JSLHR

Article

Effect of Perceptual Load on Semantic Access by Speech in Children

Susan Jerger,[a] Markus F. Damian,[b] Candice Mills,[a] James Bartlett,[a] Nancy Tye-Murray,[c] and Hervé Abdi[a]

Purpose: To examine whether semantic access by speech requires attention in children.
Method: Children (*N* = 200) named pictures and ignored distractors on a cross-modal (distractors: auditory–no face) or multimodal (distractors: auditory–static face and audiovisual–dynamic face) picture word task. The cross-modal task had a low load, and the multimodal task had a high load (i.e., respectively naming pictures displayed on a blank screen vs. below the talker's face on his T-shirt). Semantic content of distractors was manipulated to be related vs. unrelated to the picture (e.g., picture "dog" with distractors "bear" vs. "cheese"). If irrelevant semantic content manipulation influences naming times on both tasks despite variations in loads, Lavie's (2005) perceptual load model proposes that semantic access is independent of capacity-limited attentional resources; if, however, irrelevant content influences naming only on the cross-modal task (low load), the perceptual load model proposes that semantic access is dependent on attentional resources exhausted by the higher load task.
Results: Irrelevant semantic content affected performance for both tasks in 6- to 9-year-olds but only on the cross-modal task in 4- to 5-year-olds. The addition of visual speech did not influence results on the multimodal task.
Conclusion: Younger and older children differ in dependence on attentional resources for semantic access by speech.

Key Words: perceptual load, semantic access, children, picture word task, audiovisual speech, development

Understanding speech is a complex task that critically depends on a quick and reliable retrieval of the meaning of words. Speakers produce about three words per second (Bloom, 2000), and the ease with which listeners decode such a rapid stream of information suggests a remarkably efficient retrieval system. How do children develop such a system, and what are some of the changes that occur as children's understanding of spoken words matures? A central issue in this regard has concerned whether accessing the meaning of spoken words requires attention. We conceptualize attention as a capacity-limited pool of resources shared among concurrent tasks or stimuli (e.g., Cowan, 1995; Kahneman, 1973; Ricker, Cowan, & Morey, 2010; Saults & Cowan, 2007). In this article, we investigate whether younger and older children differ in their dependence on attentional resources to access the meaning of spoken words. The existing research on this issue has applied the general rationale of presenting speech that is irrelevant to a given task and assessing whether the semantic content of the to-be-unattended speech influences task performance.

[a]University of Texas at Dallas
[b]University of Bristol, United Kingdom
[c]Central Institute for the Deaf at Washington University School of Medicine, St. Louis, MO
Correspondence to Susan Jerger: sjerger@utdallas.edu

Editor: Sid Bacon
Associate Editor: Steve Goldinger
Received July 17, 2011
Revision received January 27, 2012
Accepted July 3, 2012
DOI: 10.1044/1092-4388(2012/11-0186)

Semantic Access by Irrelevant Speech

Table 1 summarizes the range of behavioral tasks, together with key references, that have investigated whether adults and children access the semantic content of to-be-ignored speech. In all tasks, participants respond to a speech or visual target while attempting to ignore an irrelevant speech stimulus or speech dimension. The semantic content of the irrelevant speech is systematically manipulated. If performance on the target task is affected, results imply that the to-be-ignored content was processed to a semantic level (e.g., Bowers, Davis, Mattys, Damian, & Hanley, 2009). Below we draw together the existing findings about semantic access by irrelevant speech (see Table 1). We also review the stimuli and the primary tasks to promote our subsequent goal of understanding whether the influence of irrelevant speech varies as a function of the processing load of primary tasks.

In *dichotic listening tasks* with two opposing speech inputs, individuals are instructed to respond to the speech targets in one ear and to ignore the irrelevant speech in the other ear. The semantic content of the irrelevant speech is varied in numerous ways, such as manipulations to bias the

Table 1. Results in literature indicating that semantic access by irrelevant speech is mandatory, occurring despite a listener's attempts to ignore.

Task	Approach	Illustrative stimuli	Results
	Two opposing speech inputs		
Dichotic listening[a]	Respond to target speech in one ear; ignore irrelevant speech with varying semantic content in other ear	Target speech: ambiguous word— *bank*; irrelevant speech: *river* vs. *money*	Semantic content of irrelevant speech influences interpretation of or performance for target words
	Two opposing dimensions of one speech input		
Speech Stroop-like[b]	Respond to target speech dimension: talker gender; ignore irrelevant speech dimension: conflicting vs. congruent semantic word content	Target talker gender: male; irrelevant word: *mommy* vs. *daddy*	Semantically conflicting content of irrelevant words slows talker-gender response times
	One speech input opposing one visual input (cross-modal)		
Cross-modal Stroop-like[c]	Respond to color of target visual stimulus; ignore irrelevant conflicting vs. congruent semantic content of speech	Target visual stimulus: blue square; irrelevant speech: *red* vs. *blue*	Semantically conflicting content of irrelevant speech slows color response times
Irrelevant speech effect (memory)[d]	Free recall of visual categorically related words; ignore irrelevant related vs. unrelated semantic content of speech	Target visual stimuli: fruit words; irrelevant speech: *bean* vs. *door*	Semantically related content of irrelevant speech disrupts visual recall
Irrelevant speech effect (reading)[e]	Read and comprehend written text; ignore irrelevant related vs. unrelated semantic content of speech	Target visual stimuli: written text; irrelevant speech: related vs. unrelated semantic prose content	Semantically related content of irrelevant speech disrupts reading comprehension
Picture word[f]	Name pictured objects; ignore irrelevant related vs. unrelated semantic content of speech	Target visual stimuli: picture of dog; irrelevant speech: *bear* vs. *flag*	Semantically related content of irrelevant speech slows picture-naming times

[a]References for dichotic listening task (see Styles [1997] for a broader discussion of dichotic listening studies/results): *Adults:* Corteen & Wood (1972); Lackner & Garrett (1972); Lewis (1970); Mackay (1973); Smith & Groen (1974). *Children:* Not applicable. [b]References for speech Stroop-like task (see MacLeod [1991] for a broader discussion of Stroop studies/results): *Adults:* Gregg & Purdy (2007); Green & Barber (1981, 1983); Henkin, Yaar-Soffer, Gilat, & Muchnik (2010); Lew, Chmiel, Jerger, Pomerantz, & Jerger (1997). *Children:* Jerger, Elizondo, Dinh, Sanchez, & Chavira (1994); Jerger, Martin, & Pirozzolo (1988); Jerger et al. (1993); Most, Sorber, & Cunningham (2007). [c]References for cross-modal Stroop-like task: *Adults:* Cowan & Barron (1987); Elliott & Cowan (2001); Elliott, Cowan, & Valle-Inclan (1998). *Children:* Hanauer & Brooks (2003). [d]References for irrelevant speech effect (memory) task (see Banbury, Macken, Tremblay, & Jones [2001] for a broader discussion of studies/results): *Adults:* Beaman (2004); Marsh, Hughes, & Jones (2008, 2009); Neely & LeCompte (1999); Sorqvist, Marsh, & Jahncke (2010). *Children:* Not applicable. [e]References for irrelevant speech effect (reading) task: *Adults:* Jones, Miles, & Page (1990); Martin, Wogalter, & Forland (1988); Oswald, Tremblay, & Jones (2000). *Children:* Not applicable. [f]References for picture word task (see Levelt et al. [1999] for a discussion of the variety of studies/results, including pictures accompanied by written words): *Adults:* Damian & Martin (1999); Hantsch, Jescheniak, & Schriefers (2009); La Heij, Dirkx, & Kramer (1990); Levelt, Roelofs, & Meyer (1999); Schriefers, Meyer, & Levelt (1990). *Children:* Hanauer & Brooks (2005); Jerger, Lai, & Marchman (2002); Jerger, Martin, & Damian (2002); Seiger-Gardner & Schwartz (2008); Tazume (1997).

interpretation of an ambiguous target stimulus. Another approach to studying semantic access by to-be-ignored speech involves variations of the popular visual Stroop (1935) task. In *auditory Stroop-like analogs* with two opposing dimensions of a single speech stimulus, participants classify the gender of a speaker while attempting to ignore a spoken word itself. In *cross-modal Stroop-like analogs* with opposing speech and visual inputs, participants name the color of a visual stimulus while ignoring an irrelevant spoken word. As illustrated in Table 1, in both cases, the irrelevant spoken words are manipulated to create conflicting or congruent semantic relationships between the two dimensions or inputs.

The *irrelevant speech effect* has also been examined with cross-modal tasks involving opposing speech and visual inputs. Participants perform target visual recall or reading tasks and ignore irrelevant spoken words or prose whose content is varied from semantically related to unrelated. Finally—and of particular relevance because this approach is used in Experiment 1—in the *cross-modal picture word task* with target pictures and irrelevant speech, participants name pictures displayed on a monitor and ignore irrelevant auditory words. The words are manipulated to be semantically related or unrelated to the target picture.

Despite the wide variation in tasks, all of the above studies (conducted with adults) have found that manipulating the semantic content of the irrelevant speech affects target performance, particularly when the primary tasks also tap semantic processes. Hence, the results support the inference that the semantic content of speech is mandatorily accessed. This interpretation is compatible with theories proposing that speech input activates lexical–semantic information as it unfolds, according to the match between the evolving input and representations in memory (e.g., Marslen-Wilson, 1987). The behavioral results also agree with electrophysiological findings in adults indicating that semantic access by speech, as indexed by the N400 potential, occurs even when the semantic content is irrelevant (e.g., Kutas & Federmeier, 2011; Relander, Rama, & Kujala, 2008).

