Audio-Visual Information Processing Across Different Age Groups

Dissertation

zur Erlangung des akademischen Grades

eines Doktors der Philosophie (Dr. phil.)

eingereicht am Fachbereich Humanwissenschaften

der Universität Kassel

vorgelegt von

Dipl.-Psych. Claudia Wille

Januar 2016

Datum der Disputation: 10.06.2016

Erstgutachterin: Prof. Dr. Mirjam Ebersbach

Zweitgutachter: Prof. Dr. Tobias Richter

Drittgutachterin: Prof. Dr. Gudrun Schwarzer

TABLE OF CONTENTS

CHAPTER 1: Synopsis

CHAPTER 1

Synopsis

1 Introduction

People possess different sensory modalities to detect, interpret, and efficiently act upon various events in a complex and dynamic environment (Fetsch, DeAngelis, & Angelaki, 2013). In recent decades, much empirical work has been done to understand the interplay of modalities in sensory integration (i.e., the auditory, visual, haptic, gustatory and olfactory modalities, see Calvert, Spence, & Stein, 2004). On the one hand, integration of multimodal input as a functional principle of the brain enables the versatile and coherent perception of the environment (Lewkowicz & Ghazanfar, 2009). On the other hand, sensory integration does not necessarily mean that input from modalities is always weighted equally (Ernst, 2008). Rather, when two or more modalities are stimulated concurrently, one often finds one modality dominating over another. Study 1 and 2 of my dissertation address the developmental trajectory of sensory dominance. In both studies 6- and 7-year-olds, 8- to 10-year-olds, and adults were tested in order to examine sensory dominance across different age groups. In Study 3 sensory dominance was put into an applied context by examining verbal and visual overshadowing effects among 4- to 6-year-olds performing a face recognition task. In the following I will review research on a) sensory dominance in adults, and b) sensory dominance in children. I will conclude the review with a description of the verbal overshadowing effect.

2 Sensory dominance in adulthood

Research on multimodal interactions in adults suggests that the visual modality often dominates other modalities (for visual dominance over the tactile system see Botvinick & Cohen, 1999; Hartcher-O’Brien, Gallace, Krings, Koppen, & Spence, 2008; Pavani, Spence, & Driver, 2000; for visual dominance over the proprioceptive system see Farnè, Pavani, Meneghello, & Làdavas, 2000). As the dissertation focusses on interactions between the auditory and the visual system, I summarize the most prominent results of previous research with adults next.

2.1 Vision alters sound-source localization: The Ventriloquist effect

The Ventriloquist effect (Howard & Templeton, 1966) refers to the perception of a sound coming from a direction other than its actual source: Speech produced by a ventriloquist is perceived as coming from the moving mouth of a puppet instead of the motionless mouth of the speaker (Callan, Callan, & Ando, 2015). Similarly, in a movie, voices are perceived to emanate from the actors on the screen, instead of the loudspeaker located somewhere else in the room (Warren, Welch, & McCarthy, 1981). Thus, auditory attention can be captured by visual stimuli, away from the actual source of the sound (Alais & Burr, 2004). In experimental settings Ventriloquism can be observed, when participants are presented with auditory and visual stimuli that occur simultaneously but at different locations. If participants have to report the perceived location of the auditory stimulus it typically shifts toward the location of the visual stimulus (Witten & Knudsen, 2005), whereas the localization of the visual stimulus is usually unaffected (or affected to a smaller amount) by the auditory stimulus (Vroomen & de Gelder, 2004). There are several possibilities to measure audio-visual shifts in localization. Typically, an experiment consists of a bimodal condition in which audio-visual stimuli are presented simultaneously, and an unimodal condition in which auditory stimuli are presented in the absence of visual stimuli. People have to localize the sound by means of pointing or visually fixating on the apparent location of the sound in both, the bimodal and the unimodal condition. When localization responses of both conditions are compared with each other, in the bimodal condition one usually finds a shift in the direction of the visual stimulus (Vroomen & de Gelder, 2004).

One explanation for the Ventriloquist effect (Howard & Templeton, 1966) is that in spatial perception, visual information has a higher acuity than auditory information and therefore the localization of an auditory source is biased by the visual system (Callan et al., 2015). Spatial information from the auditory modality can be easily distorted by environmental factors such as reverberations and needs to be extracted from a modality organized according to sound frequency rather than topography (Callan et al., 2015). In turn, visual spatial information projects directly onto the retina (Witten & Knudsen, 2015) with receptors providing a high spatial resolution. Thus, "an object in the world gives rise to a distribution of neural activity in the visual system that is tightly correlated with the true location of the object relative to the eye" (Witten & Knudsen, 2015, p. 490). As a result, vision often dominates audition in spatial perception as it provides more precise information (Witten & Knudsen, 2005).

The Ventriloquist effect (Howard & Templeton, 1966) is a robust phenomenon. It could be observed when the visual stimulus was not seen consciously as in patients with hemineglect (Bertelson, Pavani, Ladavas, Vroomen, & de Gelder, 2000), when attention toward the visual stimulus was distracted (Bertelson, Vroomen, de Gelder, & Driver, 2000), and even when participants were explicitly trained to ignore the visual stimuli (Vroomen, Bertelson, & de Gelder, 1998). For example in the study of Vroomen et al. (1998), sounds were randomly presented from a) the center of a screen and accompanied by visual stimuli presented from either left, right, or the center of the screen, or b) sounds that were presented either left or right from the screen but without accompanying visual stimuli. In each trial, participants had to judge whether they were presented with an "alternating sound" (presented from left or right) or a "central sound" (presented at the center). The main result was that the presentation of central sounds with visual stimuli appearing either left or right, increased the number of "alternating" judgements (i.e., the sound source was not perceived as coming from the center). This is surprising as participants were explicitly told to ignore the visual stimuli and also received corrective feedback after each trial. Instructions and feedback could thus not eliminate the visual bias of sound-source localization (Vroomen et al., 1998).

Although there is strong evidence for a visual dominance in spatial perception, auditory stimuli can bias the localization of visual stimuli, too. In the study of Warren et al. (1981), participants were asked to report the position of a visual stimulus which was presented with or without a spatially disparate auditory stimulus. Interestingly, when the auditory stimulus was present, the perceived position of the visual stimulus shifted toward the auditory stimulus. When participants were asked to report the position of an auditory stimulus in the presence of a spatially disparate visual stimulus instead, the perceived position of the auditory stimulus shifted toward the visual stimulus. However, the visual bias of auditory localization was much more pronounced than the auditory bias of visual localization (Warren et al., 1981).

Similar findings were reported for the “Ventriloquism aftereffect” (Recanzone, 1998). That is, repeated exposure to spatially disparate audio-visual stimuli caused a visual bias of auditory localization which even persisted when the sound was presented alone (Kopčo, Lin, Shinn-Cunningham, & Groha, 2010; Recanzone, 1998). In the opposite condition, participants were repeatedly presented with spatially disparate bimodal stimuli followed by the presentation of visual stimuli alone. When asked to localize the visual stimuli, the perceived position shifted toward the position of the sound. The shift in visual localization toward the sound, however, was much smaller than the shift in auditory localization toward the visual stimulus (Lewald, 2000). In sum, studies on sensory biases in spatial perception suggest that visual stimuli alter auditory localization to a greater extent than auditory stimuli do in visual localization.

2.2 *Vision alters speech perception: The McGurk effect*

The McGurk effect (McGurk & MacDonald, 1976) occurs when an auditory syllable is paired with an incongruent visual syllable, leading to the perception of a third syllable different from both the auditory and visual syllables (Basu Mallick, Magnotti, & Beauchamp, 2015). For example a sound of "ga" is perceived as "bga" when paired with the lip movement "ba", i.e., the result is a combination of incongruent visual and auditory syllables. Moreover a sound of "ba" is perceived as "da" when paired with the visual lip movement "ga", referred to as a fusion illusion with a third phoneme absent in either stimulus. Finally, the visual phoneme can also dominate the auditory one when the sound "ba" is presented with a visual "ga", resulting in the perception of "ga", known as a visual dominance illusion (Eskelund, McDonald, & Andersen, 2015). Thus, the McGurk effect demonstrates the relevance of visual information for speech perception (Basu Mallick, et al., 2015). Not limited to syllables, the effect can occur with isolated words and also with words in sentences (e.g., Sams, Manninen, Surakka, Helin, & Kätt, 1998), and thus might be relevant in everyday conversations, too.

However, the frequency with which individuals perceive the effect varies greatly (Strand, Cooperman, Rowe, & Simenstada, 2014; e.g., 0% to 100% of trials, Nath & Beauchamps, 2012; Basu Mallick et al., 2015). In part, this can be explained by individual differences in the ability to integrate sensory information, as people with higher abilities have been shown to be more susceptible to the effect than those with lower abilities (Grant & Seitz, 1998; but see Massaro & Cohen, 2000). Moreover, cultural background is a moderator. For example the effect was weaker for Japanese and Chinese listeners as compared to American listeners (Sekiyama, 1997; Sekiyama & Tohkura, 1993; but see Magnotti, Basu Mallick, Feng, Zhou, Zhou, & Beauchamp, 2015). One explanation for this difference is that Japanese make less use of visual information as compared to Americans due to the cultural practice of face avoidance.

Thus, vision might not alter speech perception to the same extent like it does in Americans (Sekiyama, 1997, Sekiyama & Tohkura, 1993). Another explanation refers to tonal properties. In the Japanese and Chinese language one and the same syllable can have different meanings that are distinguished by tone. Thus, Japanese and Chinese listeners might be more competent to identify incongruent audio-visual syllables than English listeners, and are therefore less prone to the effect (Sekiyama, 1997).

However, the McGurk effect (McGurk & MacDonald, 1976) was seen to persist across a vast number of experimental manipulations such as different gender of face and voice (Green, Kuhl, Meltzoff, & Stevens, 1991), when the listener was also the speaker (Aruffo & Shore, 2012), and even when participants were explicitly told that the auditory and visual stimuli may not match (Summerfield & McGrath, 1984). Thus, the McGurk effect is often cited as evidence for the powerful influence of visual input on auditory speech perception (Strand et al., 2014).

2.3 Ventriloquist- and McGurk effect: Differential integration processes

The underlying mechanisms of Ventriloquism (Howard & Templeton, 1966) and the McGurk effect (McGurk & MacDonald, 1976) have been subject of controversial debate. Whereas Massaro (1987) suggested a general integration process accounting for all audio-visual interactions similarly (Colin et al., 2001), there is evidence that interactions of the Ventriloquist and McGurk type are mediated by different cognitive mechanisms in line with with their specific functionality: Localization versus identification (Calvert, Brammer, & Iversen, 1998; Colin et al., 2001). In the study of Colin et al. (2001), participants were presented with two types of tasks. In the identification task (measuring the McGurk effect), they had to repeat aloud the syllable they had just heard. In the localization task (measuring the Ventriloquist effect), participants had to judge whether simultaneously presented auditory and visual stimuli came from the same location or not by saying “same” or “different”. For both tasks the same consonant-vowel syllables were used, articulated by a person displayed on the screen located in front of the participants. In order to assess whether the two phenomena were differently affected by spatial separation, auditory stimuli were presented through one of nine hidden loudspeakers located from straight ahead to 80 degrees left and right. Moreover, the speaker’s face was presented either upright or upside-down in order to assess the role of face inversion. In addition to incongruent syllables presented to elicit the McGurk illusion, congruent syllables were presented too, in order to examine the influence of semantic congruency on the Ventriloquist effect (Colin et al., 2001).

The Ventriloquist effect was affected by the degree of spatial separation, but unaffected by upright versus inverted presentation of the face. Moreover it was not affected by congruency of the syllables. The McGurk effect, in turn, was unaffected by spatial separation but reduced by face inversion, supporting the role of speaker's face information in speech perception: The inversion probably distorted visual cues for lip reading and hence reduced the visual salience of the syllables. The differential effects of semantic congruency, face inversion, and spatial separation on the Ventriloquist- and McGurk effect suggest that these phenomena rely on distinct cognitive mechanisms (Colin et al. 2001; for similar results see Jones & Munhall, 1997). This in turn opposes Massaro's (1987) view that a general integration process accounts for all audio-visual interactions similarly (Colin et al., 2001).

An approach that might account for modular processing of the McGurk effect (Howard & Templeton, 1966) and Ventriloquism (Howard & Templeton, 1966) is the Two-Streams Hypothesis (Ungerleider & Mishkin, 1982) which has been recently discussed by Calvert et al. (1998) and Colin et al. (2001). The Two-Streams Hypothesis states that two distinct pathways or “streams” are involved in visual processing. The ventral stream (referred to as the "what pathway") in the temporal lobe of the brain serves object identification and recognition.

The dorsal stream ("where pathway") of the parietal lobe is involved in the spatial localization of objects (Ungerleider & Mishkin, 1982). In analogy to visual processing, there is also evidence of a distinct "what" and "where" processing in the auditory modality (Clarke, Bellmann, Meuli, Assal, & Steck, 2000). Thus, as for the McGurk effect (corresponding with speech identification rather than localization) the neuronal level of processing might correspond with the “what streams”, whereas processing of the Ventriloquist effect might correspond with the “where streams” (Calvert et al., 1998; Colin et al., 2001). In fact, the modular processing of the McGurk and Ventriloquist effect in different streams could be partially supported in a study of Macaluso, George, Dolan, Spence, and Driver (2004).

2.4 Alteration of temporal aspects of vision by sound

A vast number of studies on audio-visual interactions have demonstrated the modification of auditory perception by vision (e.g., Posner, Nissen & Klein, 1976; Spence et al., 2012). However, interactions in the opposite direction have been reported, too. In particular visual perception in the temporal domain (i.e., frequency, timing, and duration) can be altered by sound (Shams, Kamitani, & Shimojo, 2002). Below, relevant findings will be illustrated.

Frequency. In a study of Welch, Dutton Hurt, and Warren (1986), audition has been found to capture vision in a phenomenon referred to as “auditory driving”, i.e., the perceived rate of visual stimuli can be driven up or down depending on the rate of simultaneously auditory stimuli (Berger, Martelli, & Pelli, 2003). In the study of Welch et al. (1986) people were presented with repetitive auditory clicks and visual flashes. The rate of visual flashes was held constant, whereas the rate of auditory clicks was either increased or decreased. Interestingly, the frequency of auditory clicks induced corresponding changes in the perceived rate of visual flashes. For example, an increasing rate of auditory clicks lead to an increase in the perceived rate of visual flashes. Changes in the frequency of visual flashes, however, did not alter auditory perception (Welch et al., 1986).

A similar phenomenon is “sound-induced illusory flashing” (Recanzone, 2002). That is, a single visual flash is perceived as multiple flashes when being accompanied by multiple auditory beeps (Shams et al., 2002). In analogy, multiple visual flashes are perceived as one flash when being accompanied by a single beep (Andersen, Tiippana, & Sams, 2004). Thus, when there is a conflict in audio-visual information regarding a temporal presentation rate, audition dominates perception. However, the authors did not control for a reversed effect. In the study of Shams et al. (2002) and Andersen et al. (2004) participants were instructed to count the number of visual flashes, and auditory beeps served as task-irrelevant distractors. It is not clear whether the instruction to count the number of auditory beeps under visual distractors would have elicited the same effect (e.g., multiple auditory beeps may be perceived as one beep when a single visual distractor is presented simultaneously).

Timing. In the study of Scheier, Nijwahan, and Shimojo (1999), two spatially disparate lights were presented with a small temporal delay ranging from - 60 ms to + 60 ms. People were asked to judge which of the two lights appeared earlier. Compared to the non-sound condition, judgments were more accurate when a sound was preceding and following the visual stimuli (A-V-V-A time order), but worse when sounds were inserted between the visual stimuli (V-A-A-V time order). This suggests that the perceived timing of a visual stimulus can be altered by sound depending on the temporal relationship between audio-visual stimuli. (Scheier et al., 1999; Shams, Kamitani, & Shimojo, 2004). Further evidence comes from a study of Fendrich and Corballis (2001). In Condition 1, participants were asked to report the location of a moving marker on a clock-face when a visual flash was seen. The reported marker position was earlier, when auditory distractors preceded the flash and later when auditory distractors followed the flash. However, one might argue whether this reflects a spatial rather than a temporal bias induced by sound since participants were not asked to estimate the time but to localize a moving marker (Morein-Zamir, Soto-Faraco, & Kingstone, 2003). In a second condition, however, participants had to judge the position of the marker when an auditory click was heard. When the click was preceded or followed by a visual distractor, similar capture effects were revealed as in Condition 1. Noteworthy, however, the auditory capture of visual flashes (Condition 1) was much more pronounced than the visual capture of auditory clicks in Condition 2 (Fendrich & Corballis, 2001).

Duration. In the study of Walker and Scott (1981) an auditory and a visual stimulus were presented for 1000 ms or 1500 ms either separately (unimodal condition) or simultaneously (bimodal condition). For both conditions participants were asked to report the perceived duration of the stimuli. In the unimodal condition, the duration of an auditory stimulus presented for 1000 ms was perceived to be longer than a visual stimulus of 1000 ms length. Interestingly, the perceived duration of bimodal stimuli (audio-visual stimuli presented in equal length) differed significantly from the perceived duration of single visual stimuli, but not from the perceived duration of single auditory stimuli. Moreover, the perceived duration of a silent gap in otherwise continuous tone (Condition 1) was longer than the perceived duration of a gap in otherwise continuous light (Condition 2). A gap occurring in a bimodal stimulus, however, was perceived equally long as in Condition 1 (i.e., gap in otherwise continous tone). Thus, the perceived duration of a bimodal stimulus presentation and the perceived duration of silent gaps in bimodal stimuli based on the auditory system (Walker & Scott, 1981).

2.5 Origins of sensory dominance: Modality appropriateness or information reliability?

Many results on audio-visual interactions can be explained by the Modality Appropriateness Hypothesis by Welch and Warren (1986). It states that the modality that is most appropriate with respect to a given task dominates processing. Thus, in spatial tasks the visual modality dominates the auditory one, because it conveys spatial information most precisely. In temporal tasks, in turn, the auditory modality dominates over the visual one (Shams et al., 2004). However, the Modality Appropriateness Hypothesis (Welch & Warren, 1986) has several limitations. For example alteration and even dominance of speech perception by vision (as demonstrated in fusion and visual dominance illusions of the McGurk effect, cf. Eskelund et al., 2015), suggest that visual dominance is not restricted to spatial tasks. Moreover, auditory dominance does not solely occur in temporal tasks, but could be demonstrated in spatial tasks, too. For example, Alais and Burr (2004) showed a reversed Ventriloquist effect in terms of auditory dominance over the visual modality, depending on visual stimulus characteristics. Basically, the aim of this study was to examine the integration of conflicting audio-visual information in bimodal space perception. For this purpose, two succeeding visual blobs were presented on a computer screen. Participants were asked which of the two blobs appeared further to the left. Next, two succeeding auditory clicks were presented and subjects had to judge which of the two sound appeared further to the left. As expected, participants were more accurate in judging the location of visual blobs than auditory clicks. In the final part of the study, visual and auditory stimuli were combined. Participants were presented with two sets of bimodal stimuli, i.e., the reference set consisting of a blob and a click, followed by a second set of a blob and a click, the so-called test set. Participants were asked which set of bimodal stimuli appeared further to the left, by comparing the location of the reference set with the location of the test set. However, there were test sets that either conflicted or not. In a non-conflict test set, the blob and the click appeared on the same side. In a conflict test set the blob appeared on the left, while the click appeared on the right, or vice versa. Importantly, visual blobs of the test stimuli varied in width of 4º (small blobs), 32º (medium blobs), and 64º (large blobs) which was a crucial point for judgements. When blobs were small, participants relied on visual rather than auditory test stimuli to judge the location. In contrast, when blobs were large, participants were more likely to make their judgements based on auditory test stimuli (Alais & Burr, 2004; Nelson, 2004; Witten & Knudsen, 2005) because large, fuzzy blobs were perceived as being less reliable than small, well defined ones (Nelson, 2004). In other words: with increasing width of the visual stimuli, visual localization lost reliability. Thus, judgements primarily based on auditory stimuli (Witten & Knudsen, 2005).

