

Is a Picture Worth a Thousand Words? The Flexible Nature of Modality Dominance in Young Children

Amanda C. Napolitano and Vladimir M. Sloutsky

When presented simultaneously with equally discriminable, but unfamiliar, visual and auditory stimuli, 4-year-olds exhibited auditory dominance, processing only auditory information (Sloutsky & Napolitano, 2003). The current study examined factors underlying auditory dominance. In 6 experiments, 4-year-olds ($N = 181$) were presented with auditory and visual compounds in which (a) the complexity and familiarity of stimuli were systematically varied (Experiments 1–5) and (b) participants were explicitly instructed to attend to a particular modality (Experiment 6). Results indicate that auditory dominance is a special case of flexible modality dominance, which may stem from automatic pulls on attention. Theoretical implications of these results for understanding the development of attention and cross-modal processing, as well as linguistic and conceptual development, are discussed.

It is well known that words play an important role in young children's conceptual organization and thinking. For example, when two entities share an auditorily presented label (e.g., both are called *a bird*), young children are more likely to consider the entities to be similar (Sloutsky & Fisher, 2004; Sloutsky & Lo, 1999), to group the entities together (Balaban & Waxman, 1997; Markman & Hutchinson, 1984; Sloutsky & Fisher, 2004), and to induce unobservable properties from one entity to another (Gelman & Markman, 1986; Sloutsky & Fisher, 2004; Sloutsky, Lo, & Fisher, 2001). It has been argued recently (Sloutsky & Napolitano, 2003) that some of these effects may stem from general auditory factors, such as privileged processing of auditory information by young children: When auditory and visual stimuli are presented simultaneously, young children are more likely to attend to auditory stimuli, which is not the case for adults. This privileged processing of auditory information manifests itself in auditory preference and auditory dominance effects.

Auditory Preference and Auditory Dominance Effects

Auditory preference is a tendency to treat the auditory component of an auditory–visual compound stimulus as a more important cue than the visual component. Auditory preference was established in a series of experiments (Sloutsky & Napolitano, 2003) that used a modified version of the switch task (see Werker, Cohen, Lloyd, Casasola, & Stager, 1998, for a description of the original task). In this task, 4-year-olds and adults were trained to select a particular combination of an auditorily presented tone pattern and visual stimulus (AUD_1-VIS_1) over another combination of an auditorily presented tone pattern and visual stimulus (AUD_2-VIS_2). When training was completed, the trained set was broken into two new sets, such that the trained auditory component was paired with a new visual component (AUD_1-VIS_1) and the trained visual component was paired with a new auditory component ($AUD_{new}VIS_1$). Participants were asked which of the two new sets was the “trained” one. It was found that 4-year-olds overwhelmingly selected AUD_1VIS_{new} thereby exhibiting auditory preference, whereas the vast majority of adults selected $AUD_{new}VIS_1$, thereby exhibiting visual preference.

Although 4-year-olds exhibited auditory preference, it remained unclear whether this preference reflected deliberate choice or whether it stemmed from auditory dominance (i.e., the failure to encode visual stimuli). On the one hand, it is possible that young children encoded both the auditory and the visual components and then strategically rejected the latter in favor of the former. On the other hand, it is

Amanda C. Napolitano and Vladimir M. Sloutsky, Center for Cognitive Science, The Ohio State University.

This research was supported by grants from the National Science Foundation (BCS 0078945 and REC 0208103) to Vladimir M. Sloutsky. We would like to thank Anna Fisher, George Hollich, Chris Robinson, and Aaron Yarlas, and three anonymous reviewers for their helpful comments.

Correspondence concerning this article should be addressed to Vladimir M. Sloutsky, Center for Cognitive Science, 208C Ohio Stadium East, 1961 Tuttle Park Place, Ohio State University, Columbus, OH 43210. Electronic mail may be sent to Sloutsky.1@osu.edu.

possible that auditory modality dominated visual modality, with auditory stimuli overshadowing (or preventing processing of) visual stimuli.

To distinguish between these possibilities, Sloutsky and Napolitano (2003) conducted an immediate old–new recognition experiment, in which they presented 4-year-olds and adults with a target stimulus, AUD_1VIS_1 , which was identical to the trained set in the modified switch task. After seeing the target, participants were presented with one of the following test items: (a) AUD_1VIS_1 , which was the old target item; (b) AUD_1VIS_{new} , which had the trained auditory component and a new visual component; (c) $AUD_{new}VIS_1$, which had the trained visual component and a new auditory component; or (d) $AUD_{new}VIS_{new}$, which had a new visual and a new auditory component. The task was to determine whether each presented test item was the same as the target (i.e., both the same auditory and visual components) or a new item (i.e., differed on one or both components). If participants encode both auditory and visual stimuli, they should correctly respond to all items by accepting old target items and rejecting all other test items. Alternatively, if they fail to encode the visual component, they should falsely accept AUD_1VIS_{new} items while correctly responding to other items. Finally, if they fail to encode the auditory component, they should falsely accept $AUD_{new}VIS_1$ items while correctly responding to other items.

Findings supported the overshadowing possibility, with the auditory modality dominating the visual modality: Four-year-olds failed to encode visual stimuli when these were accompanied by auditory stimuli, erroneously accepting AUD_1VIS_{new} items. At the same time, these children had no difficulty encoding the visual stimuli in a control condition when these stimuli were presented without corresponding auditory stimuli, which indicated that the processing of visual stimuli was not difficult per se. Note that adults ably encoded both visual and auditory stimuli, although they exhibited marked visual preference in the modified switch task.

These findings appear to contradict a well-established line of research demonstrating that when cross-modal stimuli include the same amodal relation expressed both visually and auditorily (e.g., temporal synchrony, rhythm, or tempo) even young infants efficiently process information in both modalities (see Lewkowicz, 2000a; Lickliter & Bahrick, 2000, for extensive reviews). We believe, however, that these two lines of research provide complementary evidence about cross-modal processing: Although some cross-modal stimuli in the real world

have a shared amodal relation, other important cross-modal stimuli represent arbitrary pairings (e.g., words and objects they denote or sounds produced by many artifacts and living things). Furthermore, we believe that the contradiction is apparent rather than real because amodal relations differ substantially from arbitrary pairings (and the two lines of research have used substantially different methodologies), and we discuss these differences in a greater detail in the General Discussion section.

At a more general level, the auditory preference and auditory dominance (or overshadowing) effects reported by Sloutsky and Napolitano (2003) have several important theoretical implications extending our understanding of learning and attention early in development, as well as some aspects of lexical and conceptual development. First, the fact that young children are strongly biased to process auditory information over visual information may affect our understanding of their learning of cross-modally presented information.

Second, these effects may affect our understanding of resource allocation in cross-modal processing: Young children exhibited little evidence of resource sharing between modalities (which would have resulted in an attenuated processing in both modalities); rather, one modality received full processing and the other modality received little or no processing.

Third, because words and the entities they denote represent arbitrary pairings, these effects may have important implications for our understanding of lexical development and language acquisition. Note that auditory events are transient, whereas visual objects and scenes are usually stable and their presence is protracted. Therefore, auditory dominance may play an important role in the ability of infants and young children to attend to words: They might be unable to attend to these transient auditory stimuli in the absence of auditory dominance. Finally, these effects may extend our understanding of conceptual development: They suggest that the effects of words in a variety of categorization and induction tasks may stem in part from the privileged processing of auditory information.

Although these theoretical implications are potentially significant, and the findings they stem from are novel, Sloutsky and Napolitano's (2003) research left several important questions unanswered, thus making the scope of generalization of these findings unclear. First, it is possible that auditory dominance effects are specific to the class of stimuli used by Sloutsky and Napolitano. Second, it is possible that

these findings point to an absolute auditory dominance indicating that under all (or at least most) conditions auditory stimuli overshadow visual stimuli. And third, it is possible that auditory dominance is a special case of flexible modality dominance, such that under some conditions auditory stimuli overshadow visual stimuli, whereas under other conditions visual stimuli overshadow auditory stimuli. Each of these possibilities has different theoretical implications, and the goal of this research is to distinguish among these possibilities.

Hypothesis 1: Auditory Dominance Is Stimulus Specific

Auditory dominance could be specific to the class of stimuli used by Sloutsky and Napolitano (2003). If this is the case, then in the absence of some critical properties of these stimuli, auditory dominance should disappear such that under different stimulus conditions participants would either (a) exhibit equivalent levels of processing in both modalities, or (b) exhibit visual dominance. In this section we consider the former possibility, and in the section describing Hypothesis 3 we consider the latter possibility.

If modality dominance is stimulus specific, and it disappears under different stimulus conditions, then Sloutsky and Napolitano's (2003) findings reflect properties of stimuli that either facilitate or prevent processing, and it is necessary to examine these properties. Sloutsky and Napolitano used specific visual stimuli that might be less than optimal for attending (see Figure 1 for examples of their visual stimuli). Several factors could account for the absence of encoding of these stimuli.

First, these stimuli depicted scenes and arrangements of objects rather than individual objects, and it is possible that individual objects could be more

likely to engage attention than scenes or arrangements (e.g., Scholl, Pylyshyn, & Feldman, 2001), particularly for young children. Second, both scenes and arrangements were unfamiliar, and familiar entities could be more likely than novel stimuli to engage attention (e.g., Christie & Klein, 1995). Finally, it is possible that complex stimuli, such as photographs of real entities or displays of multiple objects, are less likely to be processed than simpler stimuli because it might be more difficult or take longer to process more complex stimuli, even though children ably process these same stimuli in the absence of auditory input.

Therefore, under each of these accounts, it is possible that more object-like, more familiar, or simpler visual stimuli than those used by Sloutsky and Napolitano (2003) would eliminate auditory dominance in 4-year-olds, and these participants would either fully or partially encode information in both modalities.

Hypothesis 2: Auditory Dominance Is a General Characteristic of Bimodal Processing

It is also possible that the auditory dominance found by Sloutsky and Napolitano (2003) is a general characteristic of bimodal processing early in development. In this case, auditory stimuli would overshadow visual stimuli under most conditions.