With regard to children, results across the existing studies indicate that manipulating the semantic content of irrelevant speech affects target performance in children of at least preschool age as it does in adults. Hence, the literature is consistent with the proposal that children, like adults, access the meaning of spoken words mandatorily (see Bjorklund, 2005, and Gleitman & Landau, 1994, for discussion). This viewpoint is also consistent with the observation that young children, like adults, process speech incrementally, with the unfolding input activating lexical–semantic representations (Swingley, Pinto, & Fernald, 1999). Developmental changes across childhood with regard to semantic access by speech have not been systematically investigated by previous studies.

Overall, a wealth of evidence in adults and some support in children indicates that spoken words mandatorily access their semantic codes even when listeners wish to ignore them. That said, the broader inference that semantic access by speech is independent of attentional resources could benefit from further evidence (see Jung, Ruthruff, Tybur, Gaspelin, & Miller, 2012, for discussion). For example, it is possible that the attentional demands of a target task may influence the processing of distractors. One contemporary approach to investigating the potential interplay between the influence of irrelevant stimuli on performance and the processing load of target tasks is formalized by Lavie's perceptual load model (Lavie, 1995, 2005).

Influence of Irrelevant Stimuli Versus Perceptual Load of Relevant Tasks

According to the perceptual load model (Lavie, 1995, 2005), the extent to which a concurrent irrelevant stimulus (called a *distractor*) is processed depends on the demands of the target task and of the distractor itself for capacity-limited attentional resources. The perceptual load (i.e., resource consumption) of the target task is critical because it determines the extent to which unused resources are available for other concurrent stimuli. In studies conducted within this framework, the perceptual load of the primary task is varied. High versus low perceptual loads are defined operationally by manipulating factors such as the intricacy of the target, the complexity of the display containing the target, or the difficulty of the response (Lavie & Tsal, 1994). Evidence across different types of tasks suggests that the specific type of manipulation used to increase the relative load is not relevant (Lavie & Torralbo, 2010).

According to the perceptual load model, a task with a low perceptual load allows parallel processing of the target and the distractor, whereas a task with a high perceptual load depletes a participant's pool of capacity-limited resources and thus eliminates or reduces processing of the distractor. Hence, if a distractor influences target processing on a task with a low perceptual load but not at all or less so on a task with a high perceptual load, then distractor processing is dependent on capacity-limited resources. Conversely, if distractors consistently influence performance despite variations in the perceptual load of the target task, then (to the extent tested by the manipulation) distractor processing is independent of capacity-limited resources.

This theoretical framework raises some questions about the broad inferences of the studies in Table 1, namely, that semantic access by speech in adults and children is mandatory and independent of attention. In most or all of the tasks, the perceptual load of the primary task is low. Consider, for instance, the traditional cross-modal picture word task in which individuals name pictures while attempting to ignore spatially separate and modality-distinct auditory distractors (Experiment 1 of this study). The target stimulus consists of a single, clearly defined, and easy-to-name picture displayed on a blank monitor. The pervasive distractor processing observed in this task may, according to perceptual load theory, result from the availability of unused capacity-limited resources. According to the "spill-over hypothesis" of the theory (Lavie & Torralbo, 2010, p. 1657), processing proceeds for all concurrent stimuli until capacity-limited resources are exhausted.

An important question, then, is what might happen if the perceptual load of processing the target in the picture word task is increased. This question can be addressed with the new *multimodal* picture word task (Experiment 2 of this study) that Jerger and colleagues developed (Jerger, Damian, Spence, Tye-Murray, & Abdi, 2009). As in the cross-modal version, participants name pictures and attempt to ignore spoken distractor words. However, in the multimodal picture word task, on a given trial, a video image of the head and chest of a talker is presented, with the picture displayed on the talker's T-shirt. The talker is presented either as a still image (auditory–static face) or shown uttering the distractor word (audiovisual–dynamic face). In the auditory mode, the distractor remains the same as in the cross-modal task—spatially separate and modality distinct—but the complexity of the display containing the target is increased. In the audiovisual mode, the target and distractor are less spatially separate and modality distinct, and the increased complexity of the display containing the target now includes visual speech. Thus, picture naming in the multimodal task assumes an increased perceptual load relative to the cross-modal task.

If spoken words access their semantic codes independently of capacity-limited resources, then the different perceptual loads of the cross-modal and multimodal tasks should be irrelevant. Manipulating the semantic content of the speech distractors should consistently influence picture naming despite variations in the perceptual load. Conversely, if semantic access by speech depends on capacity-limited resources, then increasing the perceptual load of the target task may reduce or eliminate the distractor's influence on naming. In addition to this central question, the perceptual load framework also seems to raise some potentially relevant considerations for multimodal speech perception in general and for face-to-face communication by children.

Multimodal Speech Perception and Face-to-Face Communication by Children

A large body of evidence indicates that adults integrate the auditory and visual aspects of speech without conscious awareness (e.g., McGurk & MacDonald, 1976) and that the addition of visual speech benefits understanding in a variety of listening situations containing noise or unfamiliar information (e.g., Arnold & Hill, 2001; A. MacLeod & Summerfield, 1987). Some infant studies are also consistent with the idea that speech perception is a multimodal event (e.g., Burnham & Dodd, 2004; Kuhl & Meltzoff, 1982; Rosenblum, Schmuckler, & Johnson, 1997; although see Desjardins & Werker, 2004, for some precautions). What is surprising, however, is that results in the child literature suggest that speech perception is less influenced by visual speech in children than in adults and perhaps in infants. In fact, in their original research, McGurk and MacDonald (1976) noted that significantly fewer children than adults showed an influence of visual speech on perception. For instance, in response to one type of McGurk stimulus (auditory /ba/ vs. visual /ga/), 90% of adults but only 40%–60% of children reported hearing the illusion /da/. This pattern of results has been replicated and extended to other tasks (Desjardins, Rogers, & Werker, 1997; Dupont, Aubin, & Menard, 2005; Hockley & Polka, 1994; Massaro, 1984; Massaro, Thompson, Barron, & Laren, 1986; Sekiyama & Burnham, 2004; Wrightman, Kistler, & Brungart, 2006). With regard to estimating the age at which visual speech begins to exert an adultlike influence in children, studies have investigated word recognition and the McGurk effect and in general indicate that this process occurs in the preteen or teenage years (Dodd, 1977; Hockley & Polka, 1994; Tremblay et al., 2007; although see Sekiyama & Burnham, 2004).

Interpreting the literature on multimodal speech perception by young children is challenging because we do not understand whether seeing the face of the talker per se may affect the results. Despite scant direct evidence, the pioneering program of research by Doherty-Sneddon and her colleagues suggests this possibility. Younger children relative to older children have more difficulties during face-to-face communication on some types of tasks, such as describing abstract shapes (Doherty-Sneddon, Bonner, & Bruce, 2001; Doherty-Sneddon et al., 2000). In fact, when young children were trained to look away from the talker during questioning, the accuracy of their answers improved (Phelps, Doherty-Sneddon, & Warnock, 2006). Doherty-Sneddon and colleagues (e.g., Doherty-Sneddon, Phelps, & Calderwood, 2009) have speculated that their observed "face-to-face interference effect" (Doherty-Sneddon et al., 2000, p. 595) in younger children is related to the general phenomenon of gaze aversion, in which individuals attempt to reduce the perceptual load of environmental input to enhance performance (Glenberg, Schroeder, & Robertson, 1998). This possibility suggests that at least some of the existing studies on audiovisual speech perception may have underestimated children's multimodal integration if the children averted their gaze from the talker's face.

Overall, the surprising pattern of results for audiovisual speech perception and face-to-face communication by children requires significantly more research to understand. Toward this goal, the newly developed multimodal picture word task offers an opportunity to study whether semantic access by speech is influenced by the static face of the talker or visual speech. Below we more fully introduce the multimodal picture word task and provide an overview of previously reported findings on the task for phonologically related distractors, the companion distractors of the results reported in Experiment 2 (Jerger et al., 2009).

Multimodal Picture Word Task and Results With Phonologically Related Distractors

Some procedural aspects of our new multimodal picture word approach are similar to the conventional cross-modal task. Children are instructed to name a pictured object and to ignore an irrelevant spoken word. Further, the set of target pictures is held constant, and the content of the speech is manipulated to represent different types of relationships between the pictures and the distractors. In the multimodal version, however, the picture is shown on the talker's T-shirt along with the head and chest of the talker, and picture-naming performance is assessed in the presence of auditory versus audiovisual distractors coupled with a face, either static or dynamic for the two modes, respectively.

On picture word tasks, the relationship between the distractor and the picture is typically manipulated along a semantic dimension with related versus unrelated pairs (e.g., picture of "dog" paired with "bear" or "flag," respectively) or a phonological dimension with congruent or conflicting consonant-onset versus neutral vowel-onset pairs (e.g., picture of "bus" paired with "butterfly," "dump truck," or "onion," respectively). The research reported in Jerger et al. (2009) focused on phonological results. For current purposes, we overview only the effects produced by the conflicting onsets. In brief, results showed that relative to the neutral onsets, conflicting onsets significantly slowed naming in all age groups from about 4 to 12 years of age for both the auditory and audiovisual modes. The degree of phonological interference was significantly greater in the 4- and 5-year-olds than in the older children, a pattern agreeing with the general observation of greater interference in young children (e.g., see Jerger et al., 1993, for an interference function ranging from 3 to 79 years, that is, from children to older adults). With regard to assessing the influence of visual speech, the 4- and 5-year-olds showed greater interference (for at least some conditions) from audiovisual than from auditory distractors, a pattern suggesting that visual speech influenced performance in these age groups. By contrast, phonological interference effects in the older children did not show any influence of visual speech, a pattern that agrees with the findings in the literature (reviewed above).

For the two experiments reported herein, we focused on effects of the semantic, rather than the phonological, distractor words to investigate whether semantic access by speech is dependent or independent of capacity-limited attentional resources in children of varying ages. Experiment 1 reports a conventional cross-modal picture word task, and Experiment 2 reports the new multimodal procedure. Below we predict possible results on the cross-modal and multimodal tasks on the basis of perceptual load theory (Lavie, 1995) and knowledge of age-related variations in relevant cognitive skills. Then we predict possible results on the multimodal task for the auditory versus audiovisual modes.