Basically, this finding suggests that vision does not dominate audition in spatial processing because of any physiological advantage of the visual over the auditory modality as stated by the Modality Appropriateness Hypotheses (Welch & Warren, 1986). Rather, sensory dominance depends on stimulus characteristics and their reliability in a given task (Witten & Knudsen, 2005). However, as there was no manipulation of auditory stimuli in the study of Alais and Burr (2004), it remains unclear whether variations of auditory characteristics such as loudness might affect visual and auditory localization as well.

2.6 The Colavita visual dominance effect

In the preceding section, modality appropriateness (Welch & Warren, 1986) and stimulus characteristics have been discussed as potential explanations for sensory dominance. However, there is one visual dominance effect, the Colavita effect (Colavita, 1974), that cannot be explained by these approaches. In short, the effect stems from a task in which unimodal auditory, visual and bimodal stimuli need to be detected. When presented with bimodal stimuli people often miss the auditory component and only respond to the visual component (Colavita, 1974). The effect has been intensively studied and to date could not be reversed in terms of auditory dominance (Spence, Parise, & Chen, 2012). The Modality Appropriateness Hypotheses cannot account for this, as the Colavita effect does not rely on a spatial task (i.e., detecting a location) but on a modality detection task (i.e., detecting audio-visual contents). On the other hand, if the effect relied on stimulus characteristics, one should assume that it would be prone to at least some sort of stimulus manipulation. However, to date this was not the case (except for people with monocular blindness, cf. Moro & Steeves, 2012). As the Colavita effect is crucial to one study of my dissertation, it will now be discussed in further detail including explanatory approaches.

The Colavita effect rests upon the studies of Colavita (1974). In his pioneering study participants were randomly presented with unimodal auditory stimuli (i.e., a 400 Hz tone), unimodal visual stimuli (i.e., a visual angle), and bimodal stimuli consisting of both, the 400 Hz tone and the visual angle. Participants were asked to give speeded responses by pressing a particular button whenever they perceived a visual stimulus and another button for the perception of an auditory stimulus. The presentation of bimodal stimuli, however, was not mentioned in the instruction. After each response, participants had to report verbally whether they had pressed the correct button. As expected, participants were perfectly able to detect unimodal visual and auditory stimuli. Interestingly, however, for bimodal trials more visual than auditory responses were given. On average, in 49 out of 50 bimodal trials the visual response button was pressed, and, noteworthy, in 16 out of 49 visual responses participants were unaware that indeed a bimodal stimulus had been presented. In the remaining 33 bimodal trials participants responded in favor of the visual stimulus, but then verbally reported the perception of a bimodal stimulus (Colavita, 1974, Experiment 1). Moreover, even when participants were explicitly instructed on how to respond to bimodal trials (i.e., to press whichever response button is appropriate to the stimulus which was recognized first in bimodal stimuli, Colavita, 1974, Experiment 3) visual dominance persisted. This was also the case when the subjective intensity of auditory stimuli relative to that of visual stimuli was doubled (Colavita 1974, Experiment 2).

The findings of Colavita (1974) gave rise to extensive research on the boundaries of this phenomenon, and in most studies a robust visual dominance effect was revealed. Visual dominance persisted under variation of stimulus characteristics (stimulus complexity: Sinnett, Spence, & Soto-Faraco, 2007, Koppen, Alsius, & Spence, 2008; intensity: Koppen, Levitan, & Spence, 2009), and variation of response demands (i.e., pressing the auditory and visual response button simultaneously when being presented with bimodal stimuli versus a separate response button for bimodal stimuli in the study of Koppen & Spence, 2007a). It also persisted regardless of whether the stimuli were presented from the same spatial location or from different locations, although the Colavita effect was somewhat larger in the former case (Koppen & Spence, 2007b). For temporal disparities it was revealed that the Colavita effect vanished when people reliably perceived the auditory or visual stimulus as coming first, that is, the stimuli were no longer perceived as a unitary event. This suggests that there is a temporal window in which audio-visual stimuli are perceived as a single sensation. Within this window, however, visual dominance was seen to persist irrespective of temporal asynchronies, though the effect was largest when the visual stimulus was presented slightly before the onset of the auditory stimulus (Koppen & Spence, 2007c). Moreove, adults exhibited visual dominance under modulation of attention (cues: Koppen & Spence, 2007a; distraction of attention: Sinnett et al., 2007), and variation of bimodal stimulus frequencies (Koppen & Spence, 2007a; Koppen & Spence, 2007d). Although the effect sizes in more recent studies have typically been smaller than in the study of Colavita (1974; see Koppen and Spence, 2007a) and although the effect could be attenuated with some designs (e.g., attenuation of visual dominance by increasing the proportion of unimodal auditory stimuli: Sinnett et al. 2007), it could not be reversed in terms of auditory dominance over the visual modality (for a review see Spence et al., 2012).

2.7 Explanatory approaches of the Colavita effect

There has been a continuous debate of whether the Colavita effect (Colavita 1974) relies on exogenous or endogenous attentional processes. Whereas exogenous attention refers to stimulus-driven attention (i.e., properties of the stimulus itself can catch attention in a preconscious and involuntary manner), endogenous attention refers to the voluntary allocation of attentional resources by the person (Chica, Bartolomeo, & Lupiáñez, 2013).

According to Posner et al. (1976) visual dominance reflects a compensatory mechanism: as visual stimuli are less capable to alert the organism to their occurrence and thus, capture less attention than input from other modalities, people endogenously focus attention toward visual stimuli in order to compensate for their low alerting capabilities. This, in turn, requires cognitive resources and diminishes the attention to other sensory input such as auditory stimuli (Posner et al., 1976).

However, several studies have questioned this approach because the manipulation of endogenous attention could certainly change the extent to which vision dominates over audition but was not sufficient to elicit auditory dominance (Spence et al., 2012). For example in the study of Colavita (1974, Experiment 4), participants were explicitly instructed to press the auditory button in response to bimodal stimuli. Thus, endogenous attention was directed toward auditory stimuli. Given the premise that visual stimuli are less alerting than auditory stimuli and, additionally, endogenous attention is focused on auditory stimuli, than in bimodal trials auditory rather than visual dominance should occur. However, even though the visual dominance effect was reduced, it could be neither terminated nor reversed in terms of auditory dominance. This suggests that "residual biases" (Sinnett et al., 2007) - such as more alerting capabilities of visual compared to auditory stimuli - prevent audition to dominate over vision when endogenous attention is biased toward auditory stimuli.

Several studies have addressed this issue (Koppen & Spence, 2007a; Rodway, 2005; Turatto, Benso, Galfano, Gamberini, & Umiltà , 2002) in opposition to Posner et al. (1976). Basically the approach of Posner et al. (1976) relies on a study which investigated the influence of audio-visual cues on processing speed. For auditory targets following auditory cues, shorter reaction times were reported than for auditory targets following visual cues. Interestingly, there also was a trend of shorter reaction times to visual targets following auditory cues compared to visual targets following visual cues. Posner et al. (1976) took this as evidence that visual stimuli are less alerting than auditory stimuli. In turn, Turrato et al. (2002) reported that the presentation of a cue in a given modality triggered the allocation of attention to that modality. Hence, when a visual target was presented, its detection was faster when the preceding cue was visual, whereas the auditory target was detected faster when presented with a preceding auditory cue.

Moreover, in a study conducted by Koppen and Spence (2007a), visual cues were seen to be even more alerting than auditory cues: Participants were presented with 40 % unimodal visual trials, 40 % unimodal auditory trials, and 20 % bimodal trials. Each trial was preceded by a visual or an auditory cue. Both types of cues were equally often presented and equally often preceded each type of trial. Participants were instructed to press one button in response to auditory stimuli and another button in response to visual stimuli. Both buttons should be pressed when a bimodal stimulus was presented. The analysis of bimodal trials revealed that participants gave significantly more erroneous visual than erroneous auditory responses, suggesting a visual dominance effect. Moreover, the effect was significantly larger across bimodal trials preceded by a visual cue than across bimodal trials preceded by an auditory cue. Most interesting, however, auditory cues in bimodal trials could neither terminate nor reverse visual dominance in terms of auditory dominance. The authors thus concluded that visual dominance might be partially explained in terms of exogenous attention with visual stimuli to be more effective in capturing attention than auditory stimuli (Koppen & Spence, 2007a).

Moreover, decreasing sensitivity to auditory stimuli in the presence of visual stimuli has been reported in the study of Koppen et al. (2009). Intensities of auditory and visual stimuli were individually adjusted at a 75% detection threshold. Next, a Colavita task was conducted with unimodal visual stimuli presented on 25% of the trials, unimodal auditory stimuli on 25% of trials, bimodal stimuli on 25% of trials, and no stimuli on the remaining 25% trials. Participants were instructed to press a particular button for an auditory stimulus, another button for a visual stimulus, both response buttons for a bimodal stimulus, and to give no response on those trials in which no target was presented. As a manipulation check, the analysis of unimodal trials revealed that participants were equally sensitive to auditory and visual stimuli. However, in bimodal trials a robust visual dominance effect was revealed, suggesting that participants were less sensitive to the auditory stimulus when simultaneously being presented with a visual stimulus. This might be attributable to more alerting capabilities of the visual component in a bimodal stimulus that leave less attentional resources for the processing of the auditory component (Koppen et al., 2009).

In conclusion, endogenous attention can moderate the extent to which vision dominates over audition (i.e., magnitude of the visual dominance effect), whereas exogenous capture of attention by the visual stimulus seems to be more relevant regarding the origin of visual dominance.

3 Sensory dominance in infants

To examine whether sensory dominance is present in early childhood, Lewkowicz (1988a) conducted a series of experiments with 6-month-old infants. The procedure used throughout the experiments was composed of two phases: a familiarization phase followed by a test phase both consisting of 12 trials.

In each trial of the familiarization phase a bimodal, audio-visual stimulus was repeatedly presented, consisting of a pulsing tone (330 Hz) and a flashing checkerboard. The auditory were pulsed, visual stimuli were flashed at one of two rates, 2.0 Hz or 4.0 Hz. Easier said: one group of infants was presented with a bimodal stimulus consisting of a higher rate of visual flashes compared to auditory pulses, whereas the other group of infants was presented with a higher rate of visual flashes than auditory pulses (see figure 1 for the temporal distribution of the familiarization stimuli). Each trial was initiated by the infant looking at the checkerboard and terminated by the infant looking away from it for a minimum of 1 second. This procedure was continued until 12 trials were completed.

Once familiarized, infants were presented with test trials (consisting of bimodal stimuli, too) that differed from the familiarized bimodal stimulus in the following: changes in the rate of the visual components (visual test trials), changes in the rate of the auditory component (auditory test trials), or simultaneous changes in the rate of both, the auditory and the visual component (audio-visual test trials, see Figure 1). Trials were presented in the following sequence: two test trials, three re-familiarization trials (i.e., presentation of the original stimulus), two test trials, three re-familiarization trials, and again, two test trials.

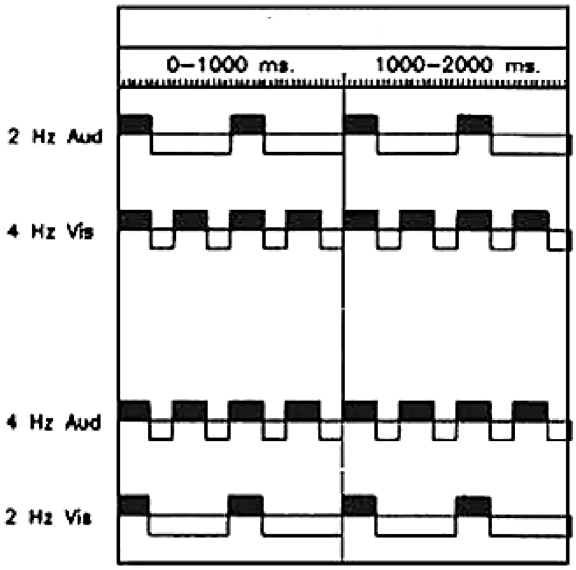

Figure 1. Schematic illustration of the temporal distribution of bimodal familiarization stimuli. Adapted from Lewkowicz, D. J. (1988a). Sensory dominance in infants: I. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155-171. Copyright 1988 by APA. Reprinted with permission.

Note. The top two lines illustrate the bimodal familiarization stimuli with a lower rate of auditory pulses (e.g., pulsed at 2 Hz, upper row) and a higher rate of visual flashes (flashed at 4 Hz, lower row). The bottom two lines show the bimodal familiarization stimuli with a higher rate of auditory pulses (pulsed at 4 Hz, upper row) and a lower rate of visual flashes (flashed at 2 Hz, lower row). Familiarization stimuli were repeatedly presented, i.e., presentation of the bimodal stimulus for 1 second and immediate repetition of the same bimodal stimulus without an inter-stimulus interval.

For the analysis each infant´s looking times in familiarization and re-familiarization trials were compared with the looking times in visual, auditory, and audio-visual test trials. If only the auditory stimulus had been attended during familiarization, then looking times should increase when auditory input changes (i.e., in auditory and audio-visual test trials). If only the visual stimulus had been attended, then looking times should increase when visual input changes (i.e., in visual and audio-visual test trials). Finally, if infants had attended both, the auditory and the visual stimuli during familiarization, then looking times should increase when there is a change in audio-visual input simultaneously (i.e., audio-visual test trials).

Indeed, the data supported the detection of temporal changes as reflected in looking times, but depending on the modality: it was revealed that infants of both familiarization groups detected changes in the rate of the auditory stimulus (auditory test trials), as well as simultaneous changes in the rate of both stimuli (audio-visual test trials), but did not recognize changes in the rate of the visual stimulus. Noteworthy, when visual stimuli were presented alone (i.e., in the absence of auditory stimuli) temporal changes were detected.

To further examine developmental changes in sensory dominance, the same experiments were conducted with 10-month-old infants (Lewkowicz, 1988b). In line with the preceding study infants were more consistent in detecting temporal changes in auditory than in visual stimuli, but, however, under some conditions also detected temporal changes in visual input (e.g., higher salience of visual stimuli relative to auditory stimuli, Lewkowicz, 1988b).

Similar findings were reported in a more recent study by Robinson and Sloutsky (2004). Infants (8-, 12-, and 16-month-olds) were familiarized with a bimodal stimulus, consisting of a visual and an auditory component. They then were presented with test trials in which the components were changed. If both components had been attended during familiarization, changes in both components should be detected. Thus, looking times in all test trials (auditory, visual and audio-visual test trials) should exceed looking times in familiarization trials. In Condition 1 the visual stimulus consisted of a single-shape pattern, in Condition 2 a more complex three shape pattern was presented, whereas auditory stimuli were the same in both conditions. In Condition 1 (single shape pattern) looking times increased in all test trials as compared to familiarization trials, but however, infants looked significantly longer in auditory and audio-visual test trials than in visual test trials. Moreover, in Condition 2 (complex three shape pattern) looking times only increased when there was a change in the auditory stimulus (auditory test trials) or a change in audio-visual stimuli simultaneously (audio-visual test trials). No increase in looking times was revealed for visual test trials, suggesting that changes in the visual stimulus were not detected. This is surprising, as infants were perfectly able to discriminate between familiarized and novel visual test stimuli when these stimuli were presented in isolation (i.e., without auditory stimuli). Thus, privileged auditory processing could not stem from any advantage in the discrimination between familiarized and novel auditory test stimuli.

Summarized, for infants an auditory dominance has been reported by Lewkowicz (1988a, 1988b), and Robinson and Sloutsky (2004, 2010). Moreover, the extent to which audition dominates over visual processing seems to depend on salience (Lewkowicz, 1988b) and the complexity of visual stimuli (Robinson & Sloustky, 2004).

However, the studies of Lewkowicz (1988a, 1988b), and Robinson and Sloutsky (2004, 2010) are only comparable to a limited extent as stimuli and task demands were quite dissimilar (Robinson & Sloutsky, 2004). Whereas infants had to detect temporal changes (i.e., changes in the rate of both, the auditory and the visual stimuli) in the studies of Lewkowicz (1988a, 1988b), infants had to detect changes in timbre and frequency of the auditory stimuli, and figural changes of the visual stimuli (i.e., changes in geometrical shapes) in the studies of Robinson and Sloutsky (2004, 2010). Thus, in order to arrive at legitimate conclusions, future studies should address a systematic investigation, i.e., differentiate between temporal and spatial tasks, using standardized procedures, and varying stimulus materials systematically.

4 Sensory dominance in 4-year-olds

To examine whether auditory dominance persists in older children, Sloutsky and Napolitano (2003) ran a series of recognition tasks with 4-year-olds. A calibration experiment was conducted in advance to make sure that the stimuli for each modality were sufficiently discriminable. Visual stimuli were pictures of unfamiliar landscapes and auditory stimuli consisted of computer generated three tone patterns. For the first experiment, stimuli were combined into several stimulus sets, each set consisting of a bimodal target stimulus, a bimodal distractor stimulus and two bimodal test stimuli. The recognition task consisted of 36 trials. In each trial, a different stimulus set was presented. Children were told that they would play a game in which they should find a prize, and were then presented with the first trial. It started with the presentation of a bimodal audio-visual target stimulus (AUD_1VIS_1), followed by a bimodal distractor stimulus (AUD_2VIS_2). Children were told that the target stimulus (AUD_1VIS_1) is the stimulus with the price. They then were presented with the target and distractor stimulus simultaneously, and asked to identify the stimulus with the price. During a short training children learned to consistently select the target stimulus (i.e., the price). Next, children were again presented with the target stimulus followed by two bimodal test stimuli. Test stimulus VIS_1AUD_{new} matched the target's visual component but contained a novel auditory component, whereas test stimulus $VIS_{new}AUD_1$ had a novel visual component but corresponded the target's auditory component. Children were asked which of the two test stimuli was the stimulus with the price (i.e., the target stimulus). In other words, they performed a forced choice-task in which they could rely either on the visual component of the bimodal target stimulus (VIS_1) or the auditory component (AUD_1). For the illustration of stimuli and procedure see figure 2. The analysis of the 36 trials (each consisting of a different stimulus set) revealed that 4-year-olds primarily selected the test stimulus $VIS_{new}AUD_1$ (i.e., novel visual component and old auditory component). This suggests that the auditory component of the bimodal target stimulus was more relevant than the visual component, and thus, children relied on the auditory stimulus when making same-different judgements. To test the robustness of these findings the same experiment was conducted with visual stimuli composed of simpler features (see figure 3). The results were largely the same, with children relying more on auditory stimuli (Sloutsky & Napolitano, 2003).