If auditory dominance is a general characteristic of young children's processing (we know of no evidence that auditory dominance characterizes processing in adults), then Sloutsky and Napolitano's (2003) findings may reflect maturational and experiential asynchronies: The auditory system starts functioning during the last trimester of gestation, allowing the fetus auditory experience in utero (Birnholtz & Benaceraff, 1983; see also Jusczyk, 1998,

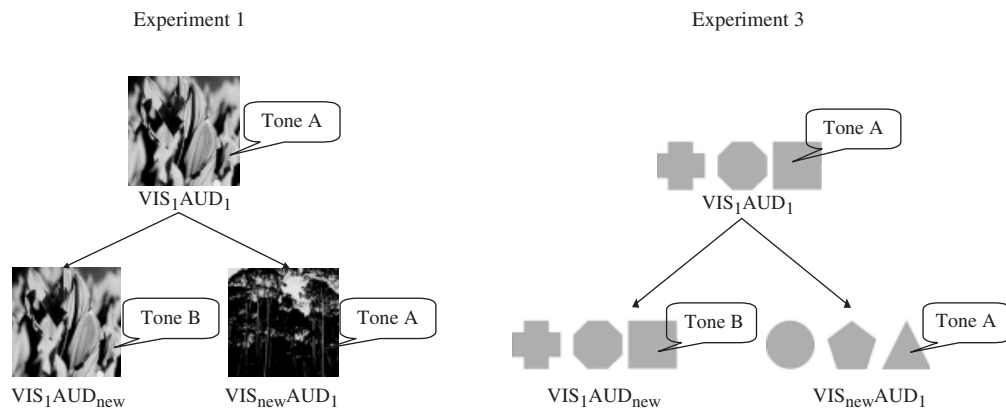


Figure 1. Example of stimulus sets used in Sloutsky and Napolitano (2003, Experiments 1 and 3). VIS = visual stimulus; AUD = auditory stimulus.

for a review), whereas the visual system does not start functioning until after birth. Therefore, early maturation may give the auditory modality an experiential head start over the visual modality. If auditory dominance stems from maturational and experiential factors, then a reversal of auditory dominance early in development should be difficult, if not impossible, as long as stimuli have comparable discriminability and salience.

Hypothesis 3: Auditory Dominance Is a Special Case of Flexible Modality Dominance

Finally, it is possible that auditory dominance is a special case of modality dominance such that (as long as visual and auditory stimuli have comparable discriminability) the auditory modality dominates the visual modality in some auditory–visual combinations, whereas the visual modality dominates the auditory modality in different combinations of the same stimuli. If such flexible modality dominance is found, then the Sloutsky and Napolitano (2003) findings reflect interesting properties of attention pertaining to (a) resource allocation in processing of multimodal stimuli and (b) flexibility of young children’s automatic attention.

First, flexible modality dominance would further extend the Sloutsky and Napolitano (2003) findings, indicating that processing of bimodal stimuli often does not result in resource sharing across modalities: One modality receives full processing rather than both modalities receiving attenuated processing.

Second, flexible modality dominance would indicate that young children’s attention shifts flexibly between the auditory and visual modalities. Although attentional shifts between auditory and visual stimuli have never been examined directly, this possibility is supported by previous research indicating that young children do shift attention flexibly between different properties of visual stimuli (Jones & Smith, 2002; Jones, Smith, & Landau, 1991; Smith, Jones, & Landau, 1996).

Third, if shifts in modality dominance are found, these shifts are more likely to stem from automatic rather than deliberate attentional flexibility. Note that flexible (or selective) attention is often defined in two distinct ways: (a) as person controlled, or deliberate, and (b) as stimulus controlled, or involuntary (see Egeth & Yantis, 1997; Pashler, Johnston, & Ruthruff, 2000, for reviews). In person-controlled flexible attention, attention is said to be flexible when a person selectively attends to a particular set of stimuli or stimulus dimensions and shifts attention to another set or dimension when instructed to do so.

For example, classical shadowing experiments with different information presented to different ears (see Pashler, 1998, for a review) found that people can exhibit such flexible selectivity and attend to a predetermined stimulus set. Alternatively, in stimulus-controlled flexible attention, attention is said to be flexible when under various stimulus conditions attention is automatically pulled to different properties or dimensions (e.g., Lamberts, 1994; Nosofsky, 1986; Smith et al., 1996; Tversky, 1977). For example, Smith et al. (1996) presented 3-year-olds with novel objects, with each object having a distinct base shape and distinct parts. They also presented participants with a nonmember of the category that by means of contrast made either the base or the parts diagnostic, and thus salient. Depending on the contrast condition, young children attended either to the shape (when the shape was more diagnostic and thus salient) or to the parts (when the parts were more diagnostic and thus salient).

There are reasons to believe that shifts in modality dominance (if found) are likely to stem from automatic pulls on attention rather than from deliberate selective attending to a particular modality. In particular, young children often fail to shift attention from one visual cue to another when instructed to do so (e.g., Kirkham, Cruess, & Diamond, 2003; Zelazo, Frye, & Rapus, 1996), whereas stimulus-driven shifts within the visual modality can be easily induced in young children (Smith et al., 1996).

If modality dominance stems from pulls on attention—automatic attending to certain properties of stimuli—then various kinds of stimuli that automatically engage attention may also generate modality dominance. For example, as mentioned earlier, familiar stimuli are likely to automatically engage attention (Christie & Klein, 1995), with familiarity being established early in the course of processing and familiar stimuli eliciting a different neuronal response in the primate brain than novel stimuli (Hölscher, Rolls, & Xiang, 2003; Xiang & Brown, 1998). Automatic attention to familiar stimuli may represent an adaptive attentional mechanism because familiarity of stimuli may reflect prominence of these stimuli in one’s environment, and it might be more efficient to notice recurring stimuli before noticing occasional stimuli.

In short, if auditory dominance is a special case of flexible modality dominance stemming from attentional factors, then when auditory stimuli are paired with visual stimuli that automatically engage attention (e.g., highly familiar visual stimuli), auditory dominance could be reversed. Furthermore, if modality dominance stems from automatic attending to

certain properties of stimuli, then explicit instructions to attend to a rejected rather than a dominant modality should not eliminate modality dominance effects.

Overview of Experiments

The goal of the reported experiments was to distinguish among these three hypotheses. To determine whether the auditory dominance reported by Sloutsky and Napolitano (2003) disappears, shifts, or sustains under different stimulus conditions, in Experiments 1 and 2, we systematically manipulated familiarity and complexity of visual stimuli while controlling for their "objecthood." In Experiments 3 to 5, we manipulated familiarity of both auditory and visual stimuli. Finally, in Experiment 6, to examine whether modality shifts stem from deliberate selective attention or from automatic pulls, we explicitly instructed participants to attend to an overshadowed modality.

Experiment 1

The goal of Experiment 1 was to examine whether complexity and familiarity could contribute to the salience of visual stimuli and thereby decrease the level of auditory dominance reported in Sloutsky and Napolitano (2003). Because Sloutsky and Napolitano used complex and unfamiliar visual stimuli, and established auditory dominance with these stimuli, Experiment 1 examined the remaining three cells of the 2 (simple vs. complex) \times 2 (familiar vs. novel) contingency table presented in Figure 2. These new conditions were: (a) simple and familiar visual stimuli (Simp+Fam), (b) simple and novel visual stimuli (Simp+Nov), and (c) complex and familiar visual stimuli (Comp+Fam). Familiarity of visual stimuli was first established in the calibration experiments described later. Complexity is defined as the amount of perceptual detail, such as the number of identifiable objects, the number of distinct parts, and the number of brightness contrasts per unit of space. Therefore, multiobject stimuli are more complex than single-object stimuli, and perceptually rich photographs of real objects are more complex than outlines of geometric shapes.

Experiment 1 used the same modified switch task as the one used by Sloutsky and Napolitano (2003, Experiments 1 and 3). Participants were first trained to select a target stimulus set (VIS₁AUD₁) composed of simultaneously presented auditory and visual components. If training was successful (i.e., children were able to reach a criterion in selecting the target),

		Overall Complexity	
		Simple	Complex
Overall Familiarity	Familiar	Condition 1: Single Geometric Figures	Condition 2: Animal Photographs
	Novel	Condition 3: Single Abstract Shapes	Sloutsky and Napolitano (2003): Landscapes and Three Shape Arrangements

Figure 2. A 2 \times 2 design crossing familiarity and complexity of visual stimuli.

children moved to a test phase. At the test phase, the target set was broken apart so that the trained image was paired with a new sound (VIS₁AUD_{new}), the trained sound was paired with a new image (VIS_{new}AUD₁), and participants were asked to select the one that was the target set. The selection of the VIS_{new}AUD₁ item would be indicative of an auditory preference, whereas the selection of the VIS₁AUD_{new} would be indicative of a visual preference. At the same time, if participants do not have a modality preference, they should respond at chance.

Method

Participants

Participants were 45 young children ($M = 4.41$ years, $SD = 0.35$ years, range = 3.8 to 5.0 years; 19 girls and 26 boys). In this and all other experiments reported here, participants were recruited from child care centers located in middle-class suburbs of the Columbus, Ohio area, and the majority of the participants were White. There were three between-participants conditions, described later, with 15 children participating in each condition. An additional group of 11 children also participated in Experiment 1, but they were either absent from their day care and did not complete all four blocks (8 children) or did not pass a single training session (3 children). These children were not included in any analyses.

Materials

Materials consisted of visual–auditory stimulus sets. Within each set, visual and auditory components were presented simultaneously, such that each image’s presentation matched the duration of the corresponding sound. The sets were created by randomly pairing an auditory and a visual component.

For each condition a total of 16 stimulus sets were used. Within each condition, there were four types of stimulus sets: (a) the training target set, VIS_1AUD_1 , which participants were trained to select; (b) the distracter set, VIS_2AUD_2 , which participants were trained to reject; (c) VIS_1AUD_{new} which was presented at test and matched the training target’s visual component but had a novel auditory component; and (d) $VIS_{new}AUD_1$, which was presented at test and had a novel visual component but matched training target’s auditory component (see Figure 3 for examples of VIS_1AUD_1 , VIS_1AUD_{new} and $VIS_{new}AUD_1$).

The auditory stimuli were identical to those used by Sloutsky and Napolitano (2003). These were computer-generated tone patterns, each consisting of three unique simple tones. Simple tones varied on timbre (sine, triangle, or sawtooth) and frequency (between 1 and 100 Hz). Each simple tone was 0.3 s in duration and was separated by .05 s of silence, with total pattern duration of 1 s. The average sound level of auditory stimuli was 67.8 dB (with a range from 66 to 72 dB), which is comparable to the sound level of human voice in a regular conversation.

Familiarity of these sounds was established in a separate calibration experiment. The calibration experiment consisted of asking a sample of twelve 4-year-olds (none of whom participated in any of the other experiments reported here) two questions about each auditory stimulus: “Have ever heard this sound before?” and “What is the sound?” For each participant, the average percentage of *yes* responses to the first question and consistent responses to the second question was calculated, and the individual participant percentages were averaged across participants to calculate an overall mean percentage. Few of the items were reported to be heard before ($M = 32.00\%$), and even fewer items were labeled consistently (i.e., different participants rarely gave the same or synonymous label to the same item; $M = 0.27\%$), although on a small percentage of trials they did attempt to identify the tones with a label ($M = 18\%$). It was therefore concluded that the tone patterns were unfamiliar.