Predicted Results

Perceptual Load Theory

On the basis of existing findings (see Table 1), we predicted an effect of semantic relatedness on the cross-modal task in all age groups. As discussed above, however, the new multimodal picture word task (Experiment 2) differs from the conventional task in terms of a higher perceptual load for the target task. Thus, predictions derived from Lavie's (1995) model were as follows: If semantic access by to-be-ignored speech is independent of attentional resources, the perceptual load should be irrelevant. Effects of the semantically related distractors should be prominent in all age groups and independent of variations in their perceptual load. In contrast, if semantic access by speech in children requires attentional resources, effects of the semantically related distractors on performance should vary as a function of the perceptual load of the task. The influence of semantic relatedness should be reduced on the multimodal task compared with the cross-modal task.

Prior research concerning the effects of perceptual load on cross-modal tasks is scant. A few previous investigations in adults reported mixed results (Klemen, Buchel, & Rose, 2009; MacDonald & Lavie, 2011; Tellinghuisen & Nowak, 2003); previous research in children is not available. One child study with a unimodal visual task observed reduced interference under a higher perceptual load in 7- to 12-year-olds (Huang-Pollock, Carr, & Nigg, 2002).

Age-Related Variations in Cognitive Skills

Predictions may also be generated on the basis of some of the well-known developmental improvements in cognitive skills. Primary relevant abilities for our tasks include processing resources, selective attention, and semantic knowledge. With regard to processing resources, the absolute amount of capacity-limited resources is significantly reduced in younger children relative to older children and adults (Bjorklund, 2005; Lavie, 2005). This evidence suggests that any absolute perceptual load will represent a relatively greater load in younger children (Huang-Pollock et al., 2002). Thus, perceptual load differences between the cross-modal and multimodal picture word tasks were predicted to exert a more powerful effect on performance in younger children than in older children.

With regard to selective attention, the ability to inhibit irrelevant distractors and resist interference is significantly reduced in younger children (Bjorklund & Harnishfeger, 1990; Dempster, 1992, 1993; Tipper, Bourque, Anderson, & Brehaut, 1989). Thus, interference effects due to semantic relatedness were predicted to be stronger in the younger children. To the extent that the developmental changes in selective attention interact with influences of the perceptual load, predictions depend on whether semantic access by speech is resource dependent or independent. If the processing of the distractors is resource independent, an age-related decrease in the degree of interference with increasing age would be predicted to occur on both tasks. If the processing of the distractors is resource dependent, however, an increased interference effect in younger children would be predicted on the cross-modal task (low perceptual load) but not on the multimodal task (higher perceptual load). To the extent that our data for the phonological distractors of the multimodal task generalize to the semantic distractors, those findings predict a greater degree of interference in the younger children for that task.

With regard to semantic knowledge, current behavioral and computational theories emphasize the dynamic developmental nature of semantic memory, with both the content and organization of semantic knowledge increasing with age. The mental lexicon is proposed to evolve gradually as a result of experience-driven learning, progressing from fragile, underspecified, and sparsely structured items to rich, elaborated, and robust representations in an intricately organized system (e.g., Bjorklund, 1987; Bloom, 2000; McGregor, Friedman, Reilly, & Newman, 2002; Plunkett, 1997). The ease with which words are accessed depends on the quality and organization of the semantic representations, with the richer, more robust, and more highly organized representations having lower thresholds of activation and easier retrieval with lessened attentional demands (e.g., Bjorklund, 1987; Cowan, 1995). Hence, semantic access by speech may be more effortful and vulnerable to retrieval failure in younger children, and thus interference effects due to semantic relatedness were predicted to be less pronounced in the younger children. To the extent that the developmental changes in semantic memory interact with influences of capacity-limited resources and the perceptual load, predictions would vary depending on whether semantic access by speech is resource dependent or independent as detailed above.

Predicted Results on Multimodal Task: Auditory Versus Audiovisual Modes

In the multimodal picture word task, the talker's face (along with his chest) is shown as a still image (auditory–static face) or while uttering the distractor (audiovisual–dynamic face). If semantic access by speech is resource independent, the face variation should not matter, and effects of semantic relatedness should not differ between the auditory and audiovisual modes. If, conversely, semantic access by speech is resource dependent and one mode represents a greater perceptual load than the other mode, then the effects of semantic relatedness should be reduced for the mode with the greater perceptual load. Whether the perceptual load of the two modes differs is difficult to determine from the existing literature.

With regard to effects of selective attention, the presence of visual speech may render the audiovisual distractors more difficult to ignore. In this case, the effects of semantic relatedness may be increased for the audiovisual mode relative to the auditory mode. With regard to age, to the extent that previous results with the phonological distractors (Jerger et al., 2009) generalize to the semantic distractors, we predicted greater interference from the audiovisual than the auditory distractors for the 4-year-olds and the 5-year-olds but not for the older children. This latter prediction also agrees with expectations based on the multimodal speech perception literature reviewed above. Finally, previous results suggest that visual speech benefits understanding primarily in degraded or conceptually difficult listening situations (Arnold & Hill, 2001; A. MacLeod & Summerfield, 1987). These findings suggest that the effects of semantic relatedness may not differ between the two modes for our clearly articulated and easy-to-understand speech distractors.

In short, the current literature does not allow clear-cut predictions for many of the above issues. Thus, our research should yield new insights into whether semantic access by speech is dependent or independent of capacity-limited resources in children. People's environments are intrinsically multimodal, and individuals typically communicate not only with auditory speech but also with visual speech. Yet most research in children has been carried out exclusively with auditory-only stimuli. Contrasting performance on the cross-modal versus multimodal tasks should positively enhance understanding of semantic access by speech during children's everyday communication.

Experiment 1: Cross-Modal Picture Word Task

Method

Participants

Participants were 100 children, 53 girls and 47 boys, ranging in age from 3 years and 11 months to 14 years and 9 months. The racial distribution was 86% White, 6% Asian, 6% Black, and 2% multiracial, with 5% of Hispanic ethnicity. The children were assigned to five groups of 20 each according to chronological age (4-year-olds, 5-year-olds, 6- to 7-year-olds, 8- to 9-year-olds, and 10- to 14-year-olds). The average Hollingshead (1975) social strata score for the children was 1.5, a value consistent with a major business and professional socioeconomic status. The criteria for participation were (a) no diagnosed or suspected disabilities and (b) English as the native language. All participants passed standardized or laboratory measures screening for normalcy of hearing sensitivity, visual acuity (including corrected to normal), oral-motor function, and spoken-word recognition. All children also correctly identified all the semantically related pairs of items on a laboratory category knowledge test requiring them to find the items out of six pictured alternatives by category membership (which ones are food, animals, etc.). The upper panel of the table in Appendix A summarizes the chronological ages and some cognitive skills in the age groups. The cognitive measures, which quantified expressive and receptive vocabulary, articulatory proficiency, and visual perception, were also obtained in the children tested on the multimodal picture word procedure (Experiment 2).

Materials and Instrumentation

Screening and Cognitive Measures

Hearing sensitivity was assessed with a standard pure-tone audiometer. Visual acuity was screened with the Rader Near Point Vision Test (Rader, 1977). Oral motor function was screened with a questionnaire designed by an otolaryngologist and speech pathologist (Peltzer, 1997). The questionnaire contained items concerning eating, swallowing, and drooling. Socioeconomic status was estimated with the Hollingshead Four Factor Index (Hollingshead, 1975). The tests of the cognitive battery are specified in Appendix A.

Picture Word Task

Test materials. To-be-named colored pictures were scanned into a computer as 8-bit PICT files and edited to achieve objects of a similar size and complexity on a white background. The rationale for colored pictures was to increase attention and interest for the child participants (e.g., Andersen, Muller, & Hillyard, 2009; Malter, 1948; Wolfe & Horowitz, 2004). The speech distractors were recorded directly into the computer by a male college student with clearly intelligible, normal speech as judged by a speech pathologist. The sampling rate was 22 kHz with 16-bit amplitude resolution. The output intensity levels of the stimuli were adjusted to equivalent peak intensities.

The to-be-named pictures were coupled to speech distractors whose content was manipulated to represent semantic or phonological relations or no relation between the pictures and the words. The development of the pictured objects and distractors has been detailed previously (Jerger, Martin, & Damian, 2002). Because this article is focused on the semantic items, we do not detail the phonological items. The semantic items for the results reported herein consisted of seven pictures and 14 speech distractors that were coupled to the pictures to represent semantically related versus unrelated picture word pairs (see Appendix B). Examples of semantically related and unrelated picture–distractor pairs are, respectively, "dog–bear" and "dog–cheese." The semantically related and unrelated distractors have comparable linguistic statistics (see Jerger, Martin, & Damian, 2002).

An additional variable that is routinely manipulated in picture word tasks (e.g., Glaser & Düngelhoff, 1984) concerns the temporal relationship between the onset of the picture and the distractor (i.e., the stimulus-onset asynchrony, or SOA). The picture and the distractor can be timed such that the onset of the distractor occurs before, simultaneous with, or after the onset of the picture. Varying the timing of the distractor relative to the picture manipulates which one of the successive cognitive stages underlying the target response is coactivated with the distractor. For our cross-modal task, performance was assessed at three SOAs. The onsets of the distractors were 150 ms before, 150 ms after, or simultaneous with the onsets of the pictures. For the results reported herein, data are reported only for the SOA with the onset of the distractors 150 ms before the onset of the pictures. The rationale for considering only this SOA is that the semantic interference effect is typically largest when the onset of the distractor occurs slightly before the onset of the picture; in contrast, little semantic interference is observed when the onset of the distractor is slightly after the onset of the picture (Damian & Martin, 1999; Schriefers, Meyer, & Levelt, 1990). The rationale for including multiple SOAs is that the phonological effects are largest at a different SOA, namely, when the onset of the distractor is slightly after the onset of the picture (Damian & Martin, 1999; Schriefers et al., 1990).