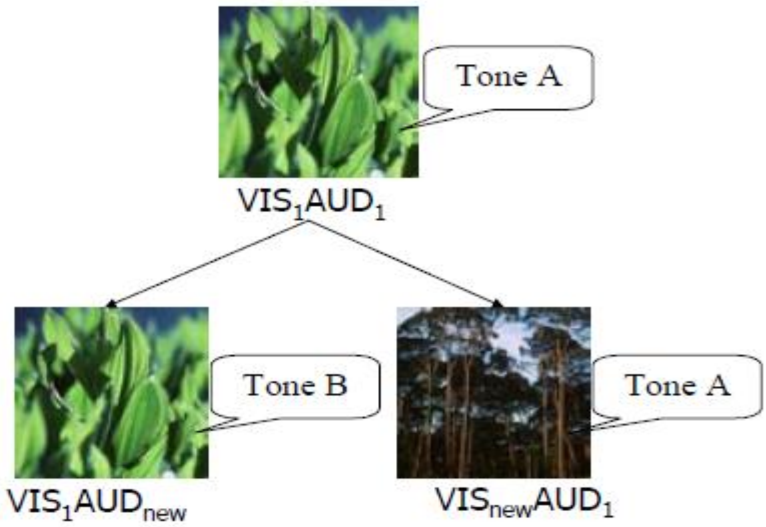

Figure 2. Example of the stimulus sets composed of complex features in the study of Sloutsky & Napolitano (2003). From Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833. Copyright 2003 by Wiley. Reprinted with permission.

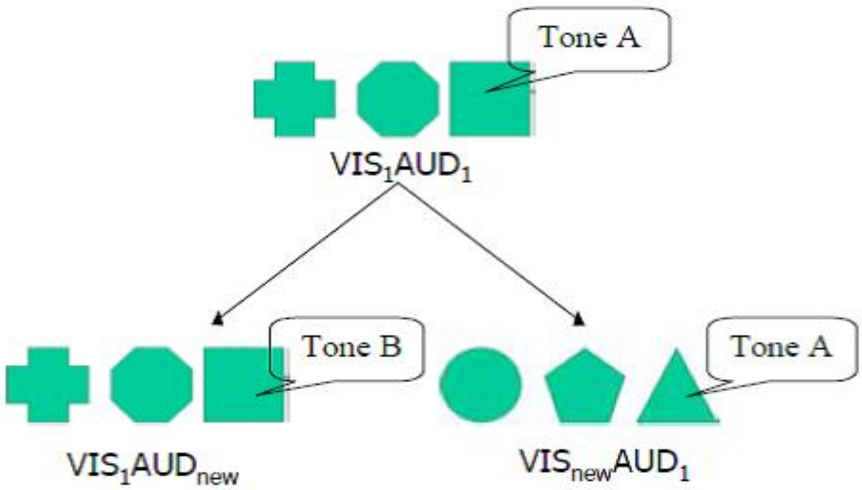

Figure 3. Example of the stimulus sets composed of simpler features in the study of Sloutsky & Napolitano (2003). From Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833. Copyright 2003 by Wiley. Reprinted with permission.

However, due to the experimental design it remained unclear whether children exhibited auditory preference or auditory dominance: It might be possible that children encoded both the auditory and the visual component of the target stimulus, but deliberately choosed the auditory modality to attend to (auditory preference). However, another possibility might be that the processing of the auditory component impaired simultaneous encoding of the visual component referred to as auditory dominance (Sloutsky & Napolitano, 2003).

To distinguish between these hypotheses a second experiment was conducted (see figure 4). The recognition task consisted of 24 trials and several stimulus sets (the stimuli, i.e., pictures of unfamiliar landscapes and three tone patterns were taken from the preceding experiment). In each trial, a stimulus set consisting of a target stimulus and four test stimuli was presented: First, the bimodal, audio-visual target stimulus VIS_1AUD_1 appeared on the screen. After presentation of the target stimulus four bimodal test stimuli were presented: the test stimulus VIS_1AUD_{new} corresponded with the target's visual component but contained a novel auditory component, whereas test stimulus $VIS_{new}AUD_1$ contained a novel visual component but corresponded with the target's auditory component. Test stimulus $VIS_{new}AUD_{new}$ contained a novel auditory component and a novel visual component, whereas test stimulus VIS_1AUD_1 was the same as the target stimulus. Children were asked to make same-different recognition judgments, i.e., whether a given test stimulus was the same as the target stimulus or different from the target stimulus. There were three assumptions:

First, if children encoded both the auditory and the visual component of the target stimulus, they should correctly accept VIS_1AUD_1, and correctly reject all other test stimuli. Second, if children failed to encode the visual component of the target stimulus, and thus made their judgements only based on the auditory component, they should erroneously accept $VIS_{new}AUD_1$, correctly accept $VIS_1AUD_{1,}$ and correctly reject all other test stimuli. Third, if they failed to encode the auditory component of the target stimulus, and made their judgements solely based on the visual component they should erroneously accept VIS_1AUD_{new}, correctly accept VIS_1AUD_1, while correctly rejecting all other test stimuli.

In all trials children were above chance in correctly accepting VIS_1AUD_1 and correctly rejecting $VIS_{new}AUD_{new}$. However, they exhibited below-chance accuracy in rejecting test stimuli with novel visual and old auditory components (VIS_{new}AUD1). That is, $VIS_{new}AUD_1$ was erroneously accepted in most of the trials, suggesting an impairment of visual encoding. On the other hand, there was an above-chance accuracy in rejecting test stimuli with novel auditory and old visual components (VIS_1AUD_{new}), suggesting there was no impairment of auditory encoding.

Thus, it seems that children made their judgements based on auditory stimuli (see assumption 2), and that the encoding of visual stimuli was impaired - although they were perfectly able to distinguish between visual stimuli in the calibration experiment (i.e., when visual stimuli were presented alone). Thus, referring back to the initial question whether children exhibit auditory preference or dominance, disrupted encoding of visual stimuli in the presence of auditory stimuli could be supported, and thus, auditory dominance (Sloutsky & Napolitano, 2003).

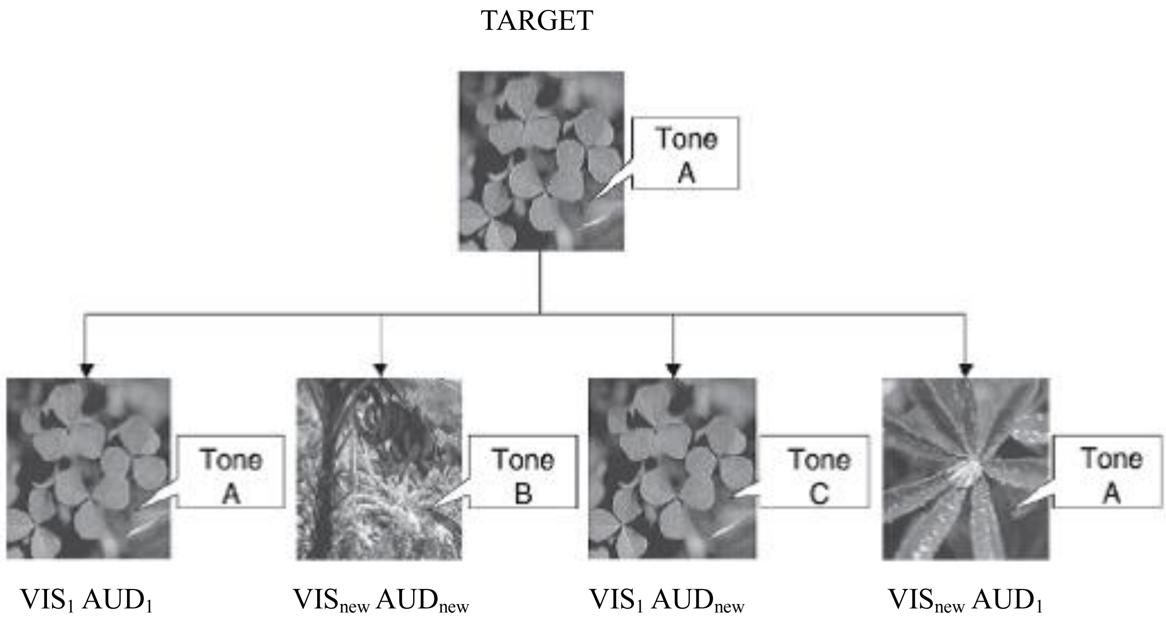

Figure 4. Example of the stimulus sets used in the study of Sloutsky and Napolitano (2003). From Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833. Copyright 2003 by Wiley. Reprinted with permission.

However, in the study of Napolitano and Sloutsky (2004) investigating the influence of stimulus familiarity on sensory dominance, 4-year-olds showed a relative modality flexibility. For example, children were presented with a recognition task which was the same as in the study of Sloutsky & Napolitano (2003, illustrated above in Figure 4), but contained different stimuli and conditions. In Condition 1 of this experiment auditory stimuli were more familiar than visual stimuli (for an example see figure 5), whereas visual stimuli were more familiar than auditory stimuli in Condition 2. It was revealed that when visual stimuli were more familiar than auditory stimuli, 4-year-olds exhibited visual dominance. In contrast, when auditory stimuli were more familiar than visual stimuli, young children showed auditory dominance.

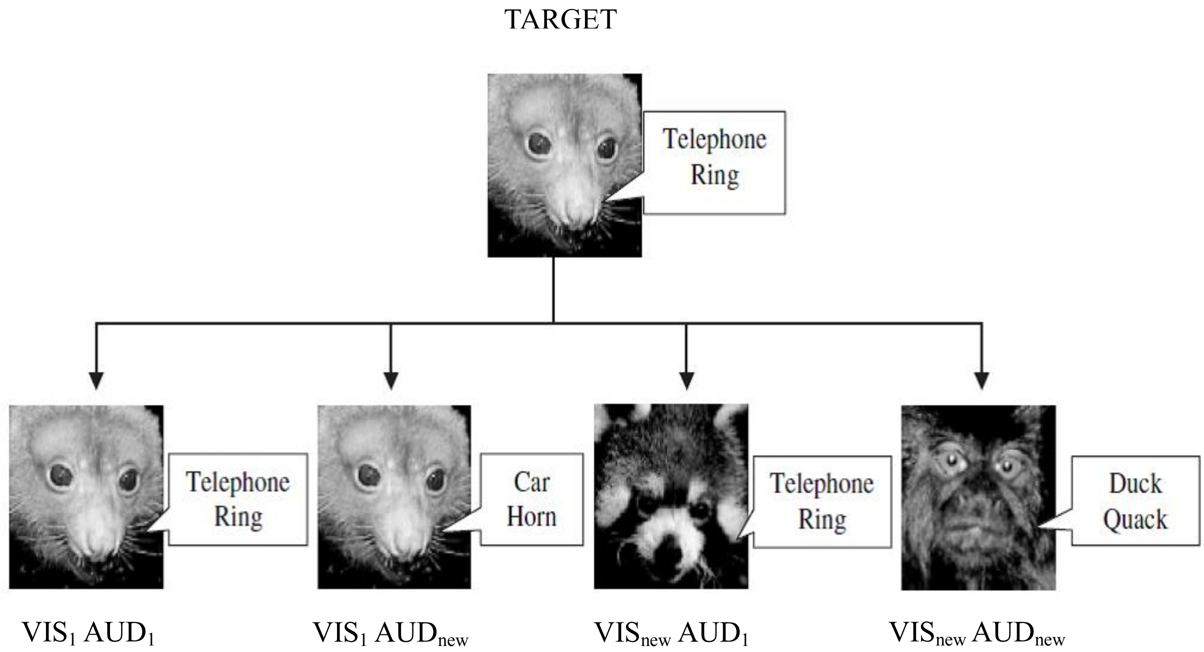

Figure 5. Example of the stimulus sets used in the study of Napolitano and Sloutsky (2004). In this particular stimulus set auditory stimuli were more familiar than visual stimuli. From Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*, 1850- 1870. Copyright 2004 by Wiley. Reprinted with permission.

Moreover, when children were presented with audio-visual stimuli that elicited auditory dominance in a preceding experiment, and were instructed to attend only to the visual component, auditory dominance was seen to persist. This, again, does not support modality preference in terms of “deliberate selective attention” to a certain modality (Napolitano & Sloutsky, 2004). Rather, children exhibited sensory dominance (auditory or visual dominance depending on stimulus familiarity) driven by automatic pulls on attention: when auditory stimuli were more familiar than visual stimuli, attention was captured by the auditory input which could not be prevented by the instruction to deliberately focus attention toward visual stimuli (i.e., endogenous attention, Napolitano & Sloutsky, 2004).

Interestingly, however, when stimuli of both modalities were unfamiliar, children exhibited auditory dominance, too. This might indicate that auditory dominance in children is a default, whereas visual dominance (depending on the familiarity of visual input) results from learning (Napolitano, 2006; Napolitano & Sloutsky, 2004). However, the question arises as to why auditory dominance should be a default. According to Napolitano and Sloutsky (2004) and Robinson and Sloutsky (2004) default auditory dominance might compensate for limited attentional resources in early childhood: Auditory input is more transient than visual input, that is, visual scences and objects usually persist longer. Considering the limited attentional ressources in young children, it might be adaptive to first allocate attention to dynamic (auditory) stimuli before processing more stable input (e.g., visual stimuli). Given this hypothesis, default auditory dominance may serve the attention to that type of input that is typically the most transient (Napolitano, 2006; Robinson & Sloutsky, 2004). During development, however, attentional resources increase and a broader range of visual objects and scenes become familiar. This might be an explanation, albeit speculative, as to why visual processing becomes more relevant in later childhood and adulthood.

5 Sensory dominance in 6- to 10-year-olds

Research on audio-visual interactions has rarely been conducted with older children. The few studies that exist indicate that auditory dominance persists beyond four years of age (Nava & Pavani, 2012). For example, children aged between 4 to 9 years (Massaro, 1984) and 5 to 9 years (Tremblay Champoux, Voss, & Bacon et al., 2007) were less prone to the McGurk effect than adults. As outlined before the McGurk effect refers to the phenomenon that visual information (i.e., a visually perceived lip movement) alters the perception of a sound (McGurk & MacDonald, 1976). Noteworthy, even when children were explicitly told to focus on the speaker´s lips in order to increase attention toward visual input, children showed a diminished McGurk effect as compared to adults. This suggests that for children visual information was less influential than auditory information, even when endogenous attention was directed toward the visual input (Massaro, 1984). However, the authors did not differentiate between 4-year-olds and 9-year-olds but children were handled as one sample in the statistical analysis. As there is quite a range between 4- and 9-year-olds regarding cognitive development, it would have been interesting to differentiate between younger and older children and examine whether there are differences within the group of children.

In fact, there is evidence that visual input gains significance in later childhood (Thompson & Massaro, 1994). In their study, 4- and 9-year-olds were presented with video-clips of a person who pointed to one of four objects and simultaneously spoke the object´s name. Speech and gesture were either congruent (e.g., pointing gesture toward the object “tea” and the verbal label “tea”) or incongruent (e.g., pointing gesture toward the object “tea” and the verbal label “pea”). Participants were asked to judge the intended referential object, and thus faced a conflict in incongruent trials by deciding for either the verbally labeled object or the visually indicated object. For both groups of children, speech had a greater influence on judgements than pointing gestures. However, gestures were significantly more relevant for judgements of older children compared to younger children. The author thus concluded that visual input becomes more relevant as children get older, and thus has more weight in multi-modal processing of 9-year-olds compared to multimodal processing in 4-year-olds.

Moreover, Nava and Pavani (2012), conducting a Colavita experiment (Colavita 1974) with a broader range of age groups, reported a switch from auditory dominance in 6- and 7-year-olds toward visual dominance in 9- and 10-year-olds. Children and adults were presented with an auditory and a visual stimulus, either separately or simultaneously, and had to report whether they perceived an unimodal auditory, visual, or a bimodal stimulus. Presented with bimodal stimuli, older children and adults more often missed the auditory component (i.e., reported the perception of a visual instead of a bimodal stimulus), whereas 6- to 7-year-old children more often missed the visual component (i.e., reported the perception of an auditory instead of a bimodal stimulus). Thus, 9- to 12-year-olds and adults exhibited a Colavita visual dominance effect, whereas 6- to 7-year-olds showed a reversed Colavita effect, i.e. auditory dominance. However, there is evidence that sensory dominance in children is more differentiated than auditory dominance in 6-year-olds and visual dominance in 9-year-olds. In a study by Gori, Sandini, and Burr (2012) several age groups (i.e., adults and 5- to 12-year-olds) were tested for audio-visual integration in space and time. All age groups exhibited visual dominance in the spatial task and an auditory dominance in the temporal task. This indicates that even in young children sensory dominance is differentiated toward auditory dominance in the temporal domain and visual dominance in the spatial domain. However, there are several methodological shortcomings: The study was run as a between-design with one sample performing the temporal task and the other one performing the spatial task. Age ranges within these samples differed and samples were not of comparable size (e.g., the performance of 22 5- to 7-year-olds in the spatial task was compared with the performance of eight 6- to 7-year-olds in the temporal task). Moreover, different age groups were related (e.g., the performance of 13- and 14-year-olds in the temporal task was compared with 12-year-olds´ performance in the spatial task), and, in part, the sample sizes were quite small (e.g. $N = 5$ in the group of 12-year-olds). Taking these shortcomings into account, the validity and reliability of the conclusions drawn from study by Gori, Sandini, and Burr (2012) seems debatable.

6 Summary

A great body of research on multimodal interactions has reported visual dominance in adults. For audio-visual interactions in particular, visual dominance has been reported in spatial tasks but also in other domians (e.g., Colavita visual dominance effect: Colavita 1976;

Mc Gurk effect: McGurk & MacDonald, 1976). In contrast, auditory dominance has been reported in temporal tasks (e.g., Walker & Scott, 1991; Welch et al., 1986). However, several studies (e.g., Shams et al., 2002; Andersen et al., 2004) tested in only one direction (i.e., tested the alteration of temporal aspects of vision by sound, but did not control for alteration of temporal aspects of sound by vision). Thus, besides the demonstration that an auditory stimulus can modify the perceived temporal characteristics of a visual stimulus, these studies do not demonstrate auditory dominance in adults.

For infants and 4-year olds auditory dominance has been reported (infants: Lewkowicz, 1988a, 1988b; Robinson & Sloutsky, 2004; 4-year-olds: Sloutsky & Napolitano, 2003; Massaro, 1984; Thompson & Massaro, 1994), whereby 4-year-olds exhibited modality flexibility depending on stimulus familiarity. Visual dominance was revealed when visual stimuli were more familiar than auditory stimuli. In contrast, higher familiarity of auditory as compared to visual stimuli, or unfamiliarity of both stimuli, elicited auditory dominance in children (Napolitano & Sloutsky, 2004).

Moreover, auditory dominance was seen to persist in 6-year-olds (Thompson & Massaro, 1994; Nava & Pavani, 2012) whereas the visual modality gains weight at the age of 9 years (Thompson & Massaro, 1994) and even starts dominating the auditory modality (Nava & Pavani, 2012).