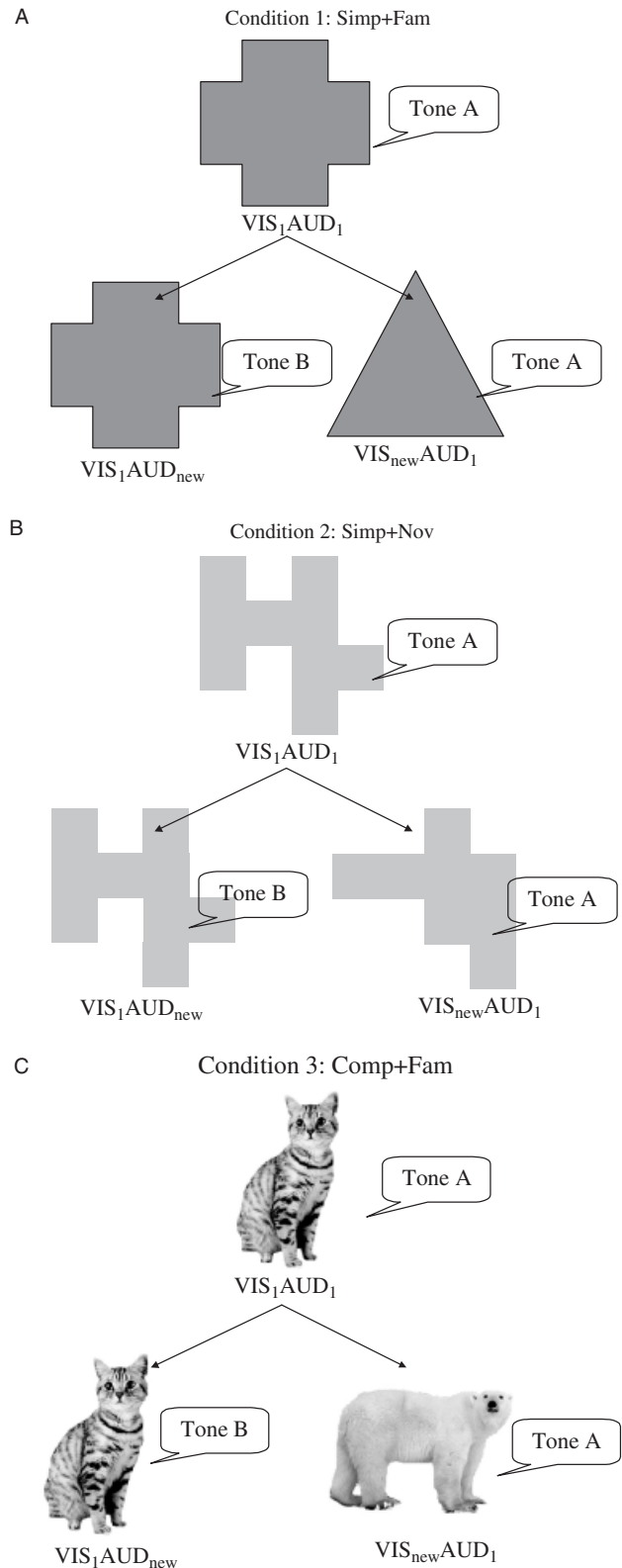


Figure 3. Example of stimulus sets for each condition of Experiment 1. VIS = visual component; AUD = auditory component; Simp = simple visual stimuli; Comp = complex visual stimuli; Nov = novel visual stimuli; Fam = familiar visual stimuli.

As mentioned earlier, visual stimuli varied across conditions in their complexity and familiarity. The complexity was established using the criteria defined earlier (i.e., the number of identifiable objects and the number of distinct parts, and the number of brightness contrasts per unit of space), whereas familiarity was established in a separate calibration experiment. The calibration experiment consisted of asking a sample of thirteen 4-year-olds (none of whom participated in other experiments reported here) two questions about each visual stimulus: "Have you ever seen one of these?" and "What is it?" For each participant, the average percentage of *yes* responses to the first question and correct responses to the second question was calculated, and the individual participant percentages were averaged across participants to calculate an overall mean percentage. For the Simp+Fam condition, most of the items were reported to be seen before ($M = 84.6\%$), and most of the items were correctly and consistently labeled ($M = 80.8\%$). For the Simp+Nov condition, fewer items were reported to be seen before ($M = 30.8\%$), and even fewer items were consistently labeled ($M = 5.7\%$), although participants attempted to label the stimuli on more than one fourth of the trials ($M = 26.9\%$). For the Comp+Fam condition, most of the items were reported to be seen before ($M = 96.2\%$), and most of the items were correctly and consistently labeled ($M = 96\%$). Based on these responses, it was concluded that the visual stimuli in Conditions 1 and 3 were indeed familiar, whereas the visual stimuli of Condition 2 were novel. Examples of visual stimuli for each of the three conditions are presented in Figure 3.

Condition 1: Simp+Fam. The simple and familiar visual stimuli for Condition 1 were computer-generated single two-dimensional geometric figures. Each shape was 10 cm \times 10 cm and was colored green.

Condition 2: Simp+Nov. The simple and novel visual stimuli for Condition 2 were computer-generated two-dimensional figures that were created by randomly coloring in 2.5 cm \times 2.5 cm squares of a 4 \times 4 grid such that: (a) each column had at least one colored square, (b) all colored squares were connected, and (c) there were no uncolored gaps within the overall colored space. Gridlines were removed to create a continuous shape. Each shape was 10 cm \times 10 cm and was colored green.

Condition 3: Comp+Fam. The complex and familiar visual stimuli for Condition 3 were photographs of familiar animals (e.g., cats, dogs). Each photograph was 10 cm \times 10 cm and varied in color.

Another calibration experiment was conducted to ascertain the comparable discriminability among the

visual and auditory stimuli. Discriminability was established using a same-different immediate recognition task in which a different sample of fifteen 4-year-olds made same-different judgments after being presented with pairs of visual or auditory stimuli. Within each trial, visual-auditory stimuli compounds were presented successively for 1 s each. Participants exhibited high levels of discrimination of visual stimuli in the Simp+Fam condition ($M = 91\%$), in the Simp+Nov condition ($M = 86\%$), and in the Comp+Fam condition ($M = 96\%$). Comparable discriminability of the auditory stimuli used here was established previously (Sloutsky & Napolitano, 2003).

Design and Procedure

The experiment had a between-subjects design with the visual stimulus condition as a factor (i.e., Simp+Fam vs. Simp+Nov vs. Comp+Fam). A female experimenter tested the participants individually in a quiet room within their day care centers. She told participants that they would play a game in which they should find the location of a prize and that they would be rewarded at the end of the game with a prize. At the end of their participation for each day, regardless of their responses, all children received a small toy as their reward.

Each participant had four blocks, with each block consisting of eight training trials (a training session) and six test trials (a testing session). Participants had two blocks per day, and the experiment was spread over a 2-week period. All stimuli were presented on a Dell Inspiron laptop computer, and presentation of stimuli and recording of responses was controlled by a Visual Basic program.

Training session. First, either VIS₁AUD₁ or VIS₂AUD₂ appeared on one side of the screen, followed by the remaining stimulus set (i.e., either VIS₂AUD₂ or VIS₁AUD₁) appearing on the other side of the screen. The order of appearance and the side of the screen on which set each appeared was counterbalanced across training trials for the two stimulus sets, such that each set could appear either first or second and on either the right or left side of the screen. A white circle icon replaced each set at the end of its presentation, and the child was asked to identify the stimulus set that "had the prize behind it" by pointing to the icon that represented the selected set. The goal of training was to teach the child to select consistently the VIS₁AUD₁ stimulus set (i.e., the location of the prize); therefore, on each trial the child was provided with *yes* feedback when this stimulus set was chosen and *no* feedback when the VIS₂AUD₂

Table 1
Overall Structure of Trials in One Block, Experiment 1

Training session ($N = 8$ trials)		Testing session ($N = 6$ trials)	
VIS_1AUD_1 (Trained set)	VIS_2AUD_2 (Distracter set)	VIS_1AUD_{new} (Test Set A)	$VIS_{new}AUD_1$ (Test Set B)

Note. VIS = visual component; AUD = auditory component.

stimulus set was chosen. Only participants making correct selections for the final four trials moved into the test session.

Test session. The test session followed immediately after the training session, where participants were presented with two novel stimulus sets. Set VIS_1AUD_{new} matched the training target's visual component but had a novel auditory component, whereas set $VIS_{new}AUD_1$ had a novel visual component but matched the training target's auditory component. The participants were asked again to identify the set where a prize was hidden: Selections of $VIS_{new}AUD_1$ would indicate auditory preference, whereas selections of VIS_1AUD_{new} would indicate visual preference. Stimuli were presented as follows. First, either $VIS_{new}AUD_1$ or VIS_1AUD_{new} appeared on one side of the screen, followed by the remaining stimulus set (i.e., either $VIS_{new}AUD_1$ or VIS_1AUD_{new}) appearing on the other side of the screen. Again, the positions of the two stimulus sets were counterbalanced across the six test trials, and a white circle icon replaced each set at the end of its presentation. When the selection was made, the experimenter pressed the keyboard key corresponding to the selection without giving feedback to the participant. The overall structure of training and testing trials is presented in Table 1. Note that the structure of each block is identical to that used by Sloutsky and Napolitano (2003).

Results and Discussion

Proportions of trials indicating auditory preference by condition are presented in Figure 4. Also included in the figure are proportions of auditory preferences from the Sloutsky and Napolitano (2003) experiments that used complex and novel stimuli (Comp+Nov). At the most general level, data in the figure point to marked familiarity effects and to no appreciable complexity effects. Proportions of "old" responses for $VIS_{new}AUD_1$ (which indicate auditory preference) were subjected to a one-way analysis of variance (ANOVA) with condition as a factor. There was a significant main effect of condition, $F(2, 42) = 13.85, p < .01$. A post hoc Tukey test pointed to a

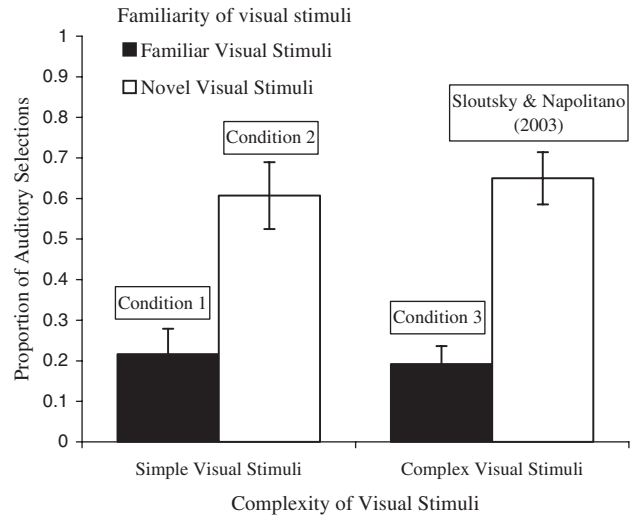


Figure 4. Proportions of auditory selections per condition of Experiment 1. Error bars represent standard errors of the mean.

significant difference between Condition 2 (i.e., Simp+Nov) and Conditions 1 and 3 (Simp+Fam and Comp+Fam), $ps < .01$, such that participants were more likely to exhibit auditory preference by selecting $VIS_{new}AUD_1$ in the former condition but not in the latter two conditions.