Instrumentation. The pictures were presented via a computer monitor. The speech distractors were presented via an audiometer and associated loudspeaker. Both the computer monitor and the loudspeaker were mounted on an adjustable-height table directly in front of the child at a distance of approximately 90 cm. When naming each picture, the child spoke into a directional microphone mounted on an adjustable stand. The microphone was placed approximately 12 in. from the participant's mouth without blocking his or her view of the monitor. To obtain naming times, the computer triggered a counter/timer with better than 1-ms resolution at the initiation of each picture. The timer was stopped by the onset of the child's vocal response into the microphone, which was fed through an amplifier and a 1-dB step attenuator to a voice-operated relay (VOR). A pulse from the VOR stopped the timer, which displayed the time in fractional seconds. We verified that the VOR was not triggered by the distractors.

Procedure

Participants were tested within a sound-treated booth. The children sat at a child-sized table with a cotester sitting alongside, keeping them on task. The tester sat at a computer workstation. The session began with the children naming each picture displayed on a 5″×5″ card. The cotester taught the child the target names of any pictures that were named incorrectly. Next the children watched the cotester model speeded naming as she flashed the picture cards quickly. The children then copied the cotester. Speeded naming practice trials alternated between the cotester and a child until the child was naming the pictures fluently.

For test trials, the children were instructed to name each picture as quickly and as accurately as possible and to ignore the distractors. All trials that were incorrect (i.e., the picture was misnamed) or flawed (e.g., the child's attention lapsed; he or she had squirmed out of position; or the VOR was triggered with a nonspeech sound, dysfluency, or an article, such as "*a* dog") were deleted online and re-administered after intervening items. Each child's speaking level, the position of the microphone or child, and/or the setting on the attenuator between the microphone and VOR were adjusted slightly to ensure that the VOR was triggering reliably. The intensity level of the distractors was approximately 70 dB SPL as measured at the imagined center of the child's head with a sound level meter.

The test items (picture–distractor pairs) were administered randomly within one "unblocked" condition also containing the three randomly intermixed SOAs. No individual picture or word distractor was allowed to recur without at least two intervening items. Again, the data below are based only on the SOA with the onset of the distractors 150 ms before the onset of the pictures.

Results and Discussion

Characteristics of the Data

Naming times larger than 3 *SD*s from an item's conditional mean were discarded. The average number of replacement trials for incorrect or flawed trials ranged from 1.23 in the 4-year-olds to 0.00 in the 10- to 14-year-olds. The number of missing trials remaining at the end because the replacement trial was also incorrect or flawed averaged 0.08 in the 4- to 5-year-olds, 1.33 in the 6- to 7-year-olds, 0.98 in the 8- to 9-year-olds, and 0.00 in the 10- to 14-year-olds.

Absolute and Adjusted Naming Times

Table 2 summarizes average picture-naming times for the unrelated and related distractor conditions in all age groups. Both the literature and the perceptual load theory predict interference from the semantically related distractors in all age groups; the age-related variations in cognitive skills additionally predict that the degree of interference should be greater in the younger children. We conducted a factorial mixed-design analysis of variance (ANOVA) with one between-participants factor (Age with groups: 4, 5, 6–7, 8–9, and 10–14 years) and one within-participants factor (Type of Distractor with conditions: unrelated vs. related). Results indicated that naming times varied significantly with age,

Table 2. Cross-modal picture word task.

	Type of distractor	
Age group in years	**Unrelated**	**Related**
4	1,861 (76)	2,042 (84)
5	1,640 (76)	1,808 (86)
6–7	1,263 (51)	1,411 (59)
8–9	1,155 (35)	1,246 (42)
10–14	858 (24)	935 (25)

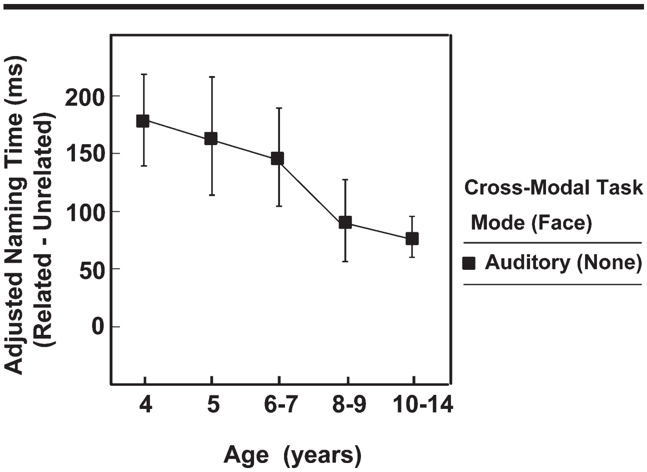
Note. Mean (standard error in parentheses) picture-naming times (in ms) in each age group for the semantically unrelated and related distractors. Data were obtained with the onset of the distractors 150 ms before the onset of the pictures. The mode used was auditory (no face).

$F(4, 95) = 54.13$, $MSE = 128{,}140.75$, $p < .0001$, partial $\eta^2 = .695$, and the type of distractor, $F(1, 95) = 58.55$, $MSE = 15{,}903.40$, $p < .0001$, partial $\eta^2 = .381$. The omnibus analysis did not yield other significant effects.

To quantify the effects of semantic relatedness, we derived adjusted naming times (related minus unrelated) as shown in Figure 1. The zero baseline of the ordinate represents absolute naming times for the unrelated distractors (Table 2). Planned orthogonal contrasts (Abdi & Williams, 2010) in each age group evaluated whether adjusted naming times differed significantly from zero. Results indicated significant semantic interference in all age groups: 4-year-olds, $F_{contrast}(1, 95) = 20.60$, $MSE = 15{,}903.40$, $p < .0001$, partial $\eta^2 = .178$; 5-year-olds, $F_{contrast}(1, 95) = 17.75$, $MSE = 15{,}903.40$, $p < .0001$, partial $\eta^2 = .157$; 6- to 7-year-olds, $F_{contrast}(1, 95) = 27.546$, $MSE = 15{,}903.40$, $p < .0001$, partial $\eta^2 = .225$; 8- to 9-year-olds, $F_{contrast}(1, 95) = 10.41$, $MSE = 15{,}903.40$, $p = .002$, partial $\eta^2 = .099$; 10- to 14-year-olds, $F_{contrast}(1, 95) = 7.46$, $MSE = 15{,}903.40$, $p = .008$, partial $\eta^2 = .073$. To scrutinize the adjusted naming times for any age-related change in the degree of semantic interference, we conducted trend analysis. Results indicated a significant linear decrease in the degree of semantic interference with increasing age, $F(1, 95) = 5.02$, $MSE = 31{,}806.80$, $p = .027$, partial $\eta^2 = .050$.

In summary, findings indicated that absolute naming times decreased significantly as age increased and that the semantically related distractors significantly slowed naming relative to the unrelated distractors. Both patterns agree with results in the literature (Brooks & MacWhinney, 2000; Jerger et al., 2009; Jescheniak, Hahne, Hoffmann, & Wagner, 2006). The planned orthogonal contrasts revealed that the effects of semantic relatedness produced a significant interference effect in all age groups. The significant linear trend characterizing the adjusted naming times (Figure 1) indicated an age-related decrease in the degree of interference, from about 180 ms in the 4-year-olds to 75 ms in the 10- to 14-year-olds. Overall, results were in line with previous results and consistent with our predictions.

Figure 1. Cross-modal picture word task: average adjusted naming times (semantically related minus unrelated) in each age group. Speech distractors were presented auditorily only (no face). Error bars represent the standard errors of the means.



Experiment 2: Multimodal Picture Word Task

Method

Participants

The participants were 100 children, 50 girls and 50 boys, ranging in age from 4 years and 3 months to 14 years and 0 months. The racial distribution was 85% White, 7% Asian, 3% Black, 2% American Indian, and 3% multiracial, with 13% of Hispanic ethnicity. The children were formed into five groups of 20 each according to age, namely, 4-year-olds, 5-year-olds, 6- to 7-year-olds, 8- to 9-year-olds, and 10- to 14-year-olds. The average Hollingshead (1975) social strata score for the children was again 1.5. The criteria for participation were the same as in Experiment 1. All participants again passed all of the screening measures for normalcy and correctly identified all of the semantically related item pairs on our laboratory category knowledge test. The bottom panel of the Appendix A table summarizes average chronological ages and cognitive skills in the multimodal age groups. As noted in Appendix A, participants in the multimodal experiment also completed measures of visual speechreading skills.

To determine whether the cognitive measures differed between the multimodal versus cross-modal groups, we conducted a factorial mixed-design ANOVA with two between-participants factors (Grand Groups: multimodal and cross-modal; and Age Groups: 4, 5, 6–7, 8–9, and 10–14 years) and one within-participants factor (standard scores for the cognitive measures). This analysis found no significant differences between the multimodal and cross-modal groups on the cognitive measures.

Materials and Instrumentation

Test materials. The semantic test items were the seven pictures and 14 speech distractors described above for the cross-modal task. Again, the phonological items are not considered. The audiovisual recordings were digitized via a Macintosh G4 computer with Final Cut Pro and Quicktime software. Color video was digitized at 30 frames per second with 24-bit resolution at 720 × 480 pixel size. Auditory input was digitized at a 22-kHz sampling rate with 16-bit amplitude resolution. The output intensity levels of the stimuli were adjusted to equivalent peak intensities. The talker was an 11-year-old boy with clearly intelligible, normal speech without pubertal characteristics as judged by a speech pathologist. As documented by the developmental literature on face perception with child faces as stimuli (e.g., Pellicano, Rhodes, & Peters, 2006), the child talker seemed to increase attention and interest for the child participants, an important consideration given that the multimodal task requires the children to complete the entire picture word task twice (once auditorily and once audiovisually). The talker looked directly into the camera, starting and ending each utterance with a neutral face and closed-mouth position. His full facial image and upper chest were recorded.