7 Verbal Overshadowing

Sensory dominance might have relevant implications on learning and memory in children. In order to examine the influence of visual and verbal processing on face memory in children, a recognition task with 4- to 6-year-olds was conducted (Study 3). The theoretical background of this study will be addressed in this section.

Although language is an important cognitive skill, one can accomplish many cognitive tasks such as face recognition without the use of language. Moreover, there is evidence that performance on non-verbal tasks is sometimes impaired by the use of language (Ryan & Schooler, 1998). Specifically, the attempt to verbally describe non-verbal (e.g., visual) stimuli, can affect later recognition of these stimuli in a detrimental way. The negative effect of verbalization on non-verbal memory is termed verbal overshadowing (Schooler, 2002). It has been demonstrated across a number of non-verbal tasks including insight problem solving (Schooler, Ohlsson & Brooks, 1993), taste memory (Melcher & Schooler, 1996), spatial memory (Fiore & Schooler, 2002), implicit learning (Fallshore & Schooler, 1995), color memory (Schooler & Engstler-Schooler, 1990), voice memory (Perfect, Hunt, & Harris, 2002), and face memory tasks (Schooler & Engstler-Schooler, 1990; Dodson, Johnson, & Schooler, 1997).

There are a number of explanatory approaches for verbal overshadowing: The content account referring to dual processing of verbal and non-verbal stimuli, the processing account reffering to the shift from holistic to feature-based processing induced by verbalization, and, finally, the criterion shift account referring to the effect of verbalization on recognition criteria (Chin & Schooler, 2008). In the following section, a brief description will be given for the processing and the criterion shift account and a more detailed description for the content account, as it holds the most important implications for verbal overshadowing effects in children.

7.1 Processing account of Verbal Overshadowing

The processing account refers to verbal overshadowing effects in face recognition tasks. Past research supports the hypothesis that memory for faces relies on holistic processing as opposed to feature-based processing (Chin & Schooler, 2008). That is, although face processing certainly results in the perception of separable facial features, features become integrated into a perceptual whole (Taubert, Apthorp, Aagten-Murphy, & Alais, 2011). Thus, “when it comes to face recognition it is the way the face looks as a whole that matters, and not as much what each individual feature looks like” (Chin & Schooler, 2008, p. 403). On the other hand, verbal face descriptions often refer to certain features, as they are easier to verbalize than for example spatial relations between those features. It has thus been suggested that verbalization of a face following its encoding causes a shift from a holistic processing toward feature-based processing which may carry over into retrieval and impairs face recognition (Chin & Schooler, 2008).

7.2 Criterion shift account of Verbal Overshadowing

The criterion shift account by Clare and Lewandowsky (2004) assumes that verbalization results in a conservative response-bias. That is, people are inclined to choose a “target not present-option” if available (Chin & Schooler, 2008). Indeed, a face recognition experiment by Clare & Lewandowsky (2004) revealed that verbalizers were more accurate than non-verbalizers when presented with a target absent lineup and provided with a “target not present-option”. A target present lineup with a “target not present-option”, however, elicited the verbal overshadowing effect: Verbalizers performed worse than non-verbalizers (Clare & Lewandowsky, 2004). Thus, verbalization can have positive effects, as it makes people less likely to pick just anyone out of the lineup. However, if the target is indeed present in the lineup, verbalizers tend to be more conservative and thus, are more hesistant in decision making than non-verbalizers (Chin & Schooler, 2008; Clare & Lewandowsky, 2004).

7.3 Content account of Verbal Overshadowing

The verbalization of a non-verbal (e.g., visual) stimulus can result in an inaccurate verbal representation of this stimulus, which impairs face recognition (Schooler & Engstler-Schooler, 1990). In other words: People confuse the verbal memory of the visual stimulus with the original visual memory. The content account partially overlaps with the misinformation effect (Loftus & Hoffman, 1989), according to which the exposure to a certain stimulus and successive verbalization of this stimulus can result in self-generated verbal representations that contain misinformation. This in turn distorts the memory of the original stimulus (Chin & Schooler, 2008). However, the content account refers not to traditional interference in which the retrieval of former to-be-remembered information is distorted by information that a person receives afterwards. Rather, the visual representation remains intact but is overshadowed by verbal representations (Schooler & Engstler-Schooler, 1990). Moreover, according to the content account verbal overshadowing routs in a mismatch between perceptual abilities and verbal abilities. That is, when perceptual (e.g., visual) abilities exceed verbal abilities, people would perceive many aspects visually, but would be unable to put all perceptual (e.g., visual) aspects into words. A poor verbal representation of a complex perceptual experience, in turn, can hamper recognition (Chin & Schooler, 2008). The content account might have interesting implications for verbal overshadowing in young children (i.e., 4- to 6-year-olds). Given their limited language skills compared to older children and adults (Brandone, Salkind, Golinkoff, & Hirsh-Pasek, 2006), one could assume that young children are especially vulnerable to verbal overshadowing: presented with a non-verbal stimulus 4- to 6-year-olds certainly perceive many aspects visually but might have problems to transfer all perceptual aspects into corresponding words. Thus, subsequent verbalization may not fully match the perceptual experience. Therefore, children would rely on the impoverished details of verbal representations (overshadowing the more detailed visual representations), which may result in impaired face recognition as described by the verbal overshadowing effect (Schooler & Engstler-Schooler, 1990). On the other hand, given auditory dominance in childhood as reported by Nava and Pavani (2012), and Sloutsky and Napolitano (2003), it might also be possible that children benefit rather than suffer from language-based processing (i.e. verbalization).

8 Focus of research

Given a visual dominance in adults and an auditory dominance in 4-year-olds modality dominance obviously changes at some point of development. As previous research has rarely addressed sensory dominance in children older than 4 years of age, the aim of my dissertation was to further examine the developmental trajectory of sensory dominance. For this purpose, I focused on sensory dominance in older children (6- to 7-year-olds, 8- to 10-year-olds) and adults in two studies.

In Study 1, children and adults performed a Colavita experiment. Contrary to the Colavita experiment conducted by Nava and Pavani (2012), stimuli with semantic content were presented. The main purpose was whether sensory dominance effects in children, as demonstrated by means of simple lights and sounds (Nava & Pavani, 2012) would persist with more complex and meaningful stimuli, too. A further question was whether semantic congruency is a modulating factor for sensory dominance effects. Previous research in adults has shown that the Colavita visual dominance effect is modulated by temporal and spatial congruency, that is, the closer audio-visual stimuli were presented in time and space the larger was the visual dominance effect. According to Koppen and Spence (2007c) this might be explained by a failure of binding: temporal and spatial congruency between audio-visual stimuli provides a window in which auditory and visual stimuli are not perceived as separate events but are bound and perceived as constituting an unitary event instead. At the same time the likelihood for a failure of binding increases. The perception of the visual stimulus may overshadow the awareness of the auditory stimulus, as it sufficiently and adequately describes the unitary audio-visual event, whereas the auditory stimulus provides only redudant information (Koppen & Spence, 2007c). However, given that temporal and spatial congruency modulates the Colavita visual dominance effect in adults, I was interested in whether semantic congruency might also be a modulating factor. So far little research exists for adults (Koppen et al., 2008), and, to my knowledge, no studies have been conducted with children.

In Study 2, I examined whether privileged auditory processing in 6- and 7-year-olds and privileged visual processing in older children and adults, as reported by Nava and Pavani (2012), occurs also for a spatial task. Moreover, I was interested in whether participants exhibited modality dominance (disrupted processing of a modality) or preference (the deliberate choice of a modality that will be attended to, cf. Napolitano & Sloutsky, 2004).

Study 3 refers to verbal overshadowing in young children. To date, research has mainly focused on adults and only few studies exist with children (e.g., Dehon, Vanootighem, & Brédart, 2013; Memon & Rose, 2002). Whereas Dehon et al. (2013) found a verbal overshadowing effect for different age groups (i.e., 7- to 8-year-olds, 9- to 10-year-olds, 13- to 14-year-olds), no effect could be shown for 8- and 9-year-olds in the study of Memon and Rose (2002). Both studies will be discussed in more detail within the scope of Study 3. The aim of this study was to investigate (a) whether the verbal overshadowing effect also occurs in younger children (i.e., 4- to 6-year-olds) (b) whether potential detrimental effects of verbalization might be mediated by verbal intelligence, and (c) whether visualization (i.e., drawing the seen face before recognizing it) might have negative effects on face memory, too.

9 References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14*, 257-262.

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21*, 301-308.

Aruffo, C., & Shore, D. I. (2012). Can you McGurk yourself? Self-face and self-voice in audiovisual speech. *Psychonomic Bulletin & Review, 19*, 66-72.

Basu Mallick, D., Magnotti, J. F., & Beauchamp, M. S. (2015). Variability and stability in the McGurk effect: Contributions of participants, stimuli, time, and response type. *Psychonomic Bulletin & Review, 22*, 1299-30

Berger, T. D., Martelli, M., & Pelli, D. G. (2003). Flicker flutter: is an illusory event as good as the real thing? *Journal of Vision, 3*, 406-412.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics, 62*, 321-332.

Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., & de Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia, 38*, 1634-1642.

Botvinick, M., & Cohen, J. D. (1998). Rubber hand ‘feels’ what eyes see. *Nature, 391*, 756.

Brandone, A. C., Salkind, S. J., Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Language development. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: development, prevention, and intervention* (pp. 499-514). Bethesda, MD: National Association of School Psychologists.

Callan, A., Callan, D., & Ando, H. (2015). An fMRI study of the ventriloquism effect. *Cerebral Cortex*, 25, 4248-4258.

Calvert, G. A., Brammer, M. J., & Iversen, S. D. (1998). Crossmodal identification. *Trends in Cognitive Sciences, 2*, 247-253.

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.) (2004). *The handbook of multisensory processing*. Cambridge, MA: MIT Press.

Chica, A. B., Bartolomeo, P., & Lupiáñez, J. (2013). Two cognitive and neural systems for endogenous and exogenous spatial attention. *Behavioral Brain Research, 15*, 107-123.

Chin, J., & Schooler, J.W. (2008). Why do words hurt? Content, process, and criterion shift accounts of verbal overshadowing. *The European Journal of Cognitive Psychology, 20*, 396-413.

Clare, J. & Lewandowsky, S. (2004). Verbalizing facial memory: criterion effects in verbal overshadowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 739-755.

Clarke, S., Bellmann, A., Meuli, R. A., Assal, G., & Steck, A. J. (2000). Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia, 38*, 797- 807.

Colavita, F. B. (1974). Human sensory dominance. *Perception and Psychophysics, 16*, 409-412.

Colin, C., Radeau, M., Deltenre, P., & Morais, J. (2001) Rules of intersensory integration in spatial scene analysis and speechreading. *Psychologica Belgica, 41*, 131-144.

Dehon, H., Vanootighem, V., & Brédart, S. (2013). Verbal overshadowing of face memory does occur in children too! *Frontiers in Psychology*, *4*. doi: 10.3389/fpsyg.2013.00970

Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: why descriptions impair face recognition. *Memory & Cognition, 25*, 129-139.

Ernst, M. O. (2008). Multisensory integration: a late bloomer. *Current Biology*, *18*, 519-521.

Eskelund, K., MacDonald, E. N., & Andersen, T. S. (2015). Face configuration affects speech perception: evidence from a McGurk mismatch negativity study. *Neuropsychologia, 66*, 48-54.

Fallshore, M., & Schooler, J. W. (1995). Verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning Memory, and Cognition, 21*, 1608-1623.

Farnè, A., Pavani, F., Meneghello, F., & Làdavas, E. (2000). Left tactile extinction following visual stimulation of a rubber hand. *Brain, 12*, 2350-2360.

Fendrich, R., & Corballis, P.M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics, 63*, 719-725.

Fetsch, C. R., De Angelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience, 14*, 429-442.

Fiore, F. M., & Schooler, J. W. (2002). How did you get here from there? Verbal overshadowing of spatial mental models. *Applied Cognitive Psychology, 16*, 897-910.

Grant, K. W., & Seitz, P. F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *Journal of the Acoustical Society of America, 104*, 2438-2450.

Green, K., Kuhl, P., Meltzoff, A., & Stevens, E. (1991). Integrating speech information across talkers, gender, and sensory modality: female faces and male voices in the McGurk effect. *Perception and Psychophysics, 50*, 524-536.

Hartcher-O'Brien, J., Gallace, A., Krings, B., Koppen, C., & Spence, C. (2008). When vision 'extinguishes' touch in neurologically-normal people: extending the colavita visual dominance effect. *Brain Research, 186*, 643-658.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. New York: Wiley.

Jones, J. A., & Munhall, K. G. (1997). The effects of separating auditory and visual sources on audio-visual integration of speech. *Canadian Acoustics, 2*, 13–19.

Kopčo, N., Lin, I. F., Shinn-Cunningham, B. G., & Groh, J. M. (2009). Reference frame of the ventriloquism aftereffect. *The Journal of Neuroscience, 29*, 13809-13814.

Koppen, C., & Spence, C. (2007a). Seeing the light: exploring the colavita visual dominance effect. *Experimental Brain Research, 180*, 737-754.

Koppen, C., & Spence, C. (2007b). Spatial coincidence modulates the colavita visual dominance effect. *Neuroscience Letters, 417*, 107-111.

Koppen, C., & Spence, C. (2007c). Audiovisual asynchrony modulates the colavita visual dominance effect. *Brain Research, 1186*, 224-232.

Koppen, C., & Spence, C. (2007d). Assessing the role of stimulus probability on the colavita visual dominance effect. *Neuroscience Letters, 418*, 266-271.

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the colavita visual dominance effect. *Experimental Brain Research, 184*, 533-546.

Koppen, C., Levitan, C. A., & Spence, C. (2009). A signal detection study of the colavita visual dominance effect. *Experimental Brain Research, 196*, 353-360.

Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learning & Memory, 9*, 268-278.

Lewkowicz, D. J. & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13*, 470-478.

Lewkowicz, D. J. (1988a). Sensory dominance in infants: I. six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155-171.

Lewkowicz, D. J. (1988b). Sensory dominance in infants: II. ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 172-182.

Loftus, E. F. & Hoffman, H. G. (1989). Misinformation and memory: the creation of new memories. *Journal of Experimental Psychology: General, 118*, 100-104.

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage 21*, 725-732.

Magnotti, J. F., Basu Mallick, D. B., Feng, G., Zhou, B., Zhou, W., & Beauchamp, M. S. (2015). Similar frequency of the McGurk effect in large samples of native mandarin chinese and american english speakers. *Experimental Brain Research, 233*, 2586-2600.

Massaro, D. W. (1987). *Speech perception by ear and by eye: paradigm for psychological inquiry*. Hillsdale: Erlbaum.

Massaro, D. W. (1984). Children´s perception of visual and auditory speech. *Child Development, 55*, 1777-1788.

Massaro, D. W., & Cohen, M. M. (2000). Tests of auditory-visual integration efficiency with in the framework of the fuzzy logical model of perception. *Journal of the Acoustical Society of America, 108*, 784-789.

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

Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language, 35*, 231-245.

Memon, A., & Rose, R. (2002). Identification abilities of children: does verbalization impair face and dog recognition? *Psychology, Crime and Law, 8*, 229-242.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research, 17*, 154-163.

Moro S. S., & Steeves, J. K. E. (2012). No colavita effect: equal auditory and visual processing in people with one eye. *Experimental Brain Research, 216*, 367-373.

Napolitano, A. (2006). *Modality dominance in young children: the underlying mechanisms and broader implications* (Unpublished doctoral dissertation). The Ohio State University, Ohio, USA. Retrieved from https://etd.ohiolink.edu/

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*, 1850-1870.

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage, 59*, 781-787.

Nava, E., & Pavani, F. (2012). Changes in sensory dominance during childhood: converging evidence from the colavita effect and the sound-induced flash illusion. *Child Development, 84*, 604-116.

Nelson, L. (2004). Reverse ventriloquism. *Nature News.* Retrieved from: http://www.nature.com/news/2004/040203/full/news040202-3.html

Pavani, F., Spence, C., & Driver. J. (2000). Visual capture of touch: out-of-the-body experiences with rubber gloves. *Psychological Science, 11*, 353-359.

Perfect, T. J., Hunt, L. J., & Harris, C. M. (2002). Verbal overshadowing in voice recognition. *Applied Cognitive Psychology, 16*, 973-980.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: an information processing account of its origins and significance. *Psychological Review, 83*, 157-171.

Recanzone, G. H. (2002). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology, 89*, 1078–1093.

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proceedings of the National Academy of Sciences USA, 95*, 869-875.

Robinson, C. W., & Sloutsky, V. M. (2010). Effects of multimodal presentation and stimulus familiarity on auditory and visual processing. *Journal of Experimental Child Psychology, 107*, 351-358.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*, 1387-1401.

Rodway, P. (2005). The modality shift effect and the effectiveness of warning signals in different modalities. *Acta Psychologica*, *120*, 199-226.

Ryan, R. S., & Schooler, J. W. (1998).Whom do words hurt? Individual differences in susceptibility to verbal overshadowing. *Applied Cognitive Psychology, 12* (7), 105-125.

Sams, M., Manninen, P., Surakka, V., Helin, P., & Katto, R. (1998). McGurk effect in finnish syllables, isolated words, and words in sentences: effects of word meaning and sentence context. *Speech Communication, 26*, 75-87.

Scheier, C. R., Nijhawan, R., & Shimojo, S. (1999). Sound alters visual temporal resolution. *Investigative Ophthalmology & Visual Science, 40*, 792-792.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology, 16*, 989-997.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology*, *22*, 36-71.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: when language overshadows insight. *Journal of Experimental Psychology: General, 122*, 166-183.

Sekiyama, K. (1997). Cultural and linguistic factors in audiovisual speech processing: the McGurk effect in chinese subjects. *Perception and Psychophysics, 59*, 73-80.

Sekiyama, K. & Tohkura, Y. (1993). Interlanguage differences in the influence of visual cues in speech perception. *Journal of Phonetics, 21*, 427-444.

Shams, L., Kamitani, Y., & Shimojo, S., (2004). Modulations of visual perception by sound. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes*, (pp. 27-34). Cambridge, MA: MIT Press.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*, 147-152.

Sinnett, S., Spence, C., &. Soto-Faraco, S. (2007). Visual dominance and attention: the colavita effect revisited. *Perception and Psychophysics, 69*, 673-686.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833.

Spence, C., Parise, C., & Chen, Y. (2012). The colavita visual dominance effect. In M. Murray & M. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 529-550). Boca Raton: CRC Press.

Strand, J., Cooperman, A., Rowe, J., & Simenstad, A. (2014). Individual differences in susceptibility to the McGurk effect: links with lipreading and detecting audiovisual incongruity. *Journal of Speech, Language, and Hearing*, 57, 2322-2331.

Summerfield, Q. A., & McGrath, M. (1984). Detection and resolution of audio-visual incompatibility in the perception of vowels. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 36,* 51-74.

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: evidence from the face inversion effect. *Vision Research, 5*, 1273-1278.

Thompson, L. A., & Massaro, D. W. (1994). Children's integration of speech and pointing gestures in comprehension. *Journal of Experimental Child Psychology, 57*, 327-354.