As a more conservative analysis of the participants' performance, we calculated the number of blocks with (a) above-chance reliance on auditory stimuli, (b) above-chance reliance on visual stimuli, and (c) chance performance. Performance was considered above chance if the same choice was made on five of six trials (binomial test, $p = .09$), otherwise it was considered at or below chance. Blocks with above-chance auditory responding were judged as exhibiting auditory preference, whereas blocks with above-chance visual responding were judged as exhibiting visual preference.

In the Simp+Fam condition, out of 60 blocks (i.e., 4 blocks \times 15 children), participants successfully completed 48 blocks, with 11% of successfully completed blocks exhibiting auditory preference, 66% exhibiting visual preference, and 23% being at chance. In the Simp+Nov condition, out of 60 blocks, participants successfully completed 44 blocks, with 36% exhibiting auditory preference, 30% exhibiting visual preference, and 34% being at chance. In the Comp+Fam condition, out of 60 blocks, participants successfully completed 55 blocks, with 2% exhibiting auditory preference, 65% exhibiting visual preference, and 33% being at chance. Proportions of blocks with above-chance performance per participant across the conditions were subjected to a one-way between-subject ANOVA with condition as a factor.

The analysis pointed to a main effect of condition, $F(2, 39) = 3.39, p < .05$, such that there were significantly more above-chance auditory blocks in the Simp+Nov condition than in the Simp+Fam or Comp+Fam conditions, post hoc Fisher's least significant different (LSD) test, both $ps = .054$.

To analyze individual patterns of responses across the conditions, participants were broken into three categories: (a) participants who relied consistently on auditory stimuli (i.e., made auditory choices on 75% or more of the trials) were categorized as auditory responders, (b) participants who relied consistently on visual stimuli (i.e., made auditory choices on 25% or less of the trials) were categorized as visual responders, and (c) participants who were inconsistent in their preferences (i.e., their auditory choices fell between 25% and 75% of the trials) were categorized as mixed responders. As shown in Table 2, there were more auditory-based responders in the Simp+Nov condition than in the other two conditions, $\chi^2(4, N = 45) = 17.09, p < .01$. The analysis of standardized residuals indicated that auditory responding was the most likely pattern in Condition 2 (Simp+Nov), whereas visual responding was the most likely pattern in Condition 1 (Simp+Fam) and Condition 3 (Comp+Fam). Furthermore, the overall pattern of results in Condition 2 (Simp+Nov) was statistically equivalent to that in Sloutsky and Napolitano (2003; Comp+Nov), $\chi^2(2, N = 28) < 1, p > .9$, whereas results of Condition 3 (Comp+Fam) differed significantly, $\chi^2(2, N = 28) = 9.40, p < .01$. These results further indicate that, unlike familiarity, complexity was not a significant factor in modality preference.

Overall, the results of the reported experiment indicate that modality preference is affected by the

familiarity of visual stimuli but not by the complexity. However, this experiment revealed only the preferences exhibited by young children, and we deemed it important to examine whether modality dominance also is affected by stimulus familiarity. To address this issue, Experiment 2 examined whether participants encode nonselected stimuli.

Experiment 2

The goal of Experiment 2 was to determine whether Experiment 1 revealed only modality preference effects, with participants strategically choosing the more familiar stimuli, or whether it revealed modality dominance effects, with participants encoding more familiar stimuli over less familiar stimuli. To achieve this goal, we presented participants with an old–new immediate recognition task under the same stimuli conditions as those used in Experiment 1.

Method

Participants

Participants were 45 young children ($M = 4.13$ years, $SD = 0.85$ years, range = 3.6 to 4.9 years; 25 girls and 20 boys). There were 3 between-participants conditions (similar to those in Experiment 1), with 15 children participating in each condition. An additional group of 5 children did not exhibit above-chance accuracy on control items described later, and these children were not included in any analyses.

Materials

The auditory and visual stimuli and conditions (i.e., Simp+Fam, Simp+Nov, and Comp+Fam) from Experiment 1 were used in this experiment. Stimuli were presented in sets composed of a simultaneous presentation of visual and auditory components identical to that in Experiment 1, such that each image's presentation matched the duration of the corresponding sound. For all conditions, six target sets (VIS_TAUD_T) were used, and for each target set, four types of test items were created: (a) a set that was identical to a target set (i.e., VIS_TAUD_T), (b) a set that had a different auditory and visual component from a target set (i.e., $VIS_{new}AUD_{new}$), (c) a set that matched a target set's visual component but had a novel auditory component (i.e., VIS_TAUD_{new}), and (d) a set that had a novel visual component but matched a target set's auditory component (i.e.,

Table 2
Percentages of Children of Each Responder Type by Condition, Experiment 1

Condition	Responder type		
	Visual	Auditory	Mixed
Condition 1:			
Simp+Fam ($N = 15$)	73.3%	6.7%	20.0%
Condition 2:			
Simp+Nov ($N = 15$)	21.4%	42.9%	35.7%
Condition 3:			
Comp+Fam ($N = 15$)	73.3%	0.0%	26.7%
Sloutsky & Napolitano (2003): Comp+Nov ($N = 13$)	15.4%	53.8%	30.8%

Note. Simp = simple visual stimuli; Comp = complex visual stimuli; Nov = novel visual stimuli; Fam = familiar visual stimuli.

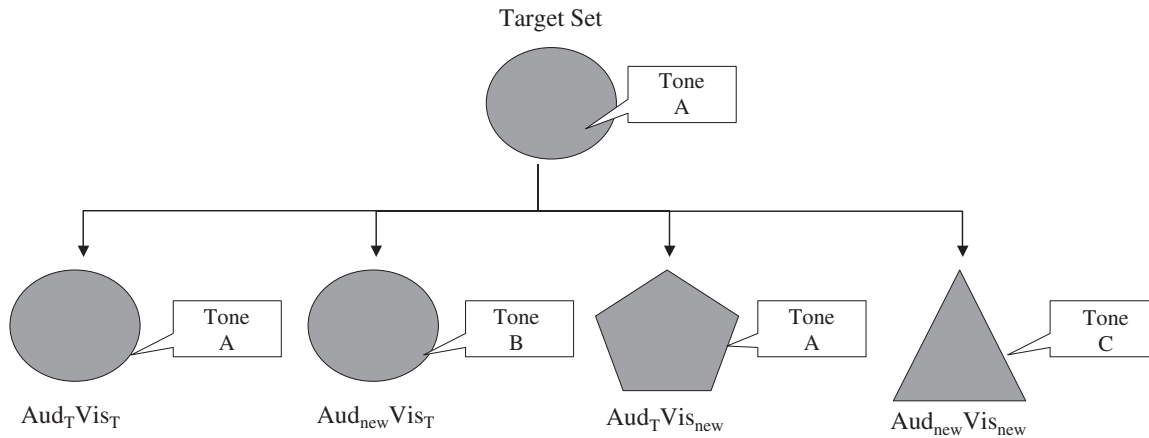


Figure 5. Example of stimulus sets in Experiment 2, Condition 1. VIS = visual component; AUD = auditory component.

VIS_{new}AUD_T). Examples of target and test items are presented in Figure 5.

Design and Procedure

The experiment had a mixed design with the three stimulus conditions varying between participants and the four test item types varying within participants. A female experimenter tested participants individually in a quiet room within their day care centers. Participants were told they would play a matching game, in which they would be shown a picture and a sound (an example was given at this point), and then another picture and sound (an example was given at this point). They would need to decide whether the second set had the same picture and sound as the first set. If it did, they should answer *same*, but if it had a different sound, a different picture, or both, they should answer *different*. The experiment included a total of 24 trials (6 target sets, with 4 test item types per set). On each trial, a target set was presented and then followed by test item, and the participant was then prompted to respond whether the test item was same as the target or different. Stimuli were displayed in the following manner. The target set, VIS_TAUD_T, was presented on the center of the screen for 1 s followed by a blank screen presented for 1 s. Next, one of the four test items was presented on the center of the screen for 1 s followed by a blank screen, and the participant was asked whether the item was exactly the same as the target or different. After receiving the participant's response, the experimenter entered this response and started the next trial. Stimuli were presented on a laptop computer running Superlab Pro 2.0 software (Cedrus Corporation, 1999). The presentation order of the six targets and corre-

sponding test items was randomized. Participants were given small toys at the end of the experiment as rewards for their participation.

Results and Discussion

Data were analyzed to determine whether children were capable of identifying differences in both auditory and visual stimuli. The VIS_TAUD_T and VIS_{new}AUD_{new} items served as controls for the overall accuracy, and whereas the VIS_TAUD_{new} and VIS_{new}AUD_T items indicated whether participants encoded (a) both components (in which cases they should accurately reject both items), (b) only the visual component (in which cases they should accurately reject VIS_{new}AUD_T but not VIS_TAUD_{new} items), or (c) only the auditory component (in which cases they should accurately reject VIS_TAUD_{new} but not VIS_{new}AUD_T items).

Across conditions, participants were accurate in accepting VIS_TAUD_T items ($M_{\text{correct}} > .88$) and in rejecting VIS_{new}AUD_{new} items ($M_{\text{correct}} > .83$), both above chance, one-sample $t_s(14) > 3.84$, $p_s < .01$, with no significant differences in accuracy across the control item types. However, participants' rejection of VIS_TAUD_{new} (indicating encoding of the auditory component) and of VIS_{new}AUD_T (indicating encoding of the visual component) differed markedly across the conditions (see Figure 6).