Each picture was displayed on the talker's T-shirt (upper chest at level of shoulder). The dimensions of the pictures on the computer monitor were about 65 mm in height and 85 mm in width; the width was intentionally comparable to the width of the talker's face at eye level, 80 mm. The picture and word distractors were paired in two presentation modes: audiovisual (dynamic face) and auditory (static face). For the audiovisual mode, the child saw 1,000 ms of the talker's still, neutral face and upper chest, followed by an audiovisual utterance of one distractor and the presentation of one picture on the talker's T-shirt, followed by 1,000 ms of still neutral face and the picture. For the auditory mode, the child saw and heard exactly the same event except that the video track was edited to contain only the talker's still neutral face for the entire trial.

With regard to the SOA, the onsets of the distractors were 165 ms before or 165 ms after the onsets of the pictures (five frames before or after the picture). Technically, a picture can be pasted onto an audiovisual stimulus only at the beginning of a frame (every 33 ms). Thus, the SOA was forced to be 165 ms, rather than the 150 ms of the cross-modal picture word task. To be consistent with the cross-modal task, we defined a distractor's onset on the basis of its auditory onset. For the results reported herein, data are reported only for the SOA with the onset of the distractors 165 ms before the onset of the pictures. Again, the rationale for considering only this SOA is that the semantic interference effect is largest when the onset of the distractors is slightly before the onset of the pictures (Damian & Martin, 1999; Schriefers et al., 1990).

Instrumentation. To administer the test items, we routed the video track of the Quicktime movie file to a high-resolution computer monitor and the auditory track through a speech audiometer to a loudspeaker. The remainder of the equipment was the same as for the cross-modal task with the exception that the counter/timer values were corrected by the amount of silence in each movie file before the onset of the picture.

Procedure

Participants were tested in two sessions, approximately 12 days apart. The mode of the speech distractors was held constant within a session and counterbalanced across sessions such that half of participants were tested with the auditory (static face) distractors first and half with the audiovisual (dynamic face) distractors first. Participants always began the initial session with the practice naming task and the second session with the minipractice task described above. Within each session, the test items (picture–distractor pairs) were administered randomly within one "unblocked" condition also containing the two randomly intermixed SOAs. No individual picture or word distractor was allowed to recur without at least two intervening items. The remainder of the procedure was the same as for the cross-modal task with the exception that the children were told that "Andy" was wearing a picture on his T-shirt and that he wanted to know what it was. Their job was to name the picture as quickly and accurately as possible.

Results and Discussion

Characteristics of the Data

Naming times that were more than 3 *SD*s from an item's conditional mean were discarded. Naming responses that were incorrect (i.e., the picture was misnamed) or flawed (e.g., the child's attention lapsed; he or she had squirmed out of position; or the VOR was triggered with a nonspeech sound, dysfluency, or an article, such as "*a* dog") were deleted online and readministered after intervening items. The average number of replacement trials ranged from 1.73 in the 4-year-olds to 0.86 in the 10- to 14-year-olds. The number of missing trials remaining at the end because the replacement trial was also incorrect or flawed averaged less than one for all groups, ranging from 0.43 in the 4-year-olds to 0.16 in the 10- to 14-year-olds.

Absolute and Adjusted Naming Times

Table 3 summarizes average naming times for the semantically unrelated and related distractors presented in the auditory (static face) or audiovisual (dynamic face) modes in the age groups. According to predictions from perceptual load theory, the semantically related distractors should interfere with picture naming in all age groups if semantic access by speech is resource independent but not if semantic access by speech is resource dependent. In the latter case, the younger children may show reduced interference relative to the older children. The age-related variations in capacity-limited attentional resources predict that any effects of perceptual load will influence performance more in the younger children. To the extent that developmental

Table 3. Multimodal picture word task.

Age group in years	Mode: Auditory (static face)		Mode: Audiovisual (dynamic face)	
	Unrelated	Related	Unrelated	Related
4	1,855 (85)	1,911 (98)	1,944 (91)	1,964 (95)
5	1,663 (85)	1,731 (80)	1,769 (86)	1,786 (84)
6–7	1,499 (62)	1,637 (77)	1,510 (79)	1,640 (75)
8–9	1,205 (82)	1,264 (81)	1,210 (75)	1,292 (71)
10–14	1,061 (62)	1,094 (56)	1,052 (58)	1,078 (56)

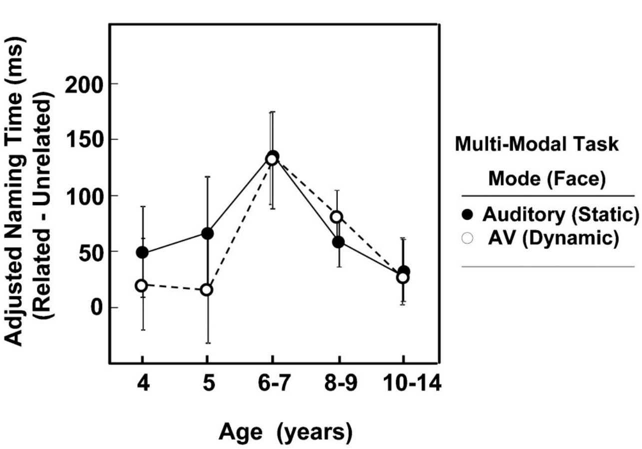
Note. Mean (standard error in parentheses) absolute naming times (in ms) for the semantically unrelated and related distractors presented in the auditory (static face) and audiovisual (dynamic face) modes. Data were obtained with the onset of the distractors 165 ms before the onset of the pictures.

differences in the other cognitive skills do not interact with the perceptual load, the degree of interference (a) should be increased in younger children with limitations in attending selectively to targets but (b) should be reduced in younger children with more impoverished and harder to access lexical semantic representations. To the extent that Jerger et al.'s (2009) results for the phonological distractors generalize to the semantic distractors, these results predict that the degree of interference should be relatively larger in the younger children particularly for the audiovisual mode. Finally, to the extent that visual speech benefits understanding primarily for degraded or conceptually difficult speech (e.g., Arnold & Hill, 2001; A. MacLeod & Summerfield, 1987), these results predict that the degree of interference should not differ as a function of mode for our high-quality, easy-to-understand auditory distractors.

We conducted a factorial mixed-design ANOVA with one between-participants factor (Age with groups: 4, 5, 6–7, 8–9, and 10–14 years) and two within-participants factors (Type of Distractor with conditions: unrelated vs. related; and Mode with levels: auditory vs. audiovisual). Results indicated a main effect of age, $F(4, 95) = 25.19$, MSE = 9,009,200.46, $p < .0001$, partial $\eta^2 = .515$; and type of distractor, $F(1, 95) = 23.75$, MSE = 305,491.45, $p < .0001$, partial $\eta^2 = .165$. The effect of the type of distractor also varied as a function of age, with a significant Age × Type of Distractor interaction, $F(4, 95) = 2.78$, MSE = 16,590.11, $p = .031$, partial $\eta^2 = .105$. The omnibus analysis did not yield a significant effect of mode nor of any other interactions.

As was the case in Experiment 1, to quantify the effects of semantic relatedness, we derived adjusted naming times (related minus unrelated) for the auditory (static face) and audiovisual (dynamic face) modes as shown in Figure 2. The zero baseline of the ordinate represents naming times for the

Figure 2. Multimodal picture word task: average adjusted naming times (semantically related minus unrelated) in each age group. Speech distractors were presented auditorily (with static face) and audiovisually (with dynamic face). Error bars represent the standard errors of the means. AV = audiovisual.



unrelated distractors (Table 3). Planned orthogonal contrasts (Abdi & Williams, 2010) evaluated whether the adjusted naming times differed significantly from zero. Results revealed significant interference (a) in the 6- to 7-year-olds for both the auditory and audiovisual modes: $F_{\text{contrast}}(1, 57) = 7.84$, MSE = 28,414.36, $p = .007$, partial $\eta^2 = .121$, and $F_{\text{contrast}}(1, 57) = 7.36$, MSE = 28,414.36, $p = .009$, partial $\eta^2 = .114$, respectively; and (b) in the 8- to 9-year-olds, again for both the auditory and audiovisual modes: $F_{\text{contrast}}(1, 57) = 4.08$, MSE = 11,329.81, $p = .048$, partial $\eta^2 = .067$, and $F_{\text{contrast}}(1, 57) = 6.34$, MSE = 11,329.81, $p = .015$, partial $\eta^2 = .100$, respectively. No other contrasts achieved significance.

In summary, as in Experiment 1, overall naming times decreased significantly as age increased, and semantically related distractors significantly slowed naming times relative to unrelated distractors. The effects of semantic relatedness, however, differed significantly across the age groups. Supporting this interaction, planned orthogonal contrasts showed significant interference in the 6- to 7-year-olds and 8- to 9-year-olds but not in the 4-year-olds, 5-year-olds, or 10- to 14-year-olds. As seen in Figure 2, adjusted naming times (collapsed across mode) were about 135 ms in the 6- to 7-year-olds and 70 ms in the 8- to 9-year-olds but only 35–45 ms in the other age groups. The observation of less semantic interference in the younger compared with older children is in contrast to the previous studies in the literature and also diverges from Jerger et al.'s (2009) results with the phonological distractors in the same children. Finally, neither the omnibus analysis nor the planned orthogonal contrasts indicated a significant effect of mode (auditory vs. audiovisual).