Turatto, M., Benso, F., Galfano, G., Gamberini, L., & Umiltà, C. (2002). Non-spatial attentional shifts between audition and vision. *Journal of Experimental Psychology: Human Perception & Performance, 28,* 628-639.

Ungerleider, L.G. & Mishkin, M. (1982) Two cortical visual systems. In D. J. Ingle, R. J. W. Mansfield, & M. D. Goodale (Eds.), *The analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.

Vroomen, J., & de Gelder, B. (2004). Perceptual effects of cross-modal stimulation: ventriloquism and the freezing phenomenon. In G. A. Calvert, C. Spence, Charles, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 141-150). Cambridge, MA: MIT Press.

Vroomen, J., Bertelson, P., & de Gelder, B. (1998). A visual influence in the discrimination of auditory location. In D. Burnham, J. Robert-Ribes, & E. Vatikiotos-Bateson (Eds.), *Proceedings of the international conference on auditory-visual speech processing, AVSP'98* (pp. 131–135) [CD]. Adelaide, Australia: Causal Productions.

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1327-1339.

Warren, D. H., Welch, R. B., & McCarthy, T. J. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: implications for transitivity among the spatial senses. *Perception & Psychophysics, 30*, 557-564.

Welch, R. B., Dutton Hurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics, 3*9, 294-300.

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions, in K. R. Boff, L. Kaufman, & J. P. Thomas (Eds), *Handbook of perception and human performance* (pp. 1-36). New York: Wiley.

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron*, *48*, 489-496.

CHAPTER 2

Study 1

Semantic congruency and the (reversed) Colavita effect in children and adults

Abstract

When presented with auditory, visual, or bimodal audiovisual stimuli in a discrimination task, adults tend to ignore the auditory component in bimodal stimuli and respond to the visual component only (i.e., Colavita visual dominance effect). The same is true for older children, whereas young children are dominated by the auditory component of bimodal audiovisual stimuli. This suggests a change of sensory dominance during childhood. The aim of the current study was to investigate, in three experimental conditions, whether children and adults show sensory dominance when presented with complex semantic stimuli and whether this dominance can be modulated by stimulus characteristics such as semantic (in)congruency, frequency of bimodal trials, and color information. Semantic (in)congruency did not affect the magnitude of the auditory dominance effect in 6-year-olds or the visual dominance effect in adults, but it was a modulating factor of the visual dominance in 9-year-olds (Conditions 1 and 2). Furthermore, the absence of color information (Condition 3) did not affect auditory dominance in 6-year-olds and hardly affected visual dominance in adults, whereas the visual dominance in 9-year-olds disappeared. Our results suggest that (a) sensory dominance in children and adults is not restricted to simple lights and sounds, as used in previous research, but can be extended to semantically meaningful stimuli and that (b) sensory dominance is more robust in 6-year-olds and adults than in 9-year-olds, implying a transitional stage around this age.

Keywords: Colavita effect, auditory dominance, visual dominance, semantic congruency, perception, multisensory processing

Introduction

Sensory integration as a functional principle of the human brain enables a versatile and coherent perception of the environment by combining multimodal information into a unified event (cf. Calvert, Spence, & Stein, 2004; Lewkowicz & Ghazanfar, 2009). Nevertheless, one sensory system often dominates another one when multiple sensory systems are stimulated concurrently. A prominent example for sensory dominance is the ventriloquist effect; the visually perceived sound source dominates the localization of the true sound source (Howard, Craske, & Templeton, 1966; Slutsky & Recanzone, 2001). In fact, research on sensory integration suggests that vision often biases the processing not only of the auditory system but also of the tactile and proprioceptive system (Botvinick & Cohen, 1998; Farnè, Pavani, Meneghello, & Làdavas, 2000; Gallace & Spence, 2005; Hartcher-O'Brien, Gallace, Krings, Koppen, & Spence, 2008). Thus, at least for adults, visual input seems to be more influential than input from other sensory modalities (but see Alais & Burr, 2004).

Another striking example of visual dominance is the Colavita effect (Colavita, 1974). When presented with bimodal stimuli consisting of an auditory component and a visual component (e.g., a 400-Hz tone and a visual angle), adults often claim that they have perceived only a visual stimulus while ignoring the auditory component. Apparently, the visual dominance of adults is strong enough to overshadow the awareness for an auditory input when being presented with synchronous bimodal stimuli. Over the years, a vast number of studies have been conducted to further examine this phenomenon, and in nearly all of the studies robust effects in favor of the visual input were shown (e.g., Koppen, Levitan, & Spence, 2009; Sinnett, Spence, & Soto-Faraco, 2007). Although effect sizes varied across studies and could be systematically reduced by some designs (Sinnett et al., 2007), the Colavita effect in adults could not be reversed in terms of auditory dominance over the visual system (for a review, see Spence, Parise, & Chen, 2012).

Interestingly, the dominance of visual perception does not seem to be innate because there is evidence that children's information processing is dominated by the auditory system (see Lewkowicz, 1988a, 1988b; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003), but because research has rarely tested children older than 4 years, several questions remain unanswered. Given a visual dominance in adults and an auditory dominance in 4-year-olds, obviously a change in modality dominance occurs during the course of development.

One of the rare studies examining the developmental trajectory of sensory dominance across older children (i.e., 6-7, 9-10, and 11-12 years) and adults using the same procedure and materials for all age groups is the study of Nava and Pavani (2012). Although challenging, standardized procedures for all age groups are essential in order to draw legitimate conclusions. Based on the design used by Colavita (1974), Nava and Pavani (2012) could show a switch from auditory dominance in 6- and 7-year-olds toward visual dominance in 9- and 10-year-olds. Children were presented with simple lights and sounds, either separately or simultaneously, and were asked whether they perceived a single auditory stimulus, a single visual stimulus, or bimodal stimuli. In the bimodal condition, older children and adults often missed the auditory component by indicating that they had perceived only a visual stimulus (Colavita effect), whereas younger children often missed the visual component of bimodal stimuli suggesting an auditory dominance (reversed Colavita effect).

In the study of Nava and Pavani (2012), auditory and visual components of bimodal stimuli were synchronously presented at the same spatial location, yielding spatial and temporal congruency. In the current study, we also examined the effect of semantic congruency on sensory dominance, which is adapted from the study of Koppen, Alsius, and Spence (2008) with adults. Thus, instead of simple lights and sounds, semantically meaningful stimuli were presented to participants of different age groups. Our main question was the following: Can effects of sensory dominance in children, as demonstrated by means of simple stimuli

(Nava & Pavani, 2012), be extended to the processing of complex and meaningful stimuli? If these dominance effects are a robust empirical phenomenon, one should expect them to emerge even if more realistic stimuli are used. Furthermore, we were interested in whether the manipulation of semantic congruency affects the magnitude of sensory dominance effects in different age groups. Multisensory stimuli of a single object or event usually share not only temporal and spatial attributes but also certain semantic features that facilitate their identification (Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004). For instance, a barking four-legged animal is labeled as a dog. In contrast, semantic congruency is lacking when being presented with a ringing four-legged animal. Whereas the ‘‘semantic mismatch’’ in semantically incongruent trials should contribute to a better detection or ‘‘pop-out’’ of both stimuli components, there should be a facilitating effect regarding sensory dominance in the semantically congruent trials; in congruent trials, older children and adults - usually exhibiting a visual dominance -should show a stronger tendency to overhear the auditory component of a bimodal stimulus combination because the auditory components contain only ‘‘redundant information.’’ This should also be the case for younger children, but due to their auditory dominance, they should neglect the visual component; in congruent trials, they should show a stronger tendency to overlook the visual component of a bimodal stimulus combination compared with incongruent trials. Thus, sensory dominance should become more evident in semantically congruent trials than in incongruent trials. An experiment with three conditions was conducted. In all of them, stimulus congruency was manipulated. In addition, the number of bimodal trials was varied between Condition 1 and Condition 2 to examine whether the relative frequency of bimodal stimuli has an impact on sensory dominance. In Condition 3, we examined whether the absence of color information could remove visual dominance in older children and adults.

Method

Participants

Participants were 65 6-year-olds (*M* = 6.40 years, *SD* = 0.49; 30 boys and 35 girls attending preschool or first grade), 71 9-year-olds (*M* = 9.00 years, *SD* = 0.81; 34 boys and 37 girls attending second or third grade), and 71 adults (*M* = 27.27 years, *SD* = 8.28; 30 men and 41 women). Children were recruited from local schools and nurseries after their parents had signed a consent form. Adults were staff members and students from a university of the same medium-sized city and were unfamiliar with the aim of the experiment. Participants were randomly assigned to one of the three conditions. For sample size, mean age, and gender distribution in the three experimental conditions, see Table 1.

Stimuli and procedure

Stimulus programming, presentation, and response collection were carried out using the program Inquisit. Visual stimuli consisted of four colored schematic images - a brown dog, a cow with brown patches, a red telephone, and a golden bell, each measuring 12 x 9 cm - presented at the center of a white-background monitor measuring 42.6 x 32.7 cm. Auditory stimuli consisted of four sounds - the bark of a dog, the moo of a cow, the ring of a phone, and the toll of a bell - presented on headphones at 65 db. Either visual or auditory stimulus was presented alone (auditory-only and visual-only trials) or a visual stimulus and an auditory stimulus were presented together (bimodal trials). Bimodal trials consisted of semantically congruent and incongruent trials. In incongruent trials, all possible combinations of visual and auditory stimuli were presented. Participants sat in front of the screen with a viewing distance of approximately 50 cm. They were instructed to press one button (marked by the symbol of a picture frame) in response to a single visual stimulus, another button in response to a single auditory stimulus (marked by the symbol of a note), or a third button whenever they perceived a visual stimulus and an auditory stimulus simultaneously (marked by the symbol of a picture frame with a note in it). An experimental session lasted approximately 10 min. Each trial started with the onset of the target presented for maximally 500 ms. If a participant did not respond within this period, the target disappeared, followed by a white screen (which was consistently presented in the auditory-only trials). A response was mandatory in order to continue with the experiment. If a participant did not give a response after 5 s, he or she was asked to remember what kind of stimulus was presented and to guess if necessary. After the response and an interstimulus interval of 750 ms, the next target emerged. No feedback was provided except for the training trials, in which the experimenter said ‘‘true’’ or ‘‘false.’’ Participants were instructed to react as fast and as accurate as possible. After completing the experiment, adults were provided with information about the aim of the study and children received an attendance certificate and stickers.

Table 1

Sample size, mean age and gender distribution in the three experimental conditions

	Condition 1	Condition 2	Condition 3
6-year-olds			
N	28	19	18
Mean age (*SD*)	6.39 (0.50)	6.42 (0.51)	6.39 (0.50)
Male / female	13 / 15	9 / 10	8 / 10
9-year-olds			
N	30	17	24
Mean age (*SD*)	8.87 (0.82)	9.06 (0.75)	9.13 (0.85)
Male / female	13 / 17	7 / 10	10 / 14
Adults			
N	28	23	20
Mean age (*SD*)	25.96 (5.81)	27.61 (10.52)	28.70 (8.48)
Male / female	12 / 16	11 / 12	7 / 13

Experimental conditions

Each condition started with 12 training trials presented in random order, among them 3 semantically congruent trials, 3 semantically incongruent trials, 3 auditory-only trials, and 3 visual-only trials.

In Condition 1 (‘‘standard’’), training trials were followed by 96 experimental trials consisting of 32 visual-only trials, 32 auditory-only trials, 16 semantically incongruent trials, and 16 semantically congruent trials that were also presented in random order.

Condition 2 (‘‘balanced trials’’) was identical to Condition 1 with the sole exception that an equal number of unimodal and bimodal stimuli were presented (i.e., 64 unimodal trials: 32 visual-only and 32 auditory-only trials; 64 bimodal trials: 32 semantically incongruent and 32 semantically congruent trials). This served to check whether sensory dominance is affected by bimodal stimulus frequency. In the study of Koppen and Spence (2007) examining this issue in adults, a robust visual dominance effect was revealed when bimodal stimuli were presented in 60% or less of all trials, but a declining effect emerged when bimodal stimuli were presented more frequently. Based on these findings, an equalized frequency of unimodal and bimodal stimuli should not affect the magnitude of the sensory dominance effect in adults but potentially could do so in children.

Condition 3 (‘‘black and white’’) was identical to Condition 2 with the sole exception that visual stimuli were presented in black and white instead of colored images. Color information plays a crucial role in detecting and recognizing objects. For example, in a study of Wurm, Legge, Isenberg, and Luebker (1993), participants needed to name food objects presented as gray-scaled or colored images. Reaction times were significantly shorter and accuracy was higher in the condition of colored images. Furthermore, shorter reaction times for colored stimuli were related to objects’ prototypicality but not to their color diagnosticity.[1] Thus, it was concluded that color does improve object recognition and that the underlying mechanism is probably sensory rather than cognitive in origin (Wurm et al., 1993). Given these premises, removing color information might affect the Colavita effect in a detrimental way; because black and white images lack salience compared with colored images, the privileged processing of visual stimuli (compared with auditory stimuli) could be restrained. Therefore, we hypothesized that the visual dominance of older children and adults would be reduced compared with Condition 2 or would even disappear, whereas the auditory dominance of younger children would not be affected and would still be present.

Results

Condition 1 (“standard”)

The Colavita effect is defined by significantly more visual-only responses (in the following called ‘‘visual errors’’) than auditory-only responses (in the following called ‘‘auditory errors’’) in bimodal trials (Koppen et al., 2008). A reversed Colavita effect in terms of an auditory dominance would be defined, in turn, by significantly more auditory errors than visual errors in bimodal trials. For a summary of the mean error rates in bimodal trials of each age group, see Table 2.

To check whether there were more visual or auditory errors across the age groups and semantic conditions, an analysis of variance (ANOVA) with repeated measures on the within-participants variables semantic congruency (congruent vs. incongruent trials) and error type (visual vs. auditory errors in bimodal trials) and on the between-participants variable age group (6-year-olds vs. 9-year-olds vs. adults) was conducted. The number of visual versus auditory errors in bimodal trials served as the dependent measure. One 6-year-old child was excluded from the analyses because he made more than 50% errors in the unimodal condition.

[1] Objects vary in the degree to which their colors are ‘‘diagnostic’’ Biederman and Ju (1988). Although virtually no object can be recognized on the basis of its color alone, some objects (e.g., a banana) are associated more strongly with a particular color than others (e.g., a telephone) Wurm et al. (1993).

There was no significant main effect of error type (p = .45), but there were significant main effects of age group, $F(2, 82) = 7.12$, $p = .001$, $\eta^2 = .15$, and congruency, $F(1, 82) = 32.04$, $p < .001$, $\eta^2 = .28$. A significant interaction emerged between error type and age group, $F(2, 82) = 9.18$, $p < .001$, $\eta^2 = .18$. To specify this effect, an ANOVA with repeated measures was conducted separately for each age group.

The 6-year-olds made significantly more auditory errors than visual errors in bimodal trials, suggesting auditory dominance, $F(1, 26) = 5.60$, $p = .026$, $\eta^2 = .18$. Furthermore, they made significantly more errors in semantically congruent trials than in incongruent trials, $F(1, 26) = 12.77$, $p = .001$, $\eta^2 = .33$. Contrary to our prediction, no significant interaction between error type and congruency emerged ($p = .16$). Thus, auditory dominance of 6-year-olds was not affected by semantic congruency on a significant level.

The 9-year-olds, in contrast, made significantly more visual errors than auditory errors in bimodal trials, suggesting visual dominance, $F(1, 29) = 7.55$, $p = .010$, $\eta^2 = .21$. Furthermore, they made significantly more errors in semantically congruent trials compared with incongruent trials, $F(1, 29) = 14.97$, $p = .001$, $\eta^2 = .34$. As expected, a significant interaction between error type and congruency was revealed, $F(1, 29) = 5.01$, $p = .033$, $\eta^2 = .15$. This was attributable to significantly more visual errors than auditory errors in semantically congruent trials, $t(29) = 2.72$, $p = .011$, $d = 0.63$, whereas no significant difference was revealed in incongruent trials ($p = .33$). Thus, the visual dominance of 9-year-olds occurred only under semantic congruency.

Similar to 9-year-olds, adults made significantly more visual errors than auditory errors in bimodal trials, suggesting visual dominance, $F(1, 27) = 6.42$, $p = .017$, $\eta^2 = .19$. Similar to the two groups of children, they made significantly more errors in semantically congruent trials than in incongruent trials, $F(1, 27) = 5.70$, $p = .024$, $\eta^2 = .17$, but no significant interaction between error type and congruency was revealed ($p = .25$). Thus, visual dominance of adults was not affected by semantic congruency on a significant level. Because the assignment of response keys was not counterbalanced in the current study, additional analyses were conducted to check whether sensory dominance was no result of different response biases in each age group. A response bias could be identified among the unimodal trials (i.e., visual-only and auditory-only trials), as reflected by an unequal distribution of visual and auditory errors due to a ‘‘preferred key.’’ However, paired *t*-tests for each age group (adjusted alpha level set at .016) revealed no significant differences between auditory and visual errors.

Table 2

Mean error rates (in %) for bimodal trials in total and separately for congruent and incongruent trials (Condition 1)

	6-year-olds		9-year-olds		Adults	
Mean error rate in bimodal trials (total)	17.1	(10.2)	10.10	(6.9)	9.49	(7.5)
Auditory errors	10.8	(9.3)	3.13	(3.1)	3.57	(5.8)
Visual errors	6.4	(3.6)	6.98	(6.6)	6.03	(2.9)
Mean error rate in incongruent trials	13.2	(7.6)	5.62	(7.0)	6.92	(7.1)
Auditory errors	7.9	(8.2)	2.29	(3.8)	2.90	(4.0)
Visual errors	5.3	(3.7)	3.33	(5.1)	4.02	(5.2)
Mean error rate in congruent trials	21.1	(14.7)	14.59	(11.2)	12.28	(11.6)
Auditory errors	13.7	(12.1)	3.96	(6.5)	4.24	(5.1)
Visual errors	7.4	(6.7)	10.63	(5.3)	8.04	(9.3)

Note. Standard deviations in parentheses.

Condition 2 ("balanced trials")

For a summary of the mean error rates in bimodal trials, see Table 3. In line with Condition 1, there were significant main effects of age group, $F(2, 56) = 9.76$, $p < .001$, $\eta^2 = .26$, and congruency, $F(1, 56) = 19.76$, $p < .001$, $\eta^2 = .26$. Moreover, a significant interaction emerged between error type and age group, $F(2, 56) = 9.77$, $p < .001$, $\eta^2 = .26$. To specify the effects of congruency on error type, an ANOVA with repeated measures was conducted for each group separately. The results were largely comparable to those of Condition 1.

The 6-year-olds made significantly more auditory errors than visual errors in bimodal trials, suggesting auditory dominance, $F(1, 18) = 6.72$, $p = .018$, $\eta^2 = .27$. Furthermore, they made significantly more errors in semantically congruent trials than in incongruent trials, $F(1, 18) = 5.11$, $p = .036$, $\eta^2 = .22$, and a marginal interaction between error type and congruency was revealed ($p = .08$). This might be due to more auditory errors than visual errors in semantically congruent trials and to a comparable amount of auditory and visual errors in incongruent trials (see Table 3).