Proportions of correct responses were subjected to a mixed 3 (condition) \times 2 (test item type: VIS_{new}AUD_T and VIS_TAUD_{new}) ANOVA with the test item type as the repeated measure. There was a significant main effect of test item type, with VIS_{new}AUD_T ($M_{\text{correct}} = .76$) $>$ VIS_TAUD_{new} ($M_{\text{correct}} = .32$), $F(1, 42) = 119$, $p < .01$. Most important, there was a significant Condition \times Test Item Type interaction,

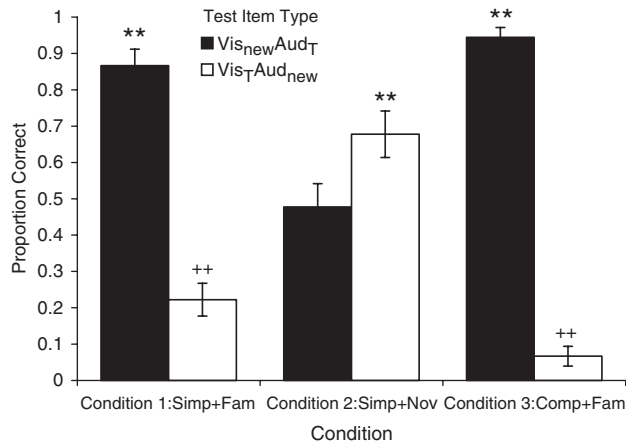


Figure 6. Proportions of correct same/different responses in Experiment 2. Error bars represent standard errors of the mean. VIS_T = target visual component; VIS_{new} = New visual component; AUD_T = Target auditory component; AUD_{new} = new auditory component.; Simp = simple visual stimuli; Comp = complex visual stimuli; Nov = novel visual stimuli; Fam = familiar visual stimuli. **Above chance, $p < .01$. ++Below chance, $p < .01$.

$F(2, 42) > 62.65$, $p < .01$. Paired sample t tests pointed to significant differences in accurately rejecting Vis_{new}Aud_T and Vis_TAud_{new} across the conditions. Although in Conditions 1 and 3 (i.e., Simp+Fam and Comp+Fam) participants were more likely to reject accurately Vis_{new}Aud_T, both paired-sample t s(14) > 10.27 , $ps < .01$, the reverse was true for Condition 2 (i.e., Simp+Nov), in which participants were more likely to reject Vis_TAud_{new} paired-sample $t = -2.15$, $p < .05$.

Recall that according to Hypothesis 1, participants could have failed to encode visual stimuli in the Sloutsky and Napolitano (2003) experiments because visual stimuli in these experiments were scenes and object arrangements. At the same time, it is possible that individual objects would more likely to be attended to rather than complex scenes or arrangements. Results of Experiment 2 eliminate this possibility. Note that all visual stimuli in this experiment (and in Experiment 1) were objects, and the auditory stimuli were the same as those used by Sloutsky and Napolitano. If objecthood was a critical factor in processing, we should have observed differences between Condition 2, where unfamiliar objects were used, and Sloutsky and Napolitano (Experiments 1 and 3), where unfamiliar scenes or arrangements were used. However, in both, Condition 2 of the current experiment (where visual stimuli were objects) and in Sloutsky and Napolitano (Experiments 1 and 3, where visual stimuli were not objects) participants exhibited auditory dominance,

thus indicating that, unlike familiarity, objecthood plays little or no role in modality dominance.

Overall, when visual stimuli were familiar (i.e., Conditions 1 and 3), children were likely to encode the visual stimuli while failing to encode the auditory stimuli. However, when visual stimuli were unfamiliar (i.e., in Condition 2 and in Sloutsky & Napolitano, 2003), children were more likely to encode the auditory stimuli than the visual stimuli. Thus, under all conditions, participants exhibited strong modality dominance effects—they processed stimuli in one modality while failing to process the other modality. Furthermore, familiarity of visual stimuli clearly contributes to the likelihood that young children would process these stimuli.

Recall, however, that all experiments reported thus far employed unfamiliar auditory stimuli. It remains unknown, therefore, whether familiarity of auditory stimuli also contributes to processing of auditory and visual information. On the one hand, it is possible that auditory processing is at ceiling, even when auditory stimuli are novel and familiarity of auditory stimuli would have little or no effect on processing. On the other hand, it is possible that familiarity of auditory stimuli would contribute to processing in the same manner as familiarity of visual stimuli.

To address this issue, we conducted Experiment 3, which uses the same old–new recognition task as in Experiment 2 but employs familiar and well-discriminable sounds as well as familiar visual stimuli. In Experiment 3, we used two conditions: (a) sounds that were more familiar than visual stimuli and (b) sounds that were less familiar than visual stimuli. If processing is controlled by relative familiarity, there should be differences in processing between the conditions. In particular, in the former condition participants would exhibit auditory overshadowing, whereas in the latter condition they would exhibit visual overshadowing.

Experiment 3

Method

Participants

Participants were 30 young children ($M = 4.24$ years, $SD = 0.27$ years, range = 3.8 to 5.0 years; 13 girls and 17 boys). There were 2 between-participants conditions (described later), with 15 children participating in each condition. An additional 2 children did not exhibit above-chance accuracy on control items, and these children were not included in any analyses.

Materials, Design, and Procedure

Auditory stimuli were 12 common sounds, such as a doorbell, dog bark, and dial tone. Each sound was 1 s in duration and matched the presentation of a visual stimulus. Discriminability was established in a calibration experiment using a same–different immediate recognition task. A different sample of fifteen 4-year-olds made correct same–different judgments after being presented with pairs of the auditory stimuli on 97% of trials.

Familiarity was established by asking a different sample of fifteen 4-year-olds two questions about each of the 12 individual auditory stimuli: “Have you ever heard this sound?” and “What is it?” Overall, most of the items were reported to be heard before ($M = 94\%$), and most of the items were labeled correctly and consistently ($M = 90\%$), indicating that the auditory stimuli were familiar.

Two of the previously tested sets of visual stimuli were used in this experiment. For Condition 1, visual stimuli were the single geometric shapes, which were likely to be consistently labeled by young children ($M = 81\%$). For Condition 2, visual stimuli were the animal photographs, which were even more likely to be consistently labeled by young children ($M = 96\%$). Recall that identical sounds were used in both conditions, which were also consistently labeled by young children ($M = 90\%$). Therefore, in the former condition sounds were slightly more familiar than visual stimuli, whereas the reverse was true in the latter condition. The experiment had the same design and procedure as Experiment 2.

Results and Discussion

Similar to Experiment 2, data from the VIS_TAUD_T and $VIS_{new}AUD_{new}$ conditions served as controls, and the data from the VIS_TAUD_{new} and $VIS_{new}AUD_T$ conditions were indicative of whether participants encoded auditory and visual stimuli. Across conditions, participants were accurate in accepting VIS_TAUD_T items ($M_{correct} > .92$) and in rejecting $VIS_{new}AUD_{new}$ items ($M_{correct} > .89$), both above chance, one-sample $t(14) > 9.2$, $ps < .01$, with no significant differences in accuracy across the control item types. At the same time, participants exhibited marked differences in processing auditory and visual stimuli across the conditions (see Figure 7).

Proportions of selections for auditory stimuli were subjected to a mixed 2 (condition) \times 2 (test item type: VIS_TAUD_{new} and $VIS_{new}AUD_T$) ANOVA with the test item type as a repeated measure. Although none of the main effects was significant, there was a sig-

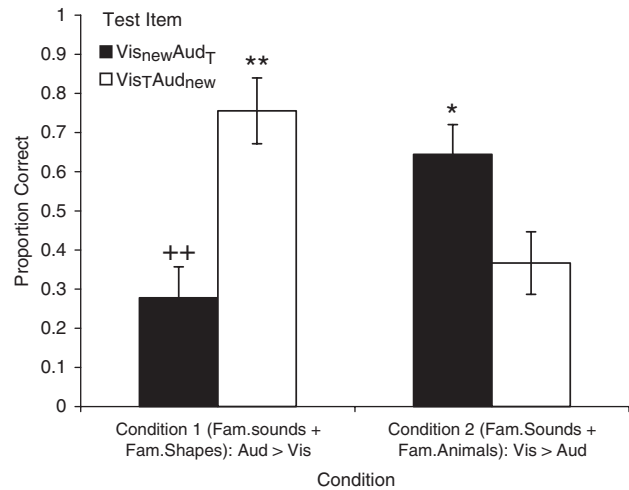


Figure 7. Proportions of correct same/different responses in Experiment 3. Error bars represent standard errors of the mean. VIS_T = target visual component; VIS_{new} = new visual component; AUD_T = target auditory component; AUD_{new} = new auditory component; Fam = familiar; Aud > Vis = auditory stimuli more familiar than visual stimuli; Vis > Aud = visual stimuli more familiar than auditory stimuli. **Above chance, $p < .02$. ++Below chance, $p < .02$. *Marginally above chance, $p < .07$.

nificant Test Item Type \times Condition interaction, $F(1, 28) = 22.12$, $p < .01$. Paired-sample t tests pointed to the following difference: In Condition 1 (i.e., when auditory stimuli were more familiar than visual stimuli) participants were more likely to reject accurately VIS_TAUD_{new} , $t(14) = -4.16$, $p < .01$, whereas in Condition 2 (i.e., when visual stimuli were more familiar than auditory stimuli) participants were more likely to reject $VIS_{new}AUD_T$, $t(14) = 2.47$, $p = .026$.

These results strongly indicate that the relative familiarity of stimuli may moderate modality dominance—more familiar stimuli overshadow processing of less familiar stimuli. The results are remarkable because the same auditory stimuli that overshadowed visual stimuli in Condition 1 (Aud > Vis), where they were more familiar than the corresponding visual stimuli, were overshadowed by visual stimuli in Condition 2 (Vis > Aud), where they were less familiar than the corresponding visual stimuli.

These findings undermine Hypothesis 2 (i.e., that modality dominance effects manifest themselves only as auditory dominance), supporting instead Hypothesis 3 (i.e., that modality dominance is flexible). Findings indicating that more familiar stimuli overshadow less familiar stimuli may have important implications for understanding the time course of processing. In particular, to establish a greater familiarity of one stimulus over the other, the familiarity of these components has to be compared, which in turn suggests that early in the course of processing both components

received some attention, with the dominant modality being selected over the nondominant modality.

Overall, the reported results indicate that when visual and auditory stimuli are presented simultaneously, young children tend to process stimuli in one modality while failing to process stimuli in the other modality, and that attention shifts flexibly between the modalities. Recall that adults were shown to process accurately both auditory and visual stimuli (Sloutsky & Napolitano, 2003).

However, it is reasonable to suggest that there is more difference between shapes and animals than just relative familiarity and that the results of Experiment 3 could also be explained by some of the stimulus types appearing more interesting than others. Thus, it is possible that pictures of animals overshadowed the sounds because the animals were more interesting to young children than the sounds, which was not the case for simple shapes. To eliminate this possibility, we conducted a control experiment (Experiment 4) that used the same sounds as those used in Experiment 3 as well as pictures of animals; these animals, however, were unfamiliar to young children. If familiarity moderates modality dominance, then in the control experiment, the more familiar auditory stimuli should overshadow the less familiar visual stimuli.

Experiment 4

Method

Participants

Participants were 13 young children ($M = 4.34$ years, $SD = 0.37$ years, range = 3.8 to 5.0 years; 8 girls and 5 boys). An additional 2 children also participated, but they did not exhibit above-chance

accuracy on control items, and these children were not included in any analyses.