Although a lack of significant semantic interference in older children and adults for some conditions has been observed previously (e.g., Hanauer & Brooks, 2005), a failure to obtain semantic effects in the current group of 10- to 14-year-olds is clearly atypical. To investigate factors that might be underlying the unusual results in the 10- to 14-year-olds, we considered the cognitive measures (Appendix A) and the auditory unrelated naming times for individuals in this age group. Uncharacteristically, four children showed unusually slowed unrelated naming times for their ages: Unrelated times averaged 1,451 ms in these four children (range: 1,333–1,639), contrasting with 963 ms in the remaining children (range: 643–1,225 ms). It is possible that these atypically slowed unrelated naming times may have affected our ability to observe the effects of semantic relatedness. Hence, we recomputed the effect of semantic relatedness in the 16 children with more typical unrelated naming times for their ages. Results of a t test indicated significant semantic interference in the 16 children, $t(15)$ = 2.43, $p = .028$, $\eta^2 = .281$. The degree of interference (68 ms) in this subset of children was more comparable to that observed in the 10- to 14-year-olds on the cross-modal task (77 ms). These results clearly seem to provide some insights into the unexpected findings in the 10- to 14-year-olds. Given the differences in findings between Experiments 1 and 2, we compared results statistically across studies.

Post Hoc Comparison of Cross-Modal and Multimodal Tasks

We conducted a post hoc analysis comparing results of the cross-modal (auditory–no face) and multimodal (auditory–static face) tasks. The analysis focused sharply on the question of whether adjusted naming times differed on the two tasks in any age group (Figure 1 vs. Figure 2). Results were analyzed with multiple *t* tests with the problem of multiple comparisons controlled by the false discovery rate procedure (Benjamini & Hochberg, 1995; Benjamini, Krieger, & Yekutieli, 2006). Performance on the two tasks differed significantly only in the 4-year-olds. In this age group, the interference effect was about 180 ms for the cross-modal task but only 50 ms for the multimodal task.

General Summary and Discussion

The question addressed by this research was whether semantic access by speech in children of various ages is dependent or independent of capacity-limited attentional resources. These results represent the first research in children that assesses semantic access by irrelevant speech under varying perceptual loads, which allows a more sensitive evaluation of whether children access the meanings of spoken words in a manner independent of attention.

Results on our cross-modal and multimodal picture word tasks varied in complex ways as a function of age and the perceptual load of the target task. On the cross-modal task with its low perceptual load, results showed significant semantic interference in all age groups, with the magnitude of the interference significantly larger in younger children than in older children. These findings are in line with previous research for the cross-modal picture word and Stroop-like tasks (Hanauer & Brooks, 2003, 2005; Jerger, Elizondo, Dinh, Sanchez, & Chavira, 1994; Jerger, Martin, & Pirozzolo, 1988; Most, Sorber, & Cunningham, 2007; Seiger-Gardner & Schwartz, 2008; Tazume, 1997). Our results support models of cognitive development proposing that children become more efficient with increasing age in inhibiting irrelevant stimuli and resisting interference (Bjorklund & Harnishfeger, 1990; Dempster, 1992; Tipper et al., 1989).

Results on our multimodal task with the same semantic items but a higher perceptual load showed significant semantic interference only in the middle-age children (6- to 7-year-olds and 8- to 9-year-olds). By contrast, results in the younger children and in the 10- to 14-year-olds did not show significant semantic effects. Clearly the higher perceptual load dramatically altered the results. The lack of a greater semantic interference in the younger children contrasts with similar studies in the literature. It should be noted, however, that age-invariant interference effects have been observed previously by studies investigating the interactions between targets and distractors representing basic perceptual attributes. For example, the degree of interference on a Garner task (Garner, 1974), in which participants identify talker gender while attempting to ignore irrelevant variability in spatial location, does not vary with age (Jerger, Pearson, & Spence, 1999). With regard to the 10- to 14-year-olds, the lack of significant semantic interference is also atypical and may reflect the influence of a few children, as discussed below. Finally, results did not differ as a function of the mode of the distractors, suggesting that the addition of visual speech did not significantly influence results.

With regard to comparing results formally on the cross-modal versus multimodal tasks for the auditory mode, Figures 1 and 2 show that the 4-year-olds, 5-year-olds, and 10- to 14-year-olds showed numerically greater effects of semantic relatedness on the cross-modal task than on the multimodal task. Performance on the two tasks differed significantly, however, only in the 4-year-olds. Prior to discussing the implications of our results for semantic access by speech in children, we should address whether the performance differences between the cross-modal and multimodal tasks might reflect influences other than the perceptual load contrast.

One alternative possibility, for example, is that the multimodal task may have slowed picture-naming times in the 4- and 5-year-olds compared with the cross-modal version. Given the assumption that a given temporal relationship (SOA) between the picture and the distractor taps into a specific processing stage of target preparation, perhaps the current SOA (–165 ms) was no longer appropriate for the assessment of semantic effects in these age groups. Two pieces of evidence argue against this possibility. First, picture-naming times for the unrelated distractors (Tables 2 and 3) did not differ statistically on the cross-modal versus multimodal tasks (1,861 ms vs. 1,855 ms in the 4-year-olds and 1,640 ms vs. 1,663 ms in the 5-year-olds, respectively). Second, as outlined in the Method section, another SOA (+165 ms) was also included in the experimental protocol. We did not report results herein because semantic effects were not expected at this picture-distractor temporal relationship. However, this latter SOA allowed us to investigate whether semantic effects may have been "delayed" in the younger children on the multimodal, compared with the cross-modal, task. Results for the +165-ms SOA also did not show effects of semantic relatedness. These observations make it unlikely that differences in overall naming times between the two tasks created the observed differences.

Another alternative we should consider is whether the methodological differences on the cross-modal and multimodal tasks may have been influencing the results. As examples, the tasks used slightly different SOA increments (165 vs. 150 ms), different numbers of SOAs (two vs. three), and different talkers (preteen vs. college student). We can address the possible impact of these methodological differences with evidence from previous studies that investigated the effects of semantic relatedness in children (Hanauer & Brooks, 2003, 2005; Jerger et al., 1994; Most et al., 2007; Seiger-Gardner & Schwartz, 2008; Tazume, 1997). In these studies, there were larger methodological variations between tasks than in our study—for example, the SOAs varied from –500 ms to +500 ms; the number of SOAs within a study varied from two to five; the types of conditions allowed two or three SOAs to vary across trials (unblocked) or held one SOA constant across trials (blocked); and the utterances included both male and female talkers (our preteen talker had a fundamental frequency of 202 Hz, which was within the range characterizing female adults; Baken, 1987). Despite this broad range of methodological variations, all of these studies observed significant semantic interference, with the degree of interference significantly greater in younger children. Overall, our results support the conclusion that performance differences between the cross-modal and multimodal tasks can be attributed to contrasting perceptual loads that differentially affected the younger and older children. Thus, below we discuss the implications of our grand pattern of results from perceptual load theory (Lavie, 1995). Results are considered separately for the children of mid-ages, oldest ages, and youngest ages.

In the mid-age children (6–9 years), effects of semantic relatedness on the cross-modal versus multimodal tasks did not differ despite the variations in the perceptual loads of the tasks. As shown in Figures 1 and 2, a significant interference effect was observed on both tasks, approximately 140–150 ms in the 6- to 7-year-olds and 60–90 ms in the 8- to 9-year-olds. This pattern of results indicates that semantic access by speech in children of this age range occurs independently of capacity-limited attentional resources to the extent tested. Findings support behavioral and connectionist proposals that spoken words activate their lexical semantic representations (Marslen-Wilson, 1987; Swingley et al., 1999). Our results also inform our understanding of the time course of semantic knowledge development. Significant effects of semantic relatedness independent of variations in the perceptual load indicate that children's lexical–semantic representations are sufficiently rich and elaborated for our set of words to render semantic access by speech robust and independent of attentional resources to the extent tested by 6 years of age.

Results in the 10- to 14-year-olds, in contrast, showed significant effects of semantic relatedness only on the cross-modal task. Numerically, the effects of semantic relatedness in the 10- to 14-year-olds on both of our tasks agree with the range of values observed by previous investigators in older children and adults (e.g., Damian & Martin, 1999; Seiger-Gardner & Schwartz, 2008). The lack of significant semantic interference on the multimodal task in the current group of 10- to 14-year-olds seemed largely attributable to a few participants with particularly slowed unrelated response times for their ages. For example, findings in the other 10- to 14-year-olds with more age-appropriate unrelated naming times (n = 16 of 20) did show significant interference. It is at present unclear what caused these unusual individual differences in unrelated naming times.

Finally and most important, results in the 4- and 5-year-olds varied significantly across the tasks. Performance on the two tasks differed significantly only in the 4-year-olds on tests controlled for multiplicity. However, both the 4- and 5-year-olds showed significant semantic interference on the cross-modal task but not on the multimodal task. According to the perceptual load model, significant semantic interference only on the cross-modal task indicates that semantic access by speech is resource dependent. The lack of a semantic interference effect in the younger children on the multimodal task also produced a developmental function that showed less interference in younger children than in older children. This contrasts not only with results in the literature but also with results for the phonological distractors in the same children showing significant phonological interference in all age groups, with the degree of interference significantly decreasing with increasing age (see Figures 4 and 5 in Jerger et al., 2009).

The difference in the pattern of interference effects on the multimodal task for the phonological onset-conflicting versus semantic distractors seems provocative to consider. Significant interference from phonologically related but not from semantically related distractors in the same 4- and 5-year-olds suggests that the phonological processing of speech is more resource independent than the semantic processing of speech. This possibility is consistent with the "phonology-first" principle within the literature, indicating that infants master language-specific phonetic categories well before they use words meaningfully (e.g., Dromi, 1987; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984). The phonology-first principle is also supported by behavioral results and theories noting that young children can generate and produce the names of objects whose meanings they do not fully understand (Bloom, 2000; McGregor et al., 2002).