The 9-year-olds made significantly more visual errors than auditory errors in bimodal trials, suggesting visual dominance, $F(1, 16) = 4.84$, $p = .043$, $\eta^2 = .23$. Furthermore, they made significantly more errors in semantically congruent trials compared with incongruent trials, $F(1, 16) = 18.00$, $p = .001$, $\eta^2 = .53$. In line with Condition 1 and our predictions, a significant interaction between error type and congruency was revealed, $F(1, 16) = 5.12$, $p = .038$, $\eta^2 = .24$. This was attributable to significantly more visual errors than auditory errors in semantically congruent trials, $t(16) = 2.64$, $p = .018$, $d = 0.85$, whereas no significant difference was revealed in incongruent trials ($p = .77$). Thus, visual dominance of 9-year-olds occurred only under semantic congruency.

Adults also made significantly more visual errors than auditory errors in bimodal trials, suggesting visual dominance, $F(1, 22) = 7.91$, $p = .010$, $\eta^2 = .26$. Similar to the two groups of children, they exhibited significantly more errors in semantically congruent trials than in incongruent trials, $F(1, 22) = 6.30$, $p = .020$, $\eta^2 = .22$. No significant interaction between error type and congruency was revealed ($p = .87$). As in Condition 1, a response bias could be ruled out.

Further analyses were conducted to examine whether the higher frequency of bimodal trials in Condition 2 affected the magnitude of the sensory dominance effects compared with Condition 1. The magnitude of a sensory dominance effect can be conceived as the relative difference between auditory and visual errors on bimodal trials (Koppen et al., 2008). Thus, the difference between auditory errors and visual errors on bimodal trials (both in %) was compared between Condition 1 and Condition 2. A 3 (Age Group: 6-year-olds vs. 9-year-olds vs. adults) x 2 (Condition: 1 vs. 2) ANOVA was conducted with sensory dominance as the dependent variable. There was a significant main effect of age, $F(2, 138) = 11.73$, $p < .001$, $\eta^2 = .15$, no significant main effect of condition ($p = .62$), and no interaction ($p = .16$). Thus, the magnitude of visual dominance in older children and adults, as well as the magnitude of auditory dominance in younger children, did not differ significantly between Condition 1 and Condition 2, suggesting that sensory dominance effects were not affected by the additional number of bimodal trials in the current experiment.

However, the mean total error rate in bimodal trials - cumulated across the percentage of visual and auditory errors - was higher in the current condition than in Condition 1, suggesting that the higher frequency of bimodal trials might have had an effect on the general accuracy. To test this, a 3 (Age Group: 6-year-olds vs. 9-year-olds vs. adults) x 2 (Condition: 1 vs. 2) ANOVA with the mean total error rate in bimodal trials as the dependent variable was conducted. Significant main effects of age group, $F(2, 138) = 11.23$, $p < .001$, $\eta^2 = .14$, and condition, $F(1, 138) = 43.73$, $p < .001$, $\eta^2 = .24$, were revealed, but no significant interaction between age group and condition was revealed ($p = .39$). Thus, all age groups exhibited a higher accuracy in bimodal trials of the current condition compared with Condition 1. This suggests facilitating effects due to the additional number of bimodal trials even if this suggestion needs to be drawn cautiously because two different samples participated in Conditions 1 and 2.

Table 3

Mean error rates (in %) for bimodal trials in total and separately for congruent and incongruent trials (Condition 2)

	6-year-olds		9-year-olds		Adults	
Mean error rate in bimodal trials (total)	7.32	(5.4)	4.14	(2.8)	2.38	(2.1)
Auditory errors	4.69	(3.9)	1.38	(1.4)	0.68	(0.9)
Visual errors	2.63	(2.3)	2.76	(2.2)	1.70	(1.7)
Mean error rate in incongruent trials	5.60	(5.2)	2.76	(2.4)	1.49	(2.1)
Auditory errors	3.13	(4.4)	1.29	(1.6)	0.27	(0.1)
Visual errors	2.47	(2.9)	1.47	(1.9)	1.22	(1.8)
Mean error rate in congruent trials	9.05	(7.3)	5.51	(3.6)	3.26	(3.2)
Auditory errors	6.25	(4.9)	1.47	(2.0)	1.09	(1.5)
Visual errors	2.80	(3.3)	4.04	(3.3)	2.17	(2.7)

Note. Standard deviations in parentheses.

Condition 3 ("black and white")

For a summary of the mean error rates in bimodal trials, see Table 4. Analogous to the preceding conditions, an ANOVA with repeated measures was conducted. Significant main effects of error type, $F(1, 59) = 4.67$, $p = .035$, $\eta^2 = .07$, age group, $F(2, 59) = 27.84$, $p < .001$, $\eta^2 = .49$, and congruency, $F(1, 59) = 14.34$, $p < .001$, $\eta^2 = .20$, were revealed. In addition, there was a significant interaction between error type and age group, $F(2, 59) = 8.63$, $p < .001$, $\eta^2 = .23$. To specify the effects of congruency on error type, an ANOVA with repeated measures was conducted for each group separately.

The results for 6-year-olds were similar to those for the two previous conditions; they made significantly more auditory errors than visual errors in bimodal trials, suggesting an auditory dominance, $F(1, 17) = 14.13$, $p = .002$, $\eta^2 = .45$. Furthermore, they made significantly more errors in semantically congruent trials than in incongruent trials, $F(1, 17) = 6.86$, $p = .018$, $\eta^2 = .29$. No significant interaction between error type and congruency was revealed ($p = .10$).

In contrast to Conditions 1 and 2, 9-year-olds showed no significant main effect of error type and no interaction, suggesting a lack of sensory dominance. A marginal main effect of congruency was revealed, $F(1, 23) = 3.04$, $p = .095$, $\eta^2 = .12$, with more errors in semantically congruent trials than in incongruent trials.

Adults - in line with previous conditions - tended to make more visual errors than auditory errors in bimodal trials, as indicated by a marginal effect of error type suggesting the tendency of visual dominance, $F(1, 19) = 4.30$, $p = .069$, $\eta^2 = .16$. Furthermore, adults tended to make more errors in semantically congruent trials than in incongruent trials, $F(1, 19) = 3.31$, $p = .052$, $\eta^2 = .18$. There was no significant interaction between error type and congruency ($p = .33$). A response bias could again be ruled out.

Table 4

Mean error rates (in %) for bimodal trials in total and separately for congruent and incongruent trials (Condition 3)

	6-year-olds		9-year-olds		Adults	
Mean error rate in bimodal trials (total)	17.54	(8.3)	8.01	(5.6)	3.05	(2.6)
Auditory errors	10.94	(5.2)	4.17	(3.6)	1.17	(1.3)
Visual errors	6.60	(4.4)	3.84	(3.7)	2.34	(2.4)
Mean error rate in incongruent trials	14.93	(8.0)	6.77	(5.3)	1.88	(3.1)
Auditory errors	8.68	(4.7)	3.65	(3.9)	.94	(1.5)
Visual errors	6.25	(4.7)	3.13	(3.8)	1.41	(3.0)
Mean error rate in congruent trials	20.14	(10.5)	9.24	(7.7)	4.22	(4.0)
Auditory errors	13.19	(7.6)	4.69	(4.9)	1.41	(1.9)
Visual errors	6.94	(5.5)	4.56	(5.5)	3.28	(4.1)

Note. Standard deviations in parentheses.

Discussion

The two main questions of our study were (a) whether visual dominance of 9-year-olds, as well as auditory dominance of 6-year-olds (cf. Nava & Pavani, 2012), can be extended to the processing of complex and semantically meaningful stimuli, and (b) whether the manipulation of semantic congruency, as well as of other stimulus characteristics (i.e., relative stimulus frequency and color information), affects the magnitude of sensory dominance effects in children and adults.

Sensory dominance effects in children and adults using complex stimuli. In all three experimental conditions, a robust visual dominance effect was revealed for adults, whereas 6-year-olds exhibited an auditory dominance. Except for Condition 3, a visual dominance effect could also be shown for 9-year-olds. This suggests that sensory dominance in children can be extended to the processing of complex and meaningful stimuli and does not occur only with simple lights and sounds used in previous research (Nava & Pavani, 2012).

The increased frequency of bimodal stimuli relative to unimodal stimuli in Condition 2 apparently reduced the mean error rate on bimodal trials but had no impact on the magnitude of sensory dominance effects. Because this was the case for all age groups, one can assume a similar robustness of sensory dominance effects regarding bimodal stimulus frequencies. The manipulation of visual salience in Condition 3 showed that the absence of color information hardly affected visual dominance in adults and had no effect on auditory dominance in 6-year-olds, but it seemed to undermine the visual dominance of 9-year-olds, who exhibited no sensory dominance in this particular condition. Obviously, there is a developmental trajectory of sensory dominance with an auditory dominance during early childhood, a visual dominance typically observed in adults, and a transition occurring at around 9 years of age - which is underlined by our current study showing that sensory dominance is vulnerable to stimulus manipulations among 9-year-olds.

However, the question remains as to why this transition takes place at this age, including its underlying mechanisms. It is rather unlikely that maturational asynchronies between the auditory and visual modalities can account for this. In fact, the subcortical auditory system matures earlier compared with the visual subcortical system (Lippé, Kovacevic, & McIntosh, 2009). Conversely, at the (higher order) thalamocortical level, the tendency seems to be reversed. The visual thalamocortical system shows relative maturity at 5 months of age, whereas the maturation of the thalamic projections to the auditory cortex continues until 6 years of age. In addition, myelination begins earlier in the occipital lobe (visual processing center) than in the temporal lobe (auditory processing center). Synapse density shows a rapid increase in the occipital lobe until approximately 8 months of age, which is followed by a decline, whereas the synaptic density of the auditory cortex increases until 4 years of age before connections are pruned (Lippé et al., 2009). In sum, the auditory system differentiates slower than the visual system. From this point of view, a weighted integration of visual and auditory inputs based on the maturational level of the two sensory systems during early childhood would rather benefit the visual one. Thus, one would expect visual dominance in young children rather than an auditory dominance.

Potentially, other factors can better account for modality dominance effects in children. For example, 4-year-olds exhibited visual dominance when visual stimuli were more familiar than auditory stimuli (Napolitano & Sloutsky, 2004). In turn, when auditory stimuli were more familiar than visual stimuli, young children exhibited auditory dominance. When stimuli of both modalities were unfamiliar, however, children exhibited auditory dominance. This might indicate that auditory dominance in children is the default, whereas visual dominance (occurring only with higher familiarity of visual input) is a result of learning (Napolitano, 2006; Napolitano & Sloutsky, 2004). Nevertheless, the question remains as to why auditory dominance should be the default. According to Napolitano and Sloutsky (2004) and Robinson and Sloutsky (2004), this might reflect an adaptive behavior based on the limited attentional resources during early childhood; auditory input is often more transient than visual input because visual scenes are usually more lasting. Thus, it seems adaptive to first allocate attention to transient (auditory) stimuli before processing more stable input (e.g., visual stimuli). Given this assumption, default auditory dominance may reflect automatic attention to the type of input that is most dynamic and thus, highly fluctuating (Robinson & Sloutsky, 2004). During the course of development, however, attentional resources increase as well as the familiarity with visual stimuli. This might explain, albeit speculatively, why visual dominance gains weight in older children and adults.

Sensory dominance effects under manipulation of semantic congruency and other stimulus characteristics. The second aim of our study was to investigate whether the manipulation of semantic congruency has an influence on the magnitude of sensory dominance effects. Whereas sensory dominance in adults and 6-year-olds was not affected by semantic congruency on a significant level, 9-year-olds were sensitive to this manipulation because visual dominance occurred under semantic congruency only. Thus, semantic congruency (besides spatiotemporal factors) seems to be a modulating factor in 9-year-olds. This is underpinned by the fact that the influence of congruency could be shown for two independent samples of 9-year-olds (Conditions 1 and 2).

The question remains as to why sensory dominance in 9-year-olds is more affected by semantic congruency compared with 6-year-olds and adults. One explanation could be the gradual change of sensory dominance. Whereas auditory dominance might have its peak at around 6 years of age and visual dominance is consolidated in adults, sensory dominance in 9-year-olds is yet in transition and may be more prone to interference by stimulus characteristics. The fact that the absence of color information (Condition 3) did not erase the visual dominance effect in adults but did so in 9-year-olds underlines that visual dominance in this age group is more vulnerable. In line with that, the sensitivity for semantic congruency could reflect the same pattern. Because all age groups exhibited significantly more errors in semantically congruent trials compared with incongruent trials, the presence of a semantic mismatch in incongruent trials seemed to provide participants with an extra cue that a bimodal stimulus had in fact been presented. The manipulation of semantic congruency, however, did not influence the type of error except for the group of 9-year-olds. The 6-year-olds exhibited significantly more auditory errors than visual errors, and adults exhibited significantly more visual errors than auditory errors, in bimodal trials regardless of whether they were presented with semantically congruent or incongruent trials. In contrast, 9-year-olds made significantly more visual errors than auditory errors on semantically congruent trials, but no significant difference, and thereby no sensory dominance, was revealed on semantically incongruent trials. Thus, in all age groups, the presence of a semantic mismatch affected the mean total error rate (taking auditory and visual errors together), but only in the group of 9-year-olds did it have a significant impact on the type of error. Because visual dominance was clearly overridden by the semantic mismatch on incongruent trials in this age group, one could assume that sensory dominance of 9-year-olds is more prone to interference.

However, because there was a clear trend for an interaction between semantic congruency and sensory dominance among 6-year-olds in all experimental conditions, one could assume that there is an effect of congruency and that the lack of significance is due to the underlying sample size.

Concluding remarks. Our results first suggest that sensory dominance in children and adults is not restricted to simple stimuli such as lights and sounds but rather can be extended to semantically meaningful stimuli. This implies a relative robustness of the sensory dominance effects first reported by Nava and Pavani (2012). Second, semantic (in)congruency did not affect the magnitude of the auditory dominance effect in 6-year-olds or of the visual dominance effect in adults, but it was a modulating factor of the visual dominance in 9-year-olds (Conditions 1 and 2). This is a novel finding because it shows that the Colavita effect (Colavita, 1974) can be modulated by factors other than spatial and temporal congruency. Third, the absence of color information (Condition 3) did not affect dominance effects in 6-year-olds and adults, whereas the visual dominance in 9-year-olds disappeared. This suggests that sensory dominance is more robust in 6-year-olds and adults than in 9-year-olds, implying a transitional stage at around this age.

Future perspectives. A longitudinal investigation of sensory dominance would be more appropriate to identify the typical trajectory of multisensory interactions. Furthermore, to solve the question of why multisensory interactions change during the course of development, one might investigate the neural substrates of the Colavita effect (Colavita, 1974) in terms of whether the auditory stimulus is in fact not perceived in the presence of a visual stimulus. Taking older children and adults into account, for example, one could compare the pattern of brain activation in the auditory cortex when the auditory component of a bimodal stimulus is detected versus when it is not (i.e., when the Colavita visual dominance effect occurs). From a clinical point of view, tracing how multisensory interactions typically develop may contribute to a better identification of deviating behavioral patterns and support the development of early diagnostic strategies (Nava & Pavani, 2012). For example, children with autism spectrum disorder (ASD) appear to have impairments in their sensory functioning (cf. Stevenson et al., 2014). Further research examining the nature and extent of processing differences in autistic children compared with non-autistic children, assessing their emergence early in development, and relating these findings to the core deficits in ASD would contribute to a broader understanding of this developmental disorder. It is also conceivable that multisensory integration plays a role in cognitive impairments such as dyscalculia and dyslexia.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14*, 257-262.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20*, 38-64.

Botvinick, M., & Cohen, J. D. (1998). Rubber hand ‘‘feels’’ what eyes see. *Nature, 391*, 756.

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processing*. Cambridge, MA: MIT Press.

Colavita, F. B. (1974). Human sensory dominance. *Perception and Psychophysics, 16*, 409-412.

Farnè, A., Pavani, F., Meneghello, F., & Làdavas, E. (2000). Left tactile extinction following visual stimulation of a rubber hand. *Brain, 12*, 2350-2360.

Gallace, A., & Spence, C. (2005). Visual capture of apparent limb position influences tactile temporal order judgments. *Neuroscience Letters, 379*, 63-68.

Hartcher-O’Brien, J., Gallace, A., Krings, B., Koppen, C., & Spence, C. (2008). When vision ‘‘extinguishes’’ touch in neurologically normal people: extending the colavita visual dominance effect. *Brain Research, 186*, 643-658.

Howard, I. P., Craske, B., & Templeton, W. B. (1966). *Human spatial orientation*. New York: John Wiley.

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the colavita visual dominance effect. *Experimental Brain Research, 184*, 533-546.

Koppen, C., Levitan, C. A., & Spence, C. (2009). A signal detection study of the colavita visual dominance effect. *Experimental Brain Research, 196*, 353-360.

Koppen, C., & Spence, C. (2007). Assessing the role of stimulus probability on the colavita visual dominance effect. *Neuroscience Letters, 418*, 266-271.

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research, 158*, 405-414.

Lewkowicz, D. J. (1988a). Sensory dominance in infants: I. six-month-old infants' response to auditory–visual compounds. *Developmental Psychology*, *24*, 155-171.

Lewkowicz, D. J. (1988b). Sensory dominance in infants: II. ten-month-old infants' response to auditory–visual compounds. *Developmental Psychology, 24*, 172-182.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, 13, 470-478.

Lippé, S., Kovacevic, N., & McIntosh, A. R. (2009). Differential maturation of brain signal complexity in the human auditory and visual system. *Frontiers in Human Neuroscience, 3*. doi: 10.3389/neuro.09.048.2009

Napolitano, A. (2006). *Modality dominance in young children: the underlying mechanisms and broader implications* (Unpublished doctoral dissertation). The Ohio State University, Ohio, USA. Retrieved from https://etd.ohiolink.edu/

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance inyoung children. *Child Development, 75*, 1850-1870.

Nava, E., & Pavani, F. (2012). Changes in sensory dominance during childhood: Converging evidence from the colavita effect and the sound-induced flash illusion. *Child Development, 84*, 604-616.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*, 1387-1401.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: the colavita effect revisited. *Perception and Psychophysics, 69*, 673-686.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*, 822-833.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport*, *12*, 7-10.

Spence, C., Parise, C., & Chen, Y. (2012). The colavita visual dominance effect. In M. Murray & M. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 529-550). Boca Raton, FL: CRC Press.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., et al. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience, 34*, 691-697.

Wurm, L. H., Legge, G. E., Isenberg, L. M., & Luebker, A. (1993). Color improves object recognition in normal and low vision. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 899-911.

CHAPTER 3

Study 2

Sensory dominance across different age groups

Abstract

The aim of this study was to examine sensory dominance across different age groups with a spatial prediction task. Based on the auditory and visual components of bimodal stimuli, participants first learned how to predict the emergence of a target either on the left or right side of the screen (congruent trials). In the test phase, congruent and incongruent trials (i.e., in which the auditory and visual component made contradictory predictions) were presented. In incongruent trials, there was a strong trend towards more auditory-based predictions among 9-year olds, suggesting privileged auditory processing. Among 6-year olds and adults no privileged modality processing could be revealed. Taking the individual level into account, however, adults (regardless of whether they exhibited privileged auditory or visual processing) were able to switch towards their non-privileged modality when being explicitly instructed, i.e., exhibited modality preference. In turn, 6-year olds exhibited dominance. Among 9-year olds a considerable part of children exhibited preference (albeit not on a significant level). This might indicate a transitional stage at around this age, potentially due to increasing attentional resources that improve the ability to encode two modalities simultaneously.