Materials, Design, and Procedure

Auditory stimuli were the same common sounds used in Experiment 3. Visual stimuli were close-up portrait photographs of unusual animals, such as a porcupine and a cuscus (see Figure 8 for an example of these stimuli). Again, to establish the degree of familiarity of the stimuli, a separate calibration study was conducted using a different sample of fifteen 4-year-olds. Similar to previous experiments, familiarity was established by asking two questions about each individual visual stimulus: "Have you ever seen this animal before?" and "What is it?" Overall, children reported that they recognized the stimuli on a little more than half of the trials ($M = 58%$), whereas they rarely correctly and consistently labeled these animals ($M = 10%$), although they attempted to label the stimuli on approximately half of the trials ($M = 47%$). Based on these responses, it was concluded that the visual stimuli were less familiar than the auditory stimuli. The experiment had the exact same design and procedure as Experiment 3.

Results and Discussion

Overall, children exhibited high levels of accuracy, correctly rejecting $VIS_{new}AUD_{new}$ items ($M_{correct} > .90$) and correctly accepting VIS_TAUD_T items ($M_{correct} > .93$), both above chance, one-sample $t_s(12) > 8.90$, $p_s < .01$, with no significant differences in accuracy across the control item types. More important, children readily noticed changes in the auditory component, correctly rejecting VIS_TAUD_{new}

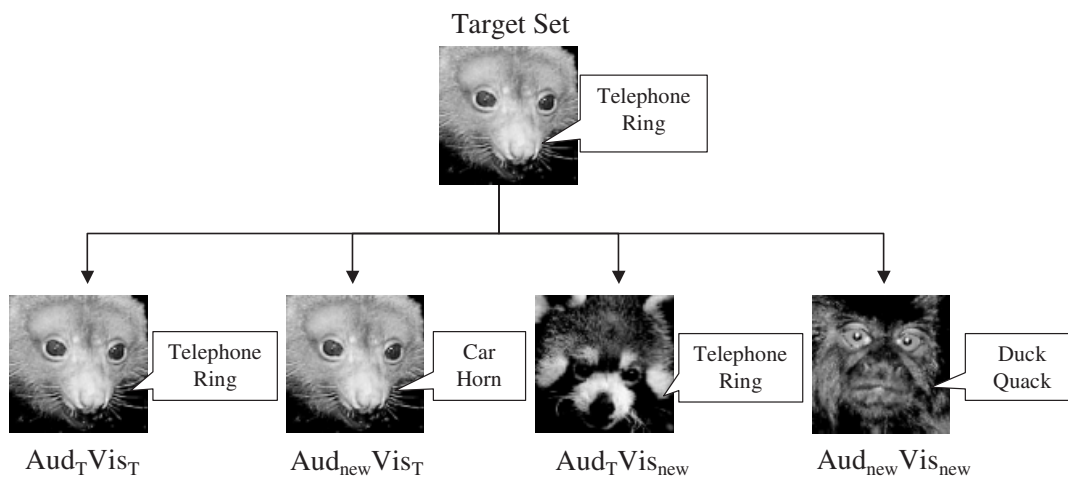


Figure 8. Example of stimulus sets in Experiment 4. VIS_T = target visual component; VIS_{new} = new visual component; AUD_T = target auditory component; AUD_{new} = new auditory component.

items, with accuracy above chance, one-sample $t(12) = 3.48, p < .01$, but did not notice changes in the visual component, erroneously accepting $VIS_{ne-wAUDT}$ items, with accuracy at chance, one-sample $t(12) = -1.2, p > .25$. In short, participants encoded accurately the auditory stimuli while failing to encode reliably the visual stimuli, thus indicating that more familiar auditory stimuli overshadowed less familiar visual stimuli. Therefore, the results of Experiment 3, where visual stimuli overshadowed auditory stimuli, are more likely to stem from greater familiarity of the visual stimuli than from them being the more interesting stimuli.

Overall, results of Experiments 1 to 4 reveal flexible modality dominance and factors underlying shifts in modality dominance. These findings extend those of Sloutsky and Napolitano (2003) and, as mentioned earlier, may have important implications for our understanding of lexical development. However all stimuli used in Experiments 1 to 4 were nonlinguistic sounds, and it is unclear whether findings with these stimuli would generalize to linguistic stimuli. For example, it could be argued that if familiarity does drive modality dominance, then the task of word learning should be impossible because new words are unfamiliar and, as such, could not be attended to when they accompany even somewhat familiar entities. Therefore, because we know that young children do acquire unfamiliar words, the findings reported in Experiments 1 to 4 may have little or no implication for lexical development.

Alternatively, it could be argued that the familiarity of auditory stimuli is determined by the familiarity of its source (see Ballas, 1993; Cykowicz & Friedman, 1998, for related discussions). If this is the case, then even unfamiliar words should represent a class of highly familiar sounds, as their source (i.e., human voice) is highly familiar to young children, and findings reported in Experiments 1 to 4 should predict processing of cross-modal compounds in which linguistic stimuli are paired with visual stimuli. In particular, we can predict that if sounds are produced by human voice (even if strings of sounds unfamiliar and not word-like), then these stimuli would overshadow less familiar visual stimuli. This hypothesis was tested in Experiment 5.

Experiment 5

Method

Participants

Participants were 30 young children ($M = 4.45$ years, $SD = 0.26$ years, range = 4.0 to 4.9 years; 16

girls and 14 boys). There were two between-participants conditions (described later), with 15 children participating in each condition. An additional 6 children also participated in Experiment 4, but they did not exhibit above-chance accuracy on control items, and these children were not included in any analyses.

Materials, Design, and Procedure

Auditory stimuli were nonsensical three-vowel sequences that did not resemble English words (e.g., “[e] - [i] - [u]”). Syllable sequences were created by recording a human speaker generating three syllables and then cutting each individual syllable to a uniform length of 0.33 s using the audio program CoolEdit 2000 so that each sound is approximately 1 s in duration. These stimuli were presented at the average sound level of 67.8 dB (with a range from 66 to 72 dB), which is comparable to the sound level of human voice in a regular conversation. Discriminability was established in a separate calibration experiment using a same-different immediate recognition task. A different sample of fifteen 4-year-olds correctly made same-different judgments after being presented with pairs of the auditory stimuli on 94% of trials.

Familiarity was established by asking a different sample of ten 4-year-olds to make a source attribution for different types of sounds on a forced-choice task. This familiarity task differed somewhat from the task described in Experiments 1, 3, and 4. Two classes of sounds were used: (a) vowel patterns (e.g., “[u] - [u] - [e]”) and (b) the familiar animal sounds (e.g., dog bark) that constituted half of the familiar sounds used in Experiments 3 and 4. In the task, participants were told they would be playing a game in which they would need to guess who was making the funny noise. Each child participated in 24 trials. For each trial participants were presented a sound with four different pictures of possible sources of the sound (see Figure 9 for an example of choice option) and asked, “Which of these do you think made this sound?” The pictures varied across trials but always included a picture of a man, a familiar animal, an unfamiliar animal, and a question mark (it was explained to the child before the task began that if they were not sure who produced the sound they should point to the question mark). In half of the trials, pointing to the picture of the man was the correct response, and in the other half, pointing to the picture of the familiar animal was the correct response. Overall, children correctly attributed vowel patterns ($M = 94%$) as well as familiar animal sounds



Figure 9. Example of picture choices for a trial in familiarity control task, Experiment 5.

($M = 98\%$). The latter measure is compatible with the measure of familiarity of familiar sounds presented in Experiment 3, thus indicating that this procedure was an adequate measure of familiarity.

Again, two of the previously tested sets of familiar visual stimuli were used in this experiment. For Condition 1, visual stimuli were the familiar single geometric shapes. For Condition 2, visual stimuli were the familiar animal photographs. The experiment had the same design and procedure as Experiments 2 to 4.

Results and Discussion

Again, data from the VIS_TAUD_T and $VIS_{new}AUD_{new}$ conditions served as controls, and the data from the VIS_TAUD_{new} and $VIS_{new}AUD_T$ conditions were indicative of whether participants encoded auditory and visual stimuli. Across conditions, participants were accurate in accepting VIS_TAUD_T items ($M_{correct} > .93$) and in rejecting $VIS_{new}AUD_{new}$ items ($M_{correct} > .88$), both above chance, one-sample $t(14) > 11.00$, $ps < .01$, with no significant differences in accuracy across the control item types. However, their rejection of VIS_TAUD_{new} (indicating encoding of the auditory component) and of $VIS_{new}AUD_T$ (indicating encoding of the visual component) differed markedly across the conditions (see Figure 10).

Proportions of selections for auditory stimuli were subjected to a mixed 2 (condition) \times 2 (test item type: VIS_TAUD_{new} and $VIS_{new}AUD_T$) ANOVA with the

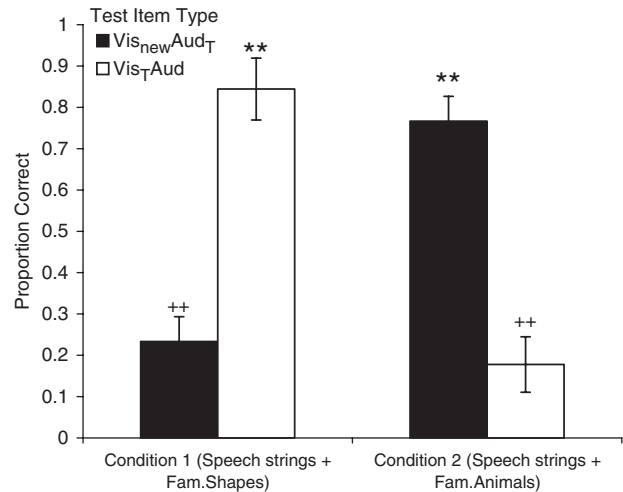


Figure 10. Proportions of correct same/different responses in Experiment 5. Error bars represent standard errors of the mean. VIS_T = target visual component; VIS_{new} = new visual component; AUD_T = target auditory component; AUD_{new} = new auditory component; Fam = familiar. **Above chance, $p < .01$. ++Below chance, $p < .01$.

test item type as a repeated measure. Although none of the main effects was significant, there was a significant Test Item Type \times Condition interaction, $F(1, 28) = 85.24$, $p < .01$. Paired-sample t tests pointed to the following difference: In Condition 1 (i.e., Speech Strings+Fam.Shapes) participants were more likely to accurately reject VIS_TAUD_{new} , $t(14) = 7.26$, $p < .01$, whereas in Condition 2 (i.e., Speech Strings+Fam.Animals) participants were more likely to reject $VIS_{new}AUD_T$, $t(14) = 6.12$, $p < .01$.