The difference in performance on the multimodal and cross-modal tasks also highlights two separate cognitive factors that may affect semantic access by irrelevant speech in young children and thus underlie developmental change on the cross-modal and multimodal picture word tasks. The first factor is selective attention, which is less well developed in younger than in older children and hence on tasks with a low perceptual load increases interference in young children. Here, performance reflects children's limitations in inhibiting distractors and resisting interference from distractors whose meaning was accessed as a result of the low load of the target task. The second factor is attentional capacity, which again is presumed to be less well developed in younger than in older children. On tasks with a high perceptual load, reduced capacity-limited attentional resources decrease interference and limit semantic access by irrelevant speech in young children. Here, performance is reflecting children's more severely limited pool of resources for distractors whose meaning was not accessed because of the high load of the target task. A possible influence for the current experiments also involves younger children's more fragile and less elaborated semantic representations that rendered semantic access by speech more effortless (e.g., Bjorklund, 1987; Bloom, 2000; McGregor et al., 2002; Plunkett, 1997). Our child theories need to take into account the perceptual loads of tasks in order to predict which of these developmental factors will predominate.

Overall, our results clearly endorse Dempster's (1992, 1993) theoretical framework, which proposes that interference is a multifaceted phenomenon. The effects of age may vary depending on interactions among the source of the interference, the nature of the target, and the perceptual load of the task. In particular, further studies seem warranted on the possible effects of a task's perceptual load on face-to-face communication and semantic access by speech in children. Such evidence seems critical to advancing understanding of how children develop in the real world with its plethora of loads.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grant DC-00421, awarded to the University of Texas at Dallas. We thank Alice O'Toole for her generous advice and assistance in recording the audiovisual stimuli. Thanks also go to the children and parents who participated and to the students who assisted—namely, Derek Hammons and Scott Hawkins (for computer programming), Sarah Joyce Bessonette, Carissa Dees, K. Meaghan Dougherty, Alycia Elkins, Irma Garza, Brittany Hernandez, Kelley Leach, Michelle McNeal, Anne Pham, Lori Pressley, Anastasia Villescas, and, in particular, Karen Banzon (for data collection, analysis, and/or presentation).

References

Abdi, H., & Williams, L. (2010). Contrast analysis. In N. Salkind (Ed.), *Encyclopedia of research design* (pp. 243–251). Thousand Oaks, CA: Sage.

Andersen, S., Muller, M., & Hillyard, S. (2009). Color-selective attention need not be mediated by spatial attention. *Journal of Vision, 9*(6), 1–7.

Arnold, P., & Hill, F. (2001). Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology, 92,* 339–355.

Ayres, J. (1978). *Southern California Figure Ground Visual Perception Test* (6th ed.). Los Angeles, CA: Western Psychological Services.

Baken, R. (1987). *Clinical measurement of speech and voice*. London, United Kingdom: Taylor & Francis.

Banbury, S., Macken, W., Tremblay, S., & Jones, D. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors, 43,* 12–29.

Beaman, C. (2004). The irrelevant sound phenomenon revisited: What role for working memory capacity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1106–1118.

Beery, K., & Beery, N. (2004). *The Beery VMI Developmental Test of Visual Perception* (5th ed.). Minneapolis, MN: NCS Pearson.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57,* 289–300.

Benjamini, Y., Krieger, A., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika, 93,* 491–507.

Bjorklund, D. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review, 7,* 93–130.

Bjorklund, D. (2005). *Children's thinking: Cognitive development and individual differences* (4th ed.). Belmont, CA: Wadsworth/Thomson Learning.

Bjorklund, D., & Harnishfeger, K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10,* 48–71.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bowers, J., Davis, C., Mattys, S., Damian, M., & Hanley, D. (2009). The activation of embedded words in spoken word identification is robust but constrained: Evidence from the picture-word interference paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1585–1597.

Brooks, P., & MacWhinney, B. (2000). Phonological priming in children's picture naming. *Journal of Child Language, 27,* 335–366.

Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test* (3rd ed.). Novato, CA: Academic Therapy Publications.

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology, 44,* 209–220.

Corteen, R., & Wood, B. (1972). Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology, 94,* 308–313.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York, NY: Oxford University Press.

Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics, 41,* 393–401.

Damian, M., & Martin, R. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 345–361.

Dempster, F. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 112,* 45–75.

Dempster, F. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. In M. Howe & R. Pasnak (Eds.), *Emerging themes in cognitive development: Vol. 1. Foundations* (pp. 3–27). New York, NY: Springer-Verlag.

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

Desjardins, R., & Werker, J. (2004). Is the integration of heard and seen speech mandatory for infants? *Developmental Psychobiology, 45,* 187–203.

Dodd, B. (1977). The role of vision in the perception of speech. *Perception, 6,* 31–40.

Doherty-Sneddon, G., Bonner, L., & Bruce, V. (2001). Cognitive demands of face monitoring: Evidence for visuospatial overload. *Memory & Cognition, 29,* 909–919.

Doherty-Sneddon, G., McAuley, S., Bruce, V., Langton, S., Blokland, A., & Anderson, A. (2000). Visual signals and children's communication: Negative effects on task outcome. *British Journal of Developmental Psychology, 18,* 595–608.

Doherty-Sneddon, G., Phelps, F., & Calderwood, L. (2009). Gaze aversion during children's transient knowledge and learning. *Cognition and Instruction, 27,* 225–238.

Dromi, E. (1987). *Early lexical development*. Cambridge, United Kingdom: Cambridge University Press.

Dunn, L., & Dunn, L. (1997). *The Peabody Picture Vocabulary Test—III*. Circle Pines, MN: AGS.

Dupont, S., Aubin, J., & Menard, L. (2005). A study of the McGurk effect in 4- and 5-year-old French Canadian children. *ZAS Papers in Linguistics, 40,* 1–17.

Elliott, E., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal color-word interference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 654–667.

Elliott, E., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal color-word interference effects. *Perception & Psychophysics, 60,* 761–767.

Garner, W. (1974). *Processing of information and structure*. New York, NY: Psychology Press.

Glaser, W., & Düngelhoff, F. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Gleitman, L., & Landau, B. (Eds.). (1994). *The acquisition of the lexicon*. Cambridge, MA: MIT Press.

Glenberg, A., Schroeder, J., & Robertson, D. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26,* 651–658.

Goldman, R., & Fristoe, M. (2000). *Goldman Fristoe Test of Articulation—Second Edition*. Circle Pines, MN: AGS.

Green, E., & Barber, P. (1981). An auditory Stroop effect with judgments of speaker gender. *Perception & Psychophysics, 30,* 459–466.

Green, E., & Barber, P. (1983). Interference effects in an auditory Stroop task: Congruence and correspondence. *Acta Psychologica, 53,* 183–194.

Gregg, M., & Purdy, K. (2007). Graded auditory Stroop effects generated by gender words. *Perceptual and Motor Skills, 105,* 549–555.

Hanauer, J., & Brooks, P. (2003). Developmental change in the cross-modal Stroop effect. *Perception & Psychophysics, 65,* 359–366.

Hanauer, J., & Brooks, P. (2005). Contributions of response set and semantic relatedness to cross-modal Stroop-like picture-word interference in children and adults. *Journal of Experimental Child Psychology, 90,* 21–47.

Hantsch, A., Jescheniak, J., & Schriefers, H. (2009). Distractor modality can turn semantic interference into semantic facilitation in the picture-word interference task: Implications for theories of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1443–1453.

Henkin, Y., Yaar-Soffer, Y., Gilat, S., & Muchnik, C. (2010). Auditory conflict processing: Behavioral and electrophysiologic manifestations of the Stroop effect. *Journal of the American Academy of Audiology, 21,* 474–486.

Hockley, N., & Polka, L. (1994). A developmental study of audiovisual speech perception using the McGurk paradigm. *The Journal of the Acoustical Society of America, 96,* 3309.

Hollingshead, A. (1975). *Four Factor Index of Social Status*. Unpublished manuscript, Yale University, New Haven, CT.

Huang-Pollock, C., Carr, T., & Nigg, J. (2002). Development of selective attention: Perceptual load influences early versus late attentional selection in children and adults. *Developmental Psychology, 38,* 363–375.

Jerger, S., Damian, M. F., Spence, M. J., Tye-Murray, N., & Abdi, H. (2009). Developmental shifts in children's sensitivity to visual speech: A new multimodal picture-word task. *Journal of Experimental Child Psychology, 102,* 40–59.

Jerger, S., Elizondo, R., Dinh, T., Sanchez, P., & Chavira, E. (1994). Linguistic influences on the auditory processing of speech by children with normal hearing or hearing impairment. *Ear and Hearing, 15,* 138–160.

Jerger, S., Lai, L., & Marchman, V. (2002). Picture naming by children with hearing loss: I. Effect of semantically-related auditory distractors. *Journal of the American Academy of Audiology, 13,* 463–477.

Jerger, S., Martin, R., & Damian, M. (2002). Semantic and phonological influences on picture naming by children and teenagers. *Journal of Memory and Language, 47,* 229–249.

Jerger, S., Martin, R., & Pirozzolo, F. (1988). A developmental study of the auditory Stroop effect. *Journal of Brain and Language, 35,* 86–104.

Jerger, S., Pearson, D., & Spence, M. (1999). Developmental course of auditory processing interactions: Garner interference and Simon interference. *Journal of Experimental Child Psychology, 74,* 44–67.

Jerger, S., Stout, G., Kent, M., Albritton, E., Loiselle, L., Blondeau, R., & Jorgenson, S. (1993). Auditory Stroop effects in children with hearing impairment. *Journal of Speech and Hearing Research, 36,* 1083–1096.

Jescheniak, J., Hahne, A., Hoffmann, S., & Wagner, V. (2006). Phonological activation of category coordinates during speech planning is observable in children but not in adults: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 373–386.

Jones, D., Miles, C., & Page, C. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal, or memory? *Applied Cognitive Psychology, 4,* 89–109.

Jung, K., Ruthruff, E., Tybur, J., Gaspelin, N., & Miller, G. (2012). Perception of facial attractiveness requires some attentional resources: Implications for the "automaticity" of psychological adaptations. *Evolution and Human Behavior, 33,* 241–250.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Klemen, J., Buchel, C., & Rose, M. (2009). Perceptual load interacts with stimulus processing across sensory modalities. *European Journal of Neuroscience, 29,* 2426–2434.