Keywords: Sensory dominance, auditory dominance, perception

Introduction

Multisensory integration as a functional principle of the human brain is central to adaptive behavior because it allows individuals to perceive a world of coherent perceptual entities (cf. Calvert, Spence, & Stein, 2004; Spence & Driver, 2004). Nevertheless, one sensory system can dominate over another when multiple sensory systems are concurrently stimulated. In adults, visual dominance seems to be prevalent (e.g., Spence, Parise, & Chen, 2012).

A prominent example for visual dominance over audition is the Colavita effect (Colavita, 1974): Adults usually have no difficulty in detecting unimodal auditory or visual stimuli. When presented with bimodal stimuli, consisting of an auditory and a visual component, however, adults often perceive a single visual stimulus while being unaware of the auditory component. Many studies were conducted to further examine this phenomenon, and the apparent prepotency of the visual over the auditory system was confirmed across many experimental conditions (for a review see Spence et al., 2012).

However, visual dominance does not seem to be innate, as there is evidence that young infants' information processing is dominated by the auditory system (e.g., Lewkowicz, 1988a, 1988b). Furthermore, there is evidence that auditory dominance persists until early childhood. Robinson and Sloutsky (2004) conducted a spatial prediction task with 4-year olds and adults. Participants were presented with two different bimodal stimuli, each consisting of an auditory and a visual component (AUD_1VIS_1 and AUD_2VIS_2). Participants learned that each bimodal stimulus predicted a certain target event (i.e., congruent trials). Later, new stimulus combinations were occasionally presented, in which auditory and visual components were switched (AUD_1VIS_2 and AUD_2VIS_1), predicting opposing target events (i.e., incongruent trials). If participants relied more on visual information, they should make predictions based on the visual rather than the auditory component in incongruent trials. Conversely, if they relied more on auditory information, predictions in incongruent trials should base on the auditory component. While adults made more visual-based predictions, 4-year olds primarily made auditory-based predictions, suggesting an auditory dominance. Overall, research indicates that for infants and 4-year olds auditory input dominates visual input (Sloutsky & Napolitano, 2003; Thompson & Massaro, 1994), although auditory dominance appears to be less robust in 4-year olds, depending on visual stimulus characteristics such as stimulus familiarity (Robinson & Sloutsky, 2004). According to Napolitano and Sloutsky (2004) and Robinson and Sloutsky (2004) auditory dominance in early childhood is the default that reflects an adaptive process based on limited attentional resources. As auditory input is often more transient than visual input from scenes and objects, it seems reasonable to allocate attention first to dynamic (auditory) stimuli before processing more stable (visual) stimuli. Therefore, default auditory dominance might reflect automatic attention to the least stable type of input (Napolitano, 2006; Robinson & Sloutsky, 2004). However, with increasing attentional resources during development, older children and adults might gain more capacities for processing visual input, too.

In fact, Nava and Pavani (2012), examining the developmental trajectory of sensory dominance with a Colavita paradigm, could show a transition from auditory dominance in 6- to 7-year olds towards visual dominance in 9- to 10-year olds. Children were presented with simple lights and sounds, either separately or simultaneously, and were asked whether they had perceived a single auditory, visual, or a bimodal stimulus. In the bimodal condition, 9- to 10-year olds, and adults often missed the auditory component, and responded only to the visual component (Colavita visual dominance effect), whereas 6-year olds often missed the visual component suggesting auditory dominance. However, the transition of sensory dominance from auditory dominance in 6- to 7-year olds to adult-like visual dominance in 9- to 10-year olds might not be as straight forward as proposed by Nava and Pavani (2012). Gori, Sandini, and Burr (2012) tested several age groups (i.e., 5- to 12-year olds, and adults) for audio-visual integration in space and time. All age groups exhibited visual dominance in the spatial task, and auditory dominance in the temporal task. These findings are in line with the Modality Appropriateness Hypothesis by Welch and Warren (1986) stating that the modality being most appropriate with respect to a given task dominates processing. Thus, in spatial tasks the visual modality dominates the auditory one, because it is more precise at determining spatial information. For temporal judgments, in turn, the auditory modality dominates over vision (cf. Shams, Kamitani, & Shimojo, 2004).

Summarized, there seems to be a default auditory dominance up to 4 years of age, which was seen to persist even in spatial tasks (Robinson & Sloutsky, 2004), while sensory dominance in children older than 4 years of age starts to differentiate towards auditory dominance in the temporal domain and visual dominance in the spatial domain.

The aim of the current study was to examine sensory dominance in children (6- and 9-year olds) and adults in a spatial prediction task. In Experiment 1, we expected visual dominance in adults, and, in line with Gori et al. (2012), visual dominance in children, too. Not only the response patterns across each age group were considered, but also individual patterns of responses (cf. Robinson & Sloutsky, 2004). Moreover, reaction times were assessed. It was assumed that participants would react faster in congruent than in incongruent trials as the latter cause a mismatch between what was learned during training and the new stimulus combinations. In Experiment 2, we examined whether children and adults exhibited sensory dominance (i.e., disrupted processing of a modality) or sensory preference (i.e., deliberately selecting a modality that will be attended to while rejecting the other, cf. Napolitano & Sloutsky, 2004). For this purpose, Experiment 1 was rerun but participants in Experiment 2 were explicitly tested on their non-privileged modality.

Experiment 1

Method

Sample. Participants were 23 6-year olds (*M* = 6.04 years, *SD* = 0.77; 10 males), 19 9-year olds (*M* = 9.05 years, *SD* = 0.85; 7 males), and 25 adults (*M* = 22.68 years, *SD* = 2.59; 9 males). Children were recruited from local schools and nurseries, after their parents had signed a consent form. Adults were students and unfamiliar with the aim of the experiment. All participants were native speakers and had normal or corrected-to normal vision and normal hearing by self-report.

Stimuli. Stimulus programming, presentation, and response collection were carried out with the program Inquisit 3.0.1.1 (Inquisit, 2008). Visual stimuli consisted of a grey and a black circle, each with a diameter of 7.5 cm, displayed at the center of a white-background computer monitor (touchscreen) measuring 42.6 cm x 32.7 cm.
Auditory stimuli consisted of a 250 and a 500 Hz tone, presented at 50 dB via headphone.

Stimuli discrimination. To ensure that participants were able to distinguish between the stimuli of each modality, a prior study was run consisting of two parts presented in random order. In one part, participants were randomly presented with 20 trials, consisting of ten black circles and ten grey circles. Participants had to decide whether they had seen a grey or a black circle by pressing the according button on the keyboard (i.e., a grey or black button). Adults were correct to 99.0 %, 6-year olds to 94.5 %, and 9-year olds to 98.9 %.

In the other part, participants were presented with 20 trials in random order, ten consisting of a 250 Hz tone and ten of a 500 Hz tone. Participants decided whether they had heard a low-pitched tone (250 Hz) or a high-pitched tone (500 Hz) by pressing the according button on the keyboard (i.e., the button marked with an arrow pointing downwards for the low-pitched tone or upwards for the high-pitched tone). Adults were correct to 98.6 %, 6-year olds to 90.2 %, and 9-year olds 91.3 %. Thus, participants of all age groups were sufficiently able to distinguish between the stimuli of each modality.

Thereafter, auditory and visual stimuli were matched and simultaneously presented as bimodal stimuli. Some participants were presented with the bimodal stimulus AUD_1VIS_1 consisting of the grey circle and the 500 Hz tone, and the bimodal stimulus AUD_2VIS_2 consisting of the black circle and the 250 Hz tone, while the others were presented with the bimodal stimulus AUD_1VIS_1 consisting of the light grey circle and the 250 Hz tone, and AUD_2VIS_2 consisting of the black circle and the 500 Hz tone. The procedure will be illustrated by the latter condition.

Procedure. All participants were tested individually. The procedure was similar to the experiment of Robinson and Sloutsky (2004), consisting of three phases: familiarization, training, and test phase. During familiarization, participants were presented with four trials. In each trial, one of the two bimodal stimuli (AUD_1VIS_1 or AUD_2VIS_2) was presented for 1000 ms, followed by two empty panels (i.e., 12.5 cm x 9 cm rectangles), simultaneously appearing on the left and right side of the touchscreen. The panels were used to mark the location where a cartoon animal, Henry, could appear. Participants were told that the stimuli would predict different target events: AUD_1VIS_1 indicates that Henry would pop up in the left panel, while AUD_2VIS_2 indicates that Henry would pop up in the right panel, which happened 2000 ms after the appearance of the panels.

During the training phase, participants were taught to make predictions based on the bimodal stimuli (see Figure 1). Adults were presented with four, children with eight trials, as we assumed that children would need more trials to learn the association between the stimuli and spatial events. After being presented with a bimodal stimulus for 1000 ms, participants were asked where Henry would appear and to respond by touching the corresponding panel on the screen. Thereafter, feedback was provided (e.g., Henry appeared in the right or left panel). Children also received verbal feedback by the experimenter, who said “Good job!” for correct predictions, and “Oops, that was not right!” for incorrect predictions. Feedback was provided automatically if no response was given within 3000 ms. In order to proceed with the test phase, children had to respond correctly within 3000 ms in each of the last four training trials, and adults in the last two trials. All participants reached the criterion.

The test phase was similar to the training phase with the following exceptions. First, no feedback was provided. Second, panels were presented simultaneously with bimodal stimuli so that participants could respond immediately. Third, there was no time restriction for the response. Fourth, not only congruent but also incongruent stimuli were provided. In total, 60 experimental trials were presented in random order, thereof 44 trials consisting of “congruent” stimuli (i.e., the same bimodal stimuli as presented during the training phase), and 16 “incongruent” stimuli (see Figure 2). In incongruent stimuli auditory and visual components were switched (i.e., the 250 Hz tone was now presented with the black circle, and the 500 Hz tone with the grey circle). Thus, incongruent stimuli made contradicting predictions. If participants relied stronger on visual information in the course of learning, they would make predictions based on the visual components in incongruent trials, whereas the effect should be reversed if they relied stronger on auditory information. Participants were asked to respond as fast as possible. Whenever participants became aware of the contradiction inherent in the incongruent trials and asked what to do, they were told to carry on by deciding between the visual or auditory component in order to make a prediction. However, this occurred only for a few participants.

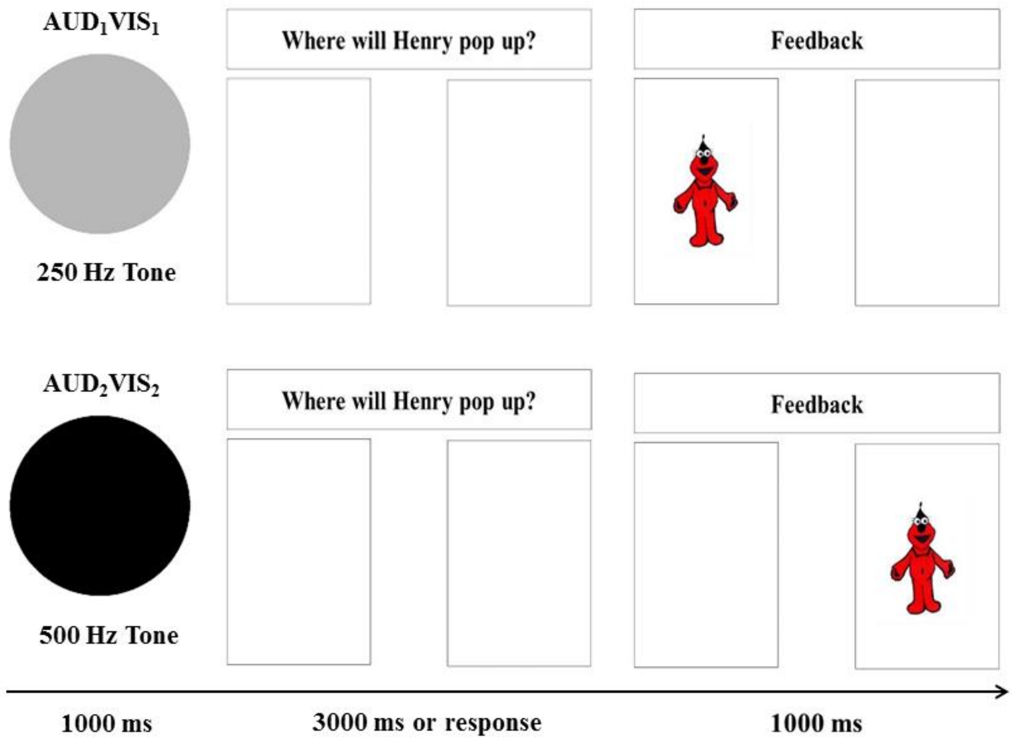

Figure 1. Procedure of the training phase.

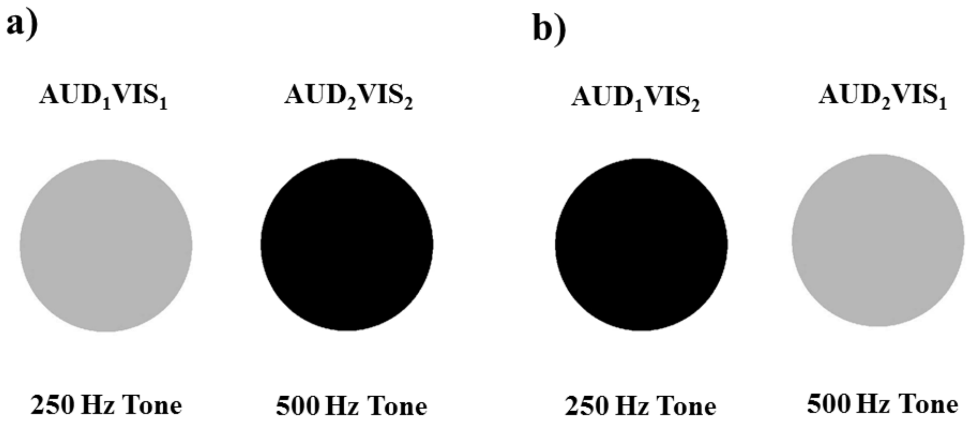

Figure 2. Bimodal stimuli presented in the test phase: (a) congruent trials, identical to training trials, (b) incongruent trials, in which the audio-visual components were switched.

Results

Participants who made more than 15% errors in congruent trials were excluded from the analysis, suggesting that they had not appropriately learned the bimodal stimulus combinations in the preceding training session. This was the case for five 6-year olds (21.7 %).

As the study was conducted by means of a touch screen, sometimes responses were not recognized by the computer if the touch was too soft. Thus, the response had to be repeated, which in turn led to the recording of longer reaction times. Therefore, trials with reaction times longer than 7000 ms were excluded from the analysis (i.e., < 1% of the trials in each age group).

Analysis of auditory- and visual-based predictions in incongruent trials

To test for sensory dominance on group level, that is, whether participants made more visual- or more auditory-based predictions in incongruent trials, one-sample *t*-tests were conducted for each age group separately (adjusted alpha level set at .016, test value 50 %). Adults made 61.50 % visual-based versus 38.50 % auditory-based predictions, $t(24) = -1.46$, $p = .16$. Six-year olds made 63.13 % auditory- versus 36.87 % visual-based predictions. $t(17) = 1.70$, $p = .11$, similar to 9-year olds with 69.41 % auditory- and 30.59 % visual-based predictions, $t(18) = 2.12$, $p = .048$ (not significant after correction of the alpha level). A one-way ANOVA confirmed that the percentage of auditory-based predictions significantly differed between age groups, $F(2, 59) = 4.19$, $p = .020$, $\eta^2 = .12$. Post-hoc tests (Games-Howell) indicated no significant differences between 6- and 9-year olds, $p > .05$, but significantly less auditory responses among adults compared to 9-year olds, $p = .038$, and, as a tendency, also compared to 6-year olds, $p = .078$.

Additional analyses were conducted to determine individual patterns of responses (cf. Robinson & Sloutsky, 2004) This was considered as the accumulation of responses on the group level might yield biased results that may not reflect individual performance. Participants who responded in favor of the auditory stimulus on at least 12 of 16 incongruent trials (e.g., 75 %) were classified as auditory responders, participants who responded in favor of the visual stimulus on at least 12 of 16 incongruent trials were classified as visual responders, and the remaining participants were classified as mixed responders (see Figure 3).

The percentages of auditory, visual, and mixed responders differed between age groups on a marginally significant level, $\chi^2(4, n = 62) = 9.50$, $p = .050$. Next, each age group was analyzed separately (adjusted alpha level set at .008). Among 6-year olds, no significant differences were revealed between mixed, auditory, and visual responders, $\chi^2(2, n = 18) = 4.33$, $p = .12$. No mixed responders were revealed among 9-year olds, percentages of auditory and visual responders did not significantly differ, $\chi^2(1, n = 19) = 2.58$, $p = .11$. Significant differences, however, were revealed among adults, $\chi^2(2, n = 25) = 10.16$, $p = .006$, with more visual than mixed responders, $\chi^2(1, n = 17) = 9.94$, $p = .002$, whereas the percentages of auditory and mixed responders, as well as visual and auditory responders did not significantly differ, ps > .008.

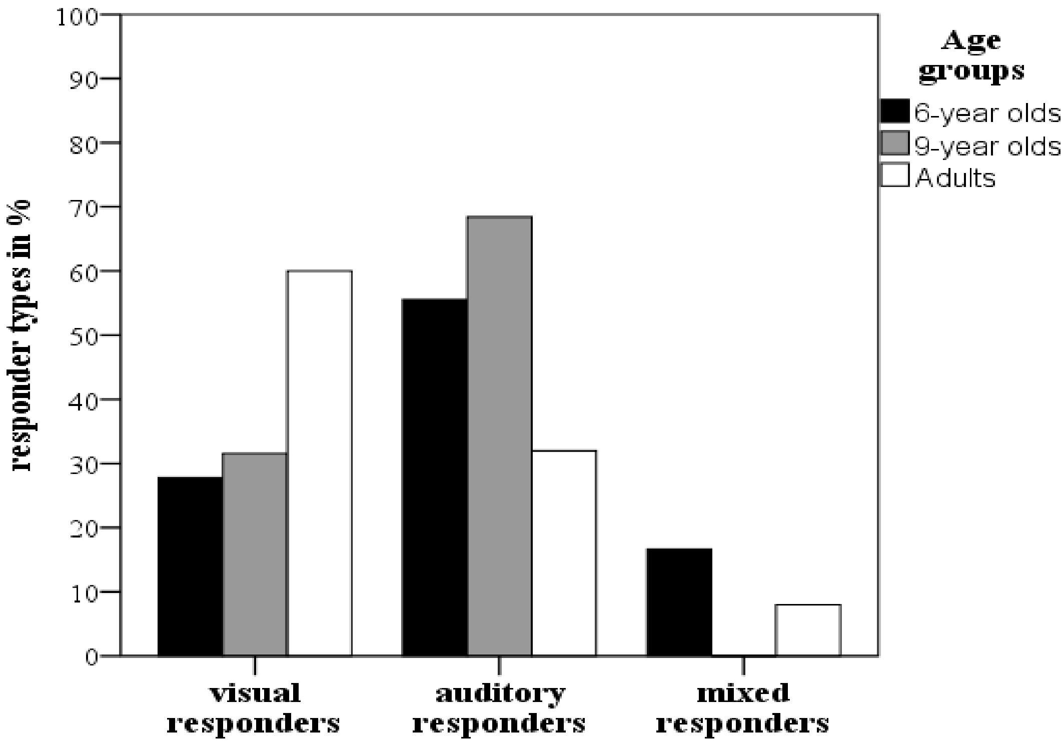

Figure 3. Responder types (in %) for each age group. *Note*: No mixed responders were revealed among 9-year olds.