These results suggest that the pattern of responses was similar to that with familiar sounds (see Experiment 3): When visual stimuli were geometric shapes, participants exhibited auditory dominance, whereas when visual stimuli were familiar animals, participants exhibited visual dominance. These results suggest that even unfamiliar linguistic stimuli (including unfamiliar words) may have an advantage over somewhat familiar visual stimuli because linguistic stimuli stem from a highly familiar source—human speech. These results, in conjunction with those of Experiment 3 (Condition 2: Vis > Aud), suggest an interesting and testable hypothesis: When word–object pairs are presented for limited time, it may be more difficult for young children to encode new words when they are paired with highly familiar objects than it is to encode new words when they are paired with less familiar or novel objects. If confirmed, this hypothesis may provide an interesting extension of research on

mutual exclusivity, or the tendency of children to extend novel words to novel objects (e.g., Markman & Wachtel, 1988; Merriman & Schuster, 1991).

Overall, in the five experiments reported here, participants exhibited modality dominance: They processed information presented in one modality, but not both. Furthermore, modality dominance was moderated by familiarity: Stimuli that were fully processed when paired with less familiar stimuli from the other modality were overshadowed when paired with more familiar stimuli from the other modality. The robustness, as well as low variability, of modality dominance effects (very few participants failed to exhibit these effects), suggests that modality dominance stems from automatic pulls on attention rather than from selective attending to a particular modality. To examine this issue directly, we conducted Experiment 6. In Experiment 6, 4-year-olds were presented with the same visual–auditory compounds as in Experiment 3, Condition 1 (familiar single geometric shapes and familiar sounds) and asked to attend to visual stimuli, though auditory stimuli were not mentioned. Recall that these were the same visual stimuli that were fully processed when paired with unfamiliar sounds in Condition 1 of Experiment 1. If modality dominance stems from deliberate selective attention to a particular modality, auditory dominance in this experiment should disappear, with participants ably encoding these visual stimuli.

Experiment 6

Method

Participants

Participants were 18 young children ($M = 4.25$ years, $SD = 0.35$ years, range = 3.8 to 4.9 years; 12 girls and 6 boys). One additional child also participated but did not exhibit above-chance accuracy on control items and was not included in any analyses.

Materials, Design, and Procedure

Stimuli were the same familiar single geometric shapes and familiar sounds as used in Experiment 3, Condition 1. The task and procedure were the same as those in Experiments 2 to 5 with one critical exception: Participants were explicitly prompted to attend to visual stimuli. Participants were given the same initial instructions before starting the experiment (i.e., they were told they were going to play a matching game where they would be shown a picture of a shape that has a sound with it followed by a

second shape with a sound and they would then be asked if the second shape with sound was the same as the first shape with sound). However, before each trial, children were explicitly asked to attend to visual stimuli, with sounds not being mentioned: “Don’t forget, you have to remember the picture!”

Results and Discussion

Overall, children exhibited high levels of accuracy, correctly rejecting $VIS_{new}AUD_{new}$ items ($M_{correct} > .95$) and correctly accepting VIS_TAUD_T items ($M_{correct} > .95$), both above chance, one-sample $t(16) > 17.26$, $ps < .01$, with no significant differences in accuracy across the control item types. They also exhibited high levels of accuracy rejecting VIS_TAUD_{new} items ($M_{correct} = .74$), above chance, one-sample $t(16) = 2.76$, $p < .05$. At the same time, participants were not above chance in correctly rejecting $VIS_{new}AUD_T$ items, one-sample $t(16) < 1$. Therefore, despite the repeated explicit instruction to attend to visual stimuli, 4-year-olds continued to exhibit auditory dominance. Recall that these same visual stimuli were ably processed when paired with different sounds in Condition 1 of Experiment 2. These results indicate that modality dominance is unlikely to stem from deliberate selective attention to a particular modality, but it is more likely to stem from automatic pulls on attention.

General Discussion

Several important findings emerged from this research. Experiment 1 examined modality preference effects using a modified switch procedure and stimuli that varied in complexity and familiarity. The results indicated that young children exhibited auditory preference when novel visual stimuli were paired with novel auditory stimuli, whereas they exhibited visual preference when familiar visual stimuli were paired with novel auditory stimuli. Experiment 2 examined whether participants encode the nonselected modality. The results indicated that when both auditory and visual stimuli are novel, young children exhibit auditory dominance, whereas when only visual stimuli are familiar, young children exhibit visual dominance. At the same time, there is little evidence that stimulus complexity plays an important role in modality dominance.

Experiments 3 and 4 indicate that relative familiarity moderates modality dominance: When auditory stimuli were more familiar than visual stimuli, young children exhibited auditory dominance, whereas when visual stimuli were more familiar

than auditory stimuli, they exhibited visual dominance. Experiment 5 expanded these findings to human speech, which represents highly familiar sounds: Even unfamiliar patterns of human speech elicited auditory dominance effects similar to those elicited by other familiar sounds. These results are remarkable because the same visual stimuli that received full processing and overshadowed corresponding auditory stimuli in Experiments 1 and 2 received little processing and were overshadowed by corresponding auditory stimuli in Experiments 3 to 5.

Finally, Experiment 6 indicated that when presented with auditory–visual compounds that yielded auditory dominance and repeatedly asked to attend to visual stimuli, young children still continued to exhibit auditory dominance. This failure to attend selectively to a particular stimulus component in conjunction with the robustness and low variability of modality dominance effects suggest that modality dominance stems from automatic pulls on attention rather than from deliberate selective attention.

At the most general level, these results point to several regularities. First, there are modality dominance effects, such that under some conditions the auditory modality dominates the visual modality, whereas under other conditions the reverse is true. Second, there are general auditory and familiarity effects, such that when visual and auditory stimuli are unfamiliar, young children exhibit auditory dominance; otherwise, more familiar stimuli dominate processing of less familiar stimuli. Third, although there is evidence for resource shifting across modalities, there is little evidence for resource sharing: Under all conditions (Experiments 2 to 6) one modality received full processing rather than both modalities receiving some processing. In particular, although a dominant modality exhibited reliable above-chance accuracy, accuracy in the other modality never exceeded chance. And finally, these shifts are likely to stem from automatic pulls on attention rather than from deliberate selective attention.

These findings support Hypothesis 3, revealing the flexible nature of modality dominance, implicating automatic attention in flexible modality dominance, and pointing to factors moderating modality dominance: When both visual and auditory stimuli are unfamiliar, young children exhibit auditory dominance; otherwise, more familiar stimuli dominate less familiar stimuli. Furthermore, these findings are novel, and they have important theoretical implications for our understanding of the

development of attention, of cross-modal processing, and of some aspects of language acquisition and conceptual development. In what follows, we discuss these points in greater detail.

Modality Dominance and Young Children's Attention

The fact that young children exhibit modality dominance effects and the fact that the dominance shifts flexibly between modalities reveal important properties of young children's attention. Overall, two important points pertaining to attention in young children stem from the reported research. First, when presented with arbitrarily paired bimodal stimuli, young children are more likely to shift flexibly resources across modalities than to share resources. Second, resource shifting stems from automatic pulls on attention rather than from deliberate selective attention to one modality.

The fact that under these conditions young children shift attentional resources across modalities rather than share them is a novel finding. However, it remains unclear whether this tendency (a) stems from increased processing demands in these cross-modal tasks (in which case, participants should process both modalities if processing demands decrease), or (b) is a more general characteristic of cross-modal processing. These issues have to be addressed in future research.

The robustness of familiarity and modality dominance effects suggests that modality dominance stems from automatic pulls on attention rather than from deliberate selective attention to a single modality. This issue was examined more directly in Experiment 6: When presented with auditory–visual sets that elicited auditory dominance in Experiment 3 (Condition 1) and instructed to attend to pictures (with sounds not being mentioned), young children continued to exhibit auditory dominance. Therefore, there is little evidence that young children deliberately select a particular modality that will be attended to, and it seems more likely that modality dominance is driven by automatic pulls on attention.

Theoretical Implications of Flexible Modality Dominance

The reported flexible modality dominance effects have important theoretical implications. These findings may affect our understanding of (a) the development of cross-modal processing, (b) some aspects of lexical development, and (c) the role of language in conceptual development.

First, cross-modal processing in infants and children has been characterized by two sets of findings.

On the one hand, there is a large body of evidence that very young infants efficiently process cross-modal stimuli, ably detecting amodal relations, including temporal synchrony, rhythm, or tempo (Bahrick, 1988, 2001, 2002; Kuhl & Meltzoff, 1982; Lewkowicz, 2000b; Meltzoff & Borton, 1979; Slater, Quinn, Brown, & Hayes, 1999; see also Lewkowicz, 2000a; Lickliter & Bahrick, 2000, for extensive reviews). On the other hand, there have been several reports pointing to the dominance of the auditory modality in infancy (Lewkowicz, 1988a, 1988b; Robinson & Sloutsky, 2004; see also Lewkowicz, 1994, for a review) and early childhood (Sloutsky & Napolitano, 2003). Findings reported here clearly extend the latter line of research.

Why do children efficiently process both modalities when there is an amodal relation but exhibit modality dominance when there are arbitrary relations? We consider several possibilities. First, the efficient processing of amodal relations may stem from intersensory redundancy created by an amodal relation. This redundancy may recruit attention, facilitating the efficient processing of both modalities (Bahrick & Lickliter, 2000). Second, it is also possible that some amodal relations used in research on cross-modal processing represent a special case of temporal processing, and there is evidence that temporal relations might be underlined by a single processing system (see Pashler, 1998).

A third possibility (which does not exclude the former two possibilities) is that the divergence stems from the significantly longer stimulus presentation in research on processing of cross-modal relations than that used in the current research. If we hypothesize that the auditory modality habituates faster than the visual modality, then this presentation time difference may play a critical role because a longer presentation may wash out auditory dominance effects: Participants may habituate faster to auditory stimuli, and this asynchronous habituation may enable them to process corresponding visual stimuli. However, more research is needed to examine fully these possibilities.

The reported modality dominance and familiarity effects may also have important implications for our understanding of some aspects of language acquisition, such as the ability of infants and young children to acquire words. In particular, visual processing is largely parallel, whereas auditory processing is largely serial. Furthermore, auditory events are dynamic and transient, whereas visual objects and scenes are usually stable. Therefore, auditory dominance may play an important role in the ability of infants and young children to encode novel words: It

might be difficult to attend to these transient auditory stimuli in the absence of the auditory dominance.

Auditory dominance effects may also be amplified by the high familiarity of human speech: Results of Experiments 5 indicate that even meaningless strings of human speech (e.g., a string of vowels "[e] - [i] - [u]") belong to a class of familiar sounds and thus are more likely to be processed than somewhat less familiar visual stimuli.