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

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992, January 31). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.

Kutas, M., & Federmeier, K. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621–647.

Lackner, J., & Garrett, M. (1972). Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition, 1,* 359–372.

La Heij, W., Dirkx, J., & Kramer, P. (1990). Categorical interference and associative priming in picture naming. *British Journal of Psychology, 81,* 511–525.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 451–468.

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences, 9,* 75–82.

Lavie, N., & Torralbo, A. (2010). Dilution: A theoretical burden or just load? A reply to Tsal and Benoni (2010). *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1657–1664.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56,* 183–197.

Levelt, W., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Lew, H., Chmiel, R., Jerger, J., Pomerantz, J., & Jerger, S. (1997). Electrophysiologic indices of Stroop and Garner interference reveal linguistic influences on auditory and visual processing. *Journal of the American Academy of Audiology, 8,* 104–118.
Lewis, J. (1970). Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology, 85,* 225–228.
MacDonald, J., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics, 73,* 1780–1789.
Mackay, D. (1973). Aspects of the theory of comprehension, memory and attention. *The Quarterly Journal of Experimental Psychology, 25,* 22–40.
MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology, 21,* 131–141.
MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.
Malter, M. (1948). Children's preferences for illustrative materials. *Journal of Educational Research, 41,* 378–385.
Marsh, J., Hughes, R., & Jones, D. (2008). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory and Language, 58,* 682–700.
Marsh, J., Hughes, R., & Jones, D. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition, 110,* 29–38.
Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition, 25,* 71–102.
Martin, R., Wogalter, M., & Forland, J. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language, 27,* 382–398.
Massaro, D. (1984). Children's perception of visual and auditory speech. *Child Development, 55,* 1777–1788.
Massaro, D., Thompson, L., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology, 41,* 93–113.
McGregor, K., Friedman, R., Reilly, R., & Newman, R. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45,* 332–346.
McGurk, H., & MacDonald, M. (1976, December 23). Hearing lips and seeing voices. *Nature, 264,* 746–748.
Most, S., Sorber, A., & Cunningham, J. (2007). Auditory Stroop reveals implicit gender associations in adults and children. *Journal of Experimental Social Psychology, 43,* 287–294.
Neely, C., & LeCompte, D. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition, 27,* 37–44.
Oswald, C., Tremblay, S., & Jones, D. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory, 8,* 345–350.
Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces? *Developmental Science, 9,* 270–277.
Peltzer, L. (1997). *Oral motor function questionnaire*. Unpublished manuscript, Baylor College of Medicine, Houston, TX.
Phelps, F., Doherty-Sneddon, G., & Warnock, H. (2006). Helping children think: Gaze aversion and teaching. *British Journal of Developmental Psychology, 24,* 577–588.
Plunkett, K. (1997). Theories of early language acquisition. *Trends in Cognitive Sciences, 1,* 146–153.
Rader, K. (1977). *Rader near point vision test*. Tulsa, OK: Modern Education Corp.
Relander, K., Rama, P., & Kujala, T. (2008). Word semantics is processed even without attentional effort. *Journal of Cognitive Neuroscience, 21,* 1511–1522.
Ricker, T., Cowan, N., & Morey, C. (2010). Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review, 17,* 516–521.
Rosenblum, L., Schmuckler, M., & Johnson, J. (1997). The McGurk effect in infants. *Perception & Psychophysics, 59,* 347–357.
Saults, J., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General, 136,* 663–684.
Schriefers, H., Meyer, A., & Levelt, W. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86–102.
Seiger-Gardner, L., & Schwartz, R. (2008). Lexical access in children with and without specific language impairment: A cross-modal picture-word interference study. *International Journal of Language & Communication Disorders, 43,* 528–551.
Sekiyama, K., & Burnham, D. (2004). Issues in the development of auditory-visual speech perception: Adults, infants, and children. *Interspeech 2004–ICSLP,* 1137–1140.
Smith, M., & Groen, M. (1974). Evidence for semantic analysis of unattended verbal items. *Journal of Experimental Psychology, 102,* 595–603.
Sorqvist, P., Marsh, J., & Jahncke, H. (2010). Hemispheric asymmetries in auditory distraction. *Brain and Cognition, 74,* 79–87.
Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.
Styles, E. (1997). *The psychology of attention*. Hove, United Kingdom: Psychology Press.
Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71,* 73–108.
Tazume, H. (1997). Information processing in Stroop-like picture naming tasks by preschool children. *Japanese Journal of Psychology, 68,* 272–280.
Tellinghuisen, D., & Nowak, E. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics, 65,* 817–828.
Tipper, S., Bourque, T., Anderson, S., & Brehaut, J. (1989). Mechanisms of attention: A developmental study. *Journal of Experimental Child Psychology, 48,* 353–378.
Tremblay, C., Champoux, R., Voss, P., Bacon, B., Lepore, F., & Theoret, H. (2007). Speech and non-speech audio-visual illusions: A developmental study. *PLoS One, 2*(8), e742.
Tye-Murray, N., & Geers, A. (2001). *Children's Audio-Visual Enhancement Test*. St. Louis, MO: Central Institute for the Deaf.
Wechsler, D. (1974). *Wechsler Intelligence Scale for Children—Revised*. New York, NY: The Psychological Corporation.
Werker, J., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 7,* 49–63.
Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory-visual integration. *The Journal of the Acoustical Society of America, 119,* 3940–3949.
Williams, K. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: AGS.
Wolfe, J., & Horowitz, T. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5,* 1–7.

Appendix A

Mean (and *SD*) Chronological Ages and Performance on a Set of Cognitive Measures in Each Age Group Participating in the Cross-Modal (*N* = 100) and Multimodal (*N* = 100) Experiments

	Cognitive measures					
Age group, in years	**Age (years;months)**	**Expressive vocabulary standard score**	**Receptive vocabulary standard score**	**Articulation proficiency no. of errors**	**Visual perception age equivalent (years;months)**	**Visual speech reading: % correct words**
			Cross-modal task			
4	4;6 (0;4)	110.00 (13.67)	113.18 (13.25)	6.05 (5.63)	6;3 (1;6)	
5	5;5 (0;3)	116.85 (15.32)	116.32 (12.39)	0.60 (1.19)	6;11 (1;11)	
6–7	7;0 (0;6)	113.50 (17.16)	116.87 (11.12)	0.50 (1.47)	8;11 (1;7)	
8–9	8;10 (0;7)	112.35 (14.00)	113.22 (9.32)	0.10 (0.45)	10;8 (2;6)	
10–14	11;10 (1;6)	118.00 (21.13)	120.05 (12.73)	0.15 (0.68)	13;9 (3;3)	
			Multimodal task			
4	4;7 (0;3)	113.90 (15.52)	115.30 (13.96)	6.75 (7.37)	5;8 (1;2)	3.75 (6.66)
5	5;6 (0;4)	117.50 (12.37)	118.30 (11.14)	2.35 (3.61)	7;3 (1;11)	3.50 (5.40)
6–7	6;9 (0;6)	114.80 (14.91)	114.42 (10.61)	1.15 (2.39)	8;10 (1;8)	12.75 (11.29)
8–9	8;11 (0;6)	115.40 (10.63)	113.42 (10.45)	0.25 (0.63)	11;11 (2;11)	14.00 (9.40)
10–14	11;11 (1;5)	115.40 (12.84)	120.35 (13.01)	0.00 (0.00)	14;0 (2;3)	22.00 (13.90)

Note. Each age group contained 20 children. Receptive vocabulary was quantified with the Peabody Picture Vocabulary Test—III (Dunn & Dunn, 1997) in all groups. Expressive vocabulary was quantified with the Expressive Vocabulary Test (Williams, 1997) in cross-modal groups and the Expressive One-Word Picture Vocabulary Test (Brownell, 2000) in multimodal groups. Articulatory proficiency was estimated with the Goldman Fristoe Test of Articulation—Second Edition (Goldman & Fristoe, 2000) in all groups. Visual perception was estimated with the Southern California Figure Ground Visual Perception Test (Ayres, 1978) in children 4 to 8 years and with the Block Design subtest of the Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974) in children 9 to 14 years in cross-modal groups and with the Beery VMI Developmental Test of Visual Perception (Beery & Beery, 2004) in multimodal groups. The difference between tests for the cross-modal versus multimodal groups reflects the availability of a new test in 2004 that spanned the entire age range. Speech reading ability (visual only) was quantified with the Children's Audiovisual Enhancement Test (Tye-Murray & Geers, 2001) in the multimodal groups.

Appendix B

Pictured Objects and Semantically Related and Unrelated Distractors

Semantic items		
	Distractors	
Pictured objects	**Related**	**Unrelated**
Boot	Slipper	Flag
Dog	Bear	Cheese
Doll	Puppet	Worm
Pants	Shirt	Horse
Pickle	Lemon	Glove
Pizza	Hotdog	Dress
Tiger	Cat	Bed

Note. The original set of semantic items for the cross-modal task contained one additional picture (''gun'') and its distractors (''knife'' and ''present''). This item was eliminated from the current data analyses, which did not alter the results. The rationale for the elimination was that, during pilot studies for the new multimodal task, some parents objected to the picture. Thus, its use was discontinued. Results reported herein for the cross-modal and multimodal picture word tasks are based on exactly the same items. Our original study represented the first children's version of the cross-modal picture word task with semantic and phonological items (Levelt, Roelofs, & Meyer, 1999). The original set of items also included filler and baseline items to be consistent with initial adult studies (see Jerger, Martin, & Damian, 2002). The filler and baseline items were subsequently eliminated because pilot studies for the new multimodal task indicated that the inclusion of these items was neither necessary nor efficient and did not influence performance.