimulus, and many which were also slower than the mean latencies found for unimodal stimuli (Study 1). There were also additional findings of asymmetrical effects due to the modality of the leading target—closely related to the relative performances of the infants' auditory and visual localization—as well as age-related differences in the degree of slowed response latency.

When the visual component was the leading stimulus, both infants and adults showed a broad, slowed response that was insensitive to the magnitude of the temporal disparity used. In the youngest infants, the mean response latencies were all equivalent to the mean latency found for visual-only targets. The older infants shared the same broad, magnitude-insensitive profile for the V_{lead} conditions, but the magnitude of the effect (the slowing of the response compared to a congruent target) increased with age, peaking at 6–8 months and dropping off, though still present, in the oldest infants. When the auditory component was the leading stimulus, infants showed a linear trend for increased slowing of the response latency, with the youngest infants showing the smallest change with increasing disparity and the maximal effect occuring, again, at 6–8 months. In the oldest infants, the general shape of the plots for both V_{lead} and A_{lead} disparities were similar to those found in adults, though the magnitude of the effects were still different. The inverted u-shaped profile of an increasing, then decreasing, magnitude of difference from the congruent audio-visual response, combined with the minimal slowing of response latencies for either modality-leading condition in the youngest infants suggests a general developmental timeline. Starting with broad level of interference in the orienting response to a wide range of temporal disparities, as infants mature the degree to which temporal asynchrony can interfere with the orienting response first increases, peaking at 6–8 months, and then decreases and becomes more finely tuned by 8–10 months.

The asymmetrical influence of the leading target's modality on response latency appears to be closely related to infants' performance in auditory versus visual spatial localization. A visual stimulus coming on before the auditory stimulus produced faster latencies that were invariant with increasing temporal displacement. However, when the auditory stimulus came on first, the response latencies were much longer, and varied directly with increasing disparity. These results are consistent with the dominance of the visual system in the parafoveal region, as well as the difficulty infants of this age were found to have in localizing auditory-only targets at 25° (Study 1). In both cases, infants seemed to be initiating their response after the onset of the leading target but were limited more by the less capable auditory localization system than the visual system. In the previous study, all infants (2–10 months) were much faster at orienting toward a visual target than an auditory target at 25°. Looking at the cumulative distribution curves for the two modalities, by 300 or 450 ms post-onset a response was much more likely to have been initiated to a visual target that to an auditory one, especially in the older age groups (60–75% versus only 25–60% for 6–10 month olds). It was therefore much more likely for a response to have already been begun to a visual-leading target by the time the auditory target was presented than the reverse; this was born out by measurements of the percentage of reaction times taking place prior to second target onset 4.4.