Modality dominance effects moderated by familiarity may have also important implications for understanding of the role of language in conceptual development. Recall that auditorily presented linguistic labels often play an important role in young children's conceptual organization and thinking. Two classes of explanations of the role of linguistic labels in conceptual and semantic tasks have been proposed (see Sloutsky & Napolitano, 2003, for a review), one arguing for language-specific effects and another arguing for general auditory effects.

First, there is a proposal claiming that effects of labels are due to the fact that labels are linguistic stimuli. This language-specific proposal has two variants: semantic and prosodic. The semantic proposal argues that the effects of linguistic labels stem from two important assumptions that young children hold: (a) that entities are members of categories and (b) that labels presented as count nouns convey category membership (Gelman & Coley, 1991). These assumptions lead young children to infer that entities that are denoted by the same count noun belong to the same category (Gelman & Markman, 1986; Markman, 1989; see also Waxman & Markow, 1995, for a discussion). The prosodic proposal argues that facilitative effects of linguistic labels might not be limited to semantic effects but that additional effects might be due to infants' and young children's special attention to the prosodic components of human speech that distinguish speech from other sounds (Balaban & Waxman, 1997).

There is also a general auditory proposal (e.g., Sloutsky & Napolitano, 2003) arguing that effects of labels might not be limited to language-specific factors but that these effects may also stem from general auditory factors, such as privileged processing of auditory information by young children.

Present findings in conjunction with the earlier report (Sloutsky & Napolitano, 2003) indicate that the prominent role of linguistic labels may be explained in part by general auditory effects and in part by familiarity effects. Of course, these results do not rule out either of the language-specific explanations, and some of the language-specific effects could be due to familiarity effects. Furthermore, it is

possible that there are different levels of familiarity, with each contributing to the familiarity of linguistic labels. First, there is a crude physical distinction between speech sounds and nonspeech sounds, with speech sounds being a more familiar set of stimuli, and as demonstrated in Experiment 5, the familiarity of human speech may contribute to modality dominance effects. Second, there is a distinction between speech patterns that approximate phonological, morphosyntactic, prosodic, and articulatory properties of the native language and those that do not, with the former being more familiar than the latter (cf. Johnson & Jusczyk, 2001). Finally, there is a distinction between known words and novel words, with the former being more familiar than the latter. In this case, familiar nouns (e.g., *cat* or *dog*) would have the highest familiarity, whereas strings that are improbable in the native language (e.g., the strings used in Experiment 5 are improbable in English) would have the lowest familiarity, with the former dominating over a wider range of visual stimuli than the latter. This auditory dominance may manifest itself in overshadowing (if visual stimuli are presented for limited time, such as in the present research) or in strong effects of word-like linguistic labels in a variety of conceptual tasks (where visual stimuli are presented for ample time). However, additional research is needed to test these contentions.

If novel auditory stimuli overshadow novel visual stimuli, and more familiar stimuli overshadow less familiar stimuli, how do children map novel words onto novel entities? If both auditory and familiarity factors attract attention only to the word, but not to the novel entity the word denotes, such mappings should be impossible. Yet young children often easily map words onto objects (e.g., Carey & Bartlett, 1978; Markson & Bloom, 1997; see Woodward & Markman, 1998, for a review). We think that the key to resolving this apparent contradiction is that such mappings occurred when objects were either presented for a longer period or were presented repeatedly. It seems that both presentation conditions may increase the probability of encoding of visual stimulus (for arguments on why modality dominance might disappear under substantially longer or repeated presentation conditions, see the previous discussion on the processing of cross-modal stimuli). However, additional research is necessary to address directly the interrelationships between overshadowing and fast mapping.

Limitations and Future Directions

Although modality dominance effects reported here are robust, more research is needed to examine

the scope of generalization of these effects. In particular, most findings reported here (Experiments 2–6) stem from a single paradigm that included the immediate recognition task with particular task conditions (i.e., synchronous presentation of cross-modal stimuli with a relatively short exposure time), and the use of a single paradigm might limit the scope of generalization of these findings.

For example, it could be argued that relatively limited presentation time could have prevented participants from encoding stimuli presented in both modalities, and modality dominance effects would disappear were participants given more time to encode cross-modal stimuli. Although substantially longer presentation time may eliminate modality dominance effects (see the previous arguments), we do not believe that modality dominance effects stem solely from participants not having enough time to process stimuli in both modalities. In particular, we have preliminary evidence that when the exposure time for both familiar visual stimuli (familiar geometric shapes) and familiar auditory stimuli was increased to 2 s, patterns of encoding remained essentially the same. Although one may argue that 2 s was not enough to encode both shapes and sounds, this possibility seems unlikely, given that under a single modality condition, it typically takes no more than 500 ms to encode these stimuli.

Note that in this research and in Sloutsky and Napolitano (2003) visual and auditory stimuli were presented synchronously, and it is unclear whether these results would generalize to asynchronous presentation of auditory and visual stimuli. In particular, it is unclear whether the current findings could be extended to situations where stimuli are presented asynchronously, such as a variety of lexical and conceptual tasks where the visual entity remains stationary, whereas the auditory stimulus is presented for a relatively short period. We have preliminary results indicating that when participants are presented with a string of vowels and familiar geometric shapes (similar to those used in Experiment 5), with visual stimulus appearing for 500 ms before the sound, 4-year-olds continue to exhibit auditory dominance. Although these preliminary results are informative, a 500-ms stimulus onset asynchrony (SOA) may be insufficient, and a more systematic investigation is necessary to examine whether a more pronounced SOA would change the reported pattern of results.

Finally, the fact that the reported effects stem from a single task may limit generalization of these findings to other types of tasks. Although it has been recently demonstrated (Robinson & Sloutsky, 2004)

that the modality preference and modality dominance effects reported here can be replicated with a completely different task, a variant of an inductive generalization task, more research is needed to examine whether modality dominance effects persist with a wider range of tasks.

Conclusions

In sum, the reported research points to several important findings. These findings replicate and further extend results reported by Sloutsky and Napolitano (2003). First, this research indicates that young children exhibit a default auditory dominance: When both auditory and visual stimuli are unfamiliar, young children tend to process auditory stimuli while failing to process visual stimuli. Second, auditory dominance is a special case of modality dominance: When auditory and visual stimuli are presented simultaneously, young children tend to process stimuli presented in one modality while failing to process the other modality. Third, in young children, modality dominance shifts flexibly: Under some conditions particular visual stimuli overshadow particular corresponding sounds, whereas under other conditions these same visual stimuli are overshadowed by different sounds. Fourth, modality dominance is moderated by stimulus familiarity: Young children process more familiar stimuli while failing to process less familiar stimuli. And fifth, modality dominance is likely to stem from automatic pulls on attention rather than from a deliberate selective process.

References

- Bahrnick, L. E. (1988). Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development, 59*, 197–209.
- Bahrnick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology, 79*, 253–270.
- Bahrnick, L. E. (2002). Generalization of learning in three-and-a-half-month-old infants on the basis of amodal relations. *Child Development, 73*, 667–681.
- Bahrnick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190–201.
- Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64*, 3–26.
- Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception & Performance, 19*, 250–267.
- Birnholtz, J. C., & Benaceraff, B. B. (1983). The development of human fetal hearing. *Science, 222*, 516–518.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17–29.
- Cedrus Corporation (1999). Super Lab Pro (version 2.0). San Pedro, CA: Author.
- Christie, J., & Klein, R. (1995). Familiarity and attention: Does what we know affect what we notice? *Memory & Cognition, 23*, 547–550.
- Cycowicz, Y. M., & Friedman, D. (1998). Effect of sound familiarity on the event-related potentials elicited by novel environmental sounds. *Brain & Cognition, 36*, 30–51.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48*, 269–297.
- Gelman, S. A., & Coley, J. (1991). Language and categorization: The acquisition of natural kind terms. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 146–196). New York: Cambridge University Press.
- Gelman, S. A., & Markman, E. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.
- Hölscher, C., Rolls, E. T., & Xiang, J.-Z. (2003). Perirhinal cortex neuronal activity related to long-term familiarity memory in the macaque. *European Journal of Neuroscience, 18*, 2037–2046.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory & Language, 44*, 548–567.
- Jones, S. S., & Smith, L. B. (2002). How children know the relevant properties for generalizing object names. *Developmental Science, 5*, 219–232.
- Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*, 499–516.
- Jusczyk, P. W. (1998). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*, 449–476.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218*, 1138–1141.
- Lamberts, K. (1994). Flexible tuning of similarity in exemplar-based categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 1003–1021.
- Lewkowicz, D. J. (1988a). Sensory dominance in infants: 1. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155–171.
- Lewkowicz, D. J. (1988b). Sensory dominance in infants: 2. Ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 172–182.

- Lewkowicz, D. J. (1994). The development of intersensory perception in human infants. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of perception: Comparative perspectives* (pp. 165–203). Hillsdale, NJ: Erlbaum.
- Lewkowicz, D. J. (2000a). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, *126*, 281–308.
- Lewkowicz, D. J. (2000b). Infants' perception of the audible, visible, and bimodal attributes of multimodal syllables. *Child Development*, *71*, 1241–1257.
- Lickliter, R., & Bahrick, L. E. (2000). The development of infant intersensory perception: Advantages of a comparative convergent-operations approach. *Psychological Bulletin*, *126*, 260–280.
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
- Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*, 1–27.
- Markman, E., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, *20*, 121–157.
- Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*, 813–815.
- Meltzoff, A. N., & Borton, R. W. (1979). Intermodal matching by human neonates. *Nature*, *282*, 403–404.
- Merriman, W. E., & Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development*, *62*, 1288–1301.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.
- Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Pashler, H., Johnston, J. C., & Ruthruff, E. (2000). Attention and performance. *Annual Review of Psychology*, *52*, 629–651.
- Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*, 1387–1401.
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, *80*, 159–177.
- Slater, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory-visual pairings. *Developmental Science*, *2*, 333–338.
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General*, *133*, 166–188.
- Sloutsky, V. M., & Lo, Y. (1999). How much does a shared name make things similar? Part 1: Linguistic labels and the development of similarity judgment. *Developmental Psychology*, *6*, 1478–1492.
- Sloutsky, V. M., Lo, Y., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*, *72*, 1695–1709.
- Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*, 822–833.
- Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism. *Cognition*, *60*, 143–171.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*, 1289–1309.
- Woodward, A. L., & Markman, E. M. (1998). Early word learning. In D. Kuhn & R. Siegler (Vol. Eds.) W. Damon (Series Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., 371–420). New York: Wiley.
- Xiang, J.-Z., & Brown, M. W. (1998). Differential neuronal encoding of novelty, familiarity, and recency in regions of the anterior temporal lobe. *Neuropharmacology*, *37*, 657–676.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*, 37–63.