Analysis of reaction times

Mean reaction times in congruent and incongruent trials were entered into a repeated measured ANOVA with age group as between-subjects variable, and congruency (congruent versus incongruent trials) as within-subjects variable. Main effects of congruency, $F(1, 59) = 10.73$, $p = .002$, $\eta^2 = .15$, and of age group were revealed, $F(2, 59) = 25.77$, $p < .001$, $\eta^2 = .47$. Furthermore, there was an interaction between congruency and age group, $F(2, 59) = 5.08$, $p = .009$, $\eta^2 = .15$. Post-hoc tests (Tukey) indicated shorter reaction times of 9- compared to 6-year olds, $p = .015$, and shorter reaction times of adults compared to both groups of children, *p*s < .001.

To specify the interaction between congruency and age group, repeated measures ANOVAs were conducted separately for each age group (adjusted alpha level set at .016). Among 6- and 9-year olds, no significant differences in reaction times between congruent and incongruent trials were revealed, *p*s > .05. However, adults had shorter reaction times in congruent trials compared to incongruent trials, $F(1, 24) = 4.27$, $p < .001$**.**

Experiment 2

Experiment 2 was run with those participants of Experiment 1 who relied on the same modality in the majority of incongruent trials (i.e., auditory and visual responders). The question was whether these participants exhibited sensory dominance (i.e., impaired encoding of the non-privileged modality due to the strong automatic pull of attention by the privileged modality) or preference (i.e., deliberately selecting a modality that will be attended to and rejecting the other, cf. Napolitano & Sloutsky, 2004). Thus, auditory and visual responders of Experiment 1 were tested on their non-privileged modality in Experiment 2 by asking them to rely in all trials on their non-privileged modality. The reliable use of the non-privileged modality might indicate sensory preference. That is, participants encoded both modalities during familiarization (i.e., the auditory as well as visual components of bimodal stimuli), usually rely on their privileged modality in the test phase (see Exp. 1), but are able to respond in favor of the non-privileged modality, if being explicitly told. Contrary, failing to use the non-privileged modality in the test phase of Experiment 2 might indicate that the privileged modality disrupted the encoding of non-privileged modality input already in the familiarization phase, yielding in sensory dominance.

Sample. Participants were the same as in Experiment 1, except for participants who could not be recruited again and mixed responders who did not show privileged processing of any modality. Hence, the sample of Experiment 2 contained 14 6-year olds (M = 6.3; SD = 7.3, 3 males), 17 9-year olds (M = 9.0, SD = 8.7, 5 males) and 23 adults (M = 22.9; SD = 2.5, 9 males).

Procedure. After a short break following the first experiment, participants were told that they would be presented with an experiment similar to the previous one. They then were instructed to focus on both modalities (i.e., auditory and visual components of the bimodal stimuli), and presented with four familiarization trials, which were identical to Experiment 1. In the test phase, visual responders of Experiment 1 were asked only to respond based on the auditory stimuli and to ignore the visual stimuli. Auditory responders of Experiment 1, in contrast, were asked only to respond based on the visual stimuli and to ignore the auditory stimuli. The test phase was identical to Experiment 1 including 44 congruent and 16 incongruent trials.

Results

Each participant was either labeled with sensory dominance or preference, depending on whether he or she was able to use the non-privileged modality in incongruent trials. Visual (or auditory) responders of Experiment 1 who remained visual (or auditory) responders in Experiment 2 although being explicitly told to respond based on their non-privileged modality were labeled with “sensory dominance”. Those who were able to switch from their privileged modality in Experiment 1 to the non-privileged modality in Experiment 2 were labeled with “sensory preference”. The percentages of participants labeled with dominance or preference significantly differed between age groups, $\chi^2(2, n = 54) = 27.00, p < .001$ (see Figure 4). Next, each age group was analyzed separately (adjusted alpha level set at .016). Among adults, there were more subjects labeled with preference than dominance $\chi^2(1, n = 23) = 19.17, p < .001$. No significant differences were revealed in the group of 9-year olds $\chi^2(1, n = 17) = 2.89, p = .090$, whereas among 6-year olds there were more participants labeled with dominance than preference, $\chi^2(1, n = 14) = 7.14, p = .008$.

To specify whether auditory and visual responders differed in dominance or preference, an additional Chi-square analysis was run. Among auditory and visual responders, there were similar percentages of participants showing sensory dominance or preference, $\chi^2(1, n = 54) = .59, p = .57$.

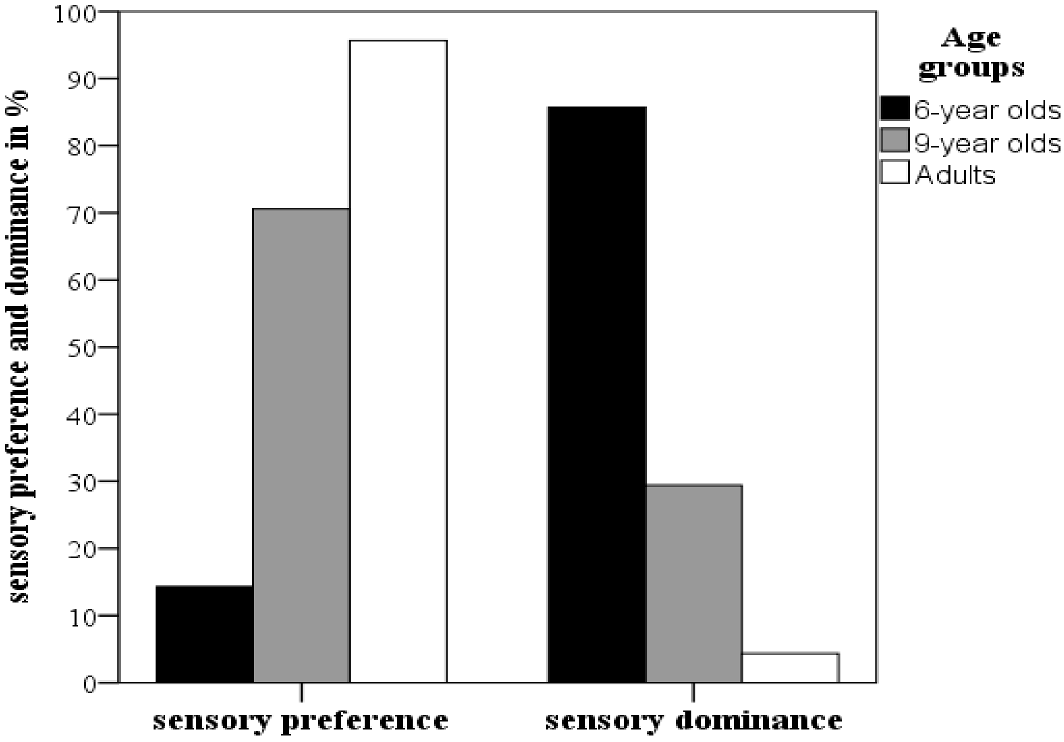

Figure 4. Participants labeled with sensory preference or dominance (in %) for each age group.

General Discussion

The aim of this study was to examine sensory dominance across three different age groups, including children older than four years and adults. In addition to analyses on the level of age groups, individual patterns of responses were taken into account in Experiment 1. Furthermore, differences in reaction times between congruent and incongruent trials were analyzed. In Experiment 2, we were interested whether children and adults exhibited sensory dominance or sensory preference, which includes the ability to enocode both sensory modalities.

Sensory dominance across different age groups

Group level. No sensory dominance was revealed among 6-year olds and adults. However, 9-year olds exhibited a trend towards privileged auditory processing. This is not in line with the study of Gori et al. (2012) demonstrating auditory dominance in a temporal task and visual dominance in a spatial task among children and adults. However, whereas the findings of Gori et al. (2012) supported the Modality Appropriateness Hypothesis (Welch &

Warren, 1986), according to which the modality being most appropriate with respect to a given task dominates processing (i.e., the visual modality dominates in spatial tasks, the auditory modality in temporal tasks), one has to take into account that this approach has limitations. That is, visual dominance is not restricted to spatial tasks (e.g., fusion and visual dominance illusions of the McGurk effect; Eskelund, MacDonald, & Andersen, 2004), and auditory dominance was revealed in spatial tasks, too. For example, Alais and Burr (2004) could show a reversed Ventriloquist effect in terms of auditory dominance over the visual modality, depending on visual stimulus reliability. When the visual stimuli were small and well defined, participants relied on visual rather than auditory stimuli to judge a location. In turn, when visual stimuli were large and fuzzy, participants were more likely to make their judgements based on auditory stimuli (Alais & Burr, 2004; Nelson, 2004; Witten & Knudsen, 2005). This was the case because large, fuzzy visual stimuli were perceived as being less reliable than small, well defined ones when judging a location (Nelson, 2004). Summarized, the study of Alais and Burr (2004) indicates that vision does not dominate audition in spatial processing because of any physiological advantage (as stated by the Modality Appropriateness Hypotheses, Welch and Warren, 1986), but depending on stimulus reliability in a given task (Witten & Knudsen, 2005). It thus might be possible that a replication of our study under variation of stimulus characteristics such as color information or width in visual stimuli and frequency in auditory stimuli elicits different outcomes such as privileged visual processing in 9-year olds.

Referring to the study of Nava and Pavani (2012), however, auditory dominance was reported in 6-year olds, a change towards visual dominance in 9-year olds, and a robust visual dominance effect in 12-year olds and adults. Note that 6-year olds in the current study exhibited 63 % auditory responses versus 37 % visual responses in incongruent trials on the group level, and that there were twice as many auditory responders than visual responders on the individual level. Thus, albeit not on a significant level, these results might support auditory dominance among 6-year olds as reported by Nava and Pavani (2012). However, whereas 9-year olds in the current study exhibited a trend towards auditory dominance, visual dominance has been reported for this age group by Nava and Pavani (2012). One explanation for these contradicting findings might be that visual dominance in 9-year olds is yet in transition and, thus, not as robust. This was confirmed by a study examining the influence of semantic congruency on sensory dominance (Wille & Ebersbach, 2016). Here, adults and 9-year olds exhibited visual dominance when colored pictures were presented together with meaningful tones (e.g., a colored picture of a dog presented together with a barking sound). However, with the absence of color information, visual dominance in 9-year olds was extinguished, whereas visual dominance in adults was hardly affected. This might be an indicator that visual dominance in 9-year olds is less marked yet and more prone to interference due to stimulus characteristics, such as color information, than visual dominance in adults. Thus, as for 9-year olds it might be possible that a replication of the current study with colored instead of monochrome visual stimuli elicits visual instead of auditory dominance.

Individual level. No significant differences in the percentages of visual, auditory, and mixed responders were revealed among 6- and 9-year olds. However, as there were twice as many auditory than visual responders among children, it might be possible that the lack of significance reflects a lack of power. Among adults no significant differences were revealed between visual and auditory responders, though there were more than twice as many visual than auditory responders.

Reaction times

It was assumed that participants would react faster in congruent than in incongruent trials as the latter cause a mismatch between what was learned during training and the new stimulus combinations. This effect emerged only for adults. Interestingly, when participants were non-systematically asked whether they noticed something in particular during the experiment, most of the adults said that they took notice of the incongruent trials, whereas hardly any child was aware of them. In line with that, no differences in reaction times were revealed as plenty of children obviously did not perceive a mismatch between what was learned and the new stimulus combinations.

Sensory dominance or preference

In the second experiment, participants were tested on their non-privileged modality to examine whether they exhibit sensory dominance or preference. A similar experiment was conducted by Robinson and Sloutsky (2004, Experiment 3) with 4-year olds, who exhibited sensory dominance. The current study demonstrated dominance in 6-year olds, too. Among 9-year olds, however, there were comparable percentages of children exhibiting preference or dominance. This in turn might indicate a transitional stage: Given a considerable part of 9-year olds exhibiting adult-like preference instead of dominance, one might speculate whether this is due to increasing attentional resources in some of the children enabling them to encode two modalities simultaneously. Finally, there were significantly more adults exhibiting preference than dominance. This is in line with adults' detection of incongruent trials and a slowdown in reaction times on incongruent trials as it indicates that indeed adults encoded both modalities but deliberately chose the visual modality over the auditory to attend to.

However, as for 6-year olds and some of the 9-year olds exhibiting sensory dominance, one could argue that failure to use the non-privileged modality in the test phase does not necessarily mean that input of the non-privileged modality was not encoded during familiarization. Rather, the automatic pull of attention by the privileged modality might have prevailed over deliberate attention towards the non-privileged modality in the test phase. Thus, rather than a lack of encoding of non-privileged modality input, maybe children simply were not able to inhibit the privileged modality in order to respond in favor of the non-privileged modality during the test phase. However, this requires further research.

Conclusions

On the group level there was a trend towards auditory dominance among 9-year olds, whereas no sensory dominance was revealed among 6-year olds and adults. Taking the individual level into account, however, adults (whether or not they were visual or auditory responders) were able to switch towards their non-privileged modality when being explicitly instructed, implying sensory preference rather than dominance. In turn, 6-year olds (regardless of responder type) exhibited dominance. Albeit not on a significant level, a considerable part of 9-year olds exhibited (adult-like) preference. This might indicate increased attentional resources in some children of this age group that improve the simultaneous encoding of two modalities.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14*, 257-262.

Calvert, G. A., Spence, C., & Stein, B. E. (Eds.) (2004). *The handbook of multisensory processing*. Cambridge, MA: MIT Press.

Colavita, F. B. (1974). Human sensory dominance. *Perception and Psychophysics, 16*, 409-412.

Eskelund, K., MacDonald, E. N., & Andersen, T. S. (2015). Face configuration affects speech perception: evidence from a McGurk mismatch negativity study. *Neuropsychologia, 66*, 48-54.

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience, 6*, 28-35.

Inquisit 3.0.1.1. [Computer software] (2012). Seattle, WA: Millisecond Software LLC.

Lewkowicz, D. J. (1988a). Sensory dominance in infants: I. six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155–171.

Lewkowicz, D. J. (1988b). Sensory dominance in infants: II. ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 172–182.

Massaro, D. W. (1984). Children’s perception of visual and auditory speech. *Child Development, 55*, 1777–1788.

Massaro, D. W., Thompson, L. A., Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. *Journal of Experimental Child Psychology, 41*, 93–113.

Napolitano, A. (2006). *Modality dominance in young children: The underlying mechanisms and broader implications* (Unpublished doctoral dissertation). The Ohio State University, Ohio, USA. Retrieved from https://etd.ohiolink.edu/

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*, 1850-1870.

Nava, E., & Pavani, F. (2012). Changes in sensory dominance during childhood: converging evidence from the colavita effect and the sound-induced flash illusion. *Child Development, 84*, 604-116.

Nelson, L. (2004). Reverse ventriloquism. *Nature News.* Retrieved from: http://www.nature.com/news/2004/040203/full/news040202-3.html

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*, 1387-1401.

Shams, L., Kamitani, Y., & Shimojo, S. (2004). Modulations of visual perception by sound. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes,* (pp. 27-34). Cambridge, MA: MIT Press.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833.

Spence, C., & Driver, J. (Eds.) (2004). *Crossmodal space and crossmodal attention*. Oxford: Oxford University Press.

Spence, C., Parise, C., & Chen, Y. (2012). The colavita visual dominance effect. In M. Murray & M. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 529-550). Boca Raton: CRC Press.

Thompson, L. A., & Massaro, D. W. (1994). Children's integration of speech and pointing gestures in comprehension. *Journal of Experimental Child Psychology, 57*, 327 – 354.

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In: K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 1-36). New York: Wiley.

Wille, C. & Ebersbach, M. (2016). Semantic congruency and the (reversed) colavita effect in children and adults. *Journal of Experimental Child Psychology, 141*, 23-33.

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron, 48,* 489-496.

CHAPTER 4

Study 3

Verbal facilitation effects instead of verbal overshadowing in

face memory of 4- to 6-year olds

Abstract

Research on eye witness memory in older children and adults revealed that verbally describing unfamiliar faces impairs later recognition of these faces, known as the “verbal overshadowing effect”. The aim of this study was to investigate whether a verbal overshadowing effect occurs in 4- to 6-year olds, too, and whether visualization (i.e., drawing the seen face) might elicit a visual overshadowing effect. Instead of a verbal overshadowing effect, a verbal facilitation effect was revealed with verbal intelligence being a significant predictor for recognition accuracy in the verbalization group but not in the control group. No effect of visualization was observed on recognition accuracy. Potential explanations for the results are discussed.

Keywords: Face memory, verbal overshadowing, visual overshadowing

Introduction

Research on eye witness memory in adults identified plenty of factors impairing the accuracy of face recognition (e.g., Sporer, Malpass, & Koehnken, 1996). A well-known phenomenon is, for instance, the “verbal overshadowing effect.” Witnesses who verbally described a face and then have to identify it in a recognition task are more prone to make errors than witnesses who did not make a verbal description in advance. This phenomenon was initially reported by Schooler and Engstler-Schooler (1990). In the study, adults were presented with a video of a bank robbery, followed by an unrelated filler task. After 20 min, subjects of the experimental group were asked to give a detailed description of the robber’s face, while the control group continued with another filler task. All subjects were then presented with photographs of eight similar faces including the robber’s face and were asked to identify the latter. In fact, the experimental group was found to be less able to recognize the robber compared to the control group. Over the years, a vast number of studies have been conducted to further examine this phenomenon. For example, a recent study by Alogna et al. (2014) consisting of a large-scale, multiple-laboratory replication of the original experiments of Schooler and Engstler-Schooler (1990, Experiment 1 and 4) revealed a robust verbal overshadowing effect in adults, whereas Brown and Lloyd-Jones (2005) reported a verbal facilitation effect (i.e., higher recognition accuracy after verbalization) rather than a verbal overshadowing effect. A meta-analysis by Meissner and Brigham (2001) indicated a small, but reliable negative effect of verbalization on face memory, which was, however, questioned by Francis (2012), who suggested that this effect was still overestimated due to a potential publication bias. Thus, the verbal overshadowing effect - and in particular its underlying mechanisms - remains subject of controversial debate.

Currently, three explanatory approaches for this effect are discussed (Chin & Schooler, 2008). The processing account (Schooler, 2002) refers to a shift in processing caused by verbalization; that is, verbalization triggers the use of feature-based strategy instead of holistic processing, which impedes face recognition. The criterion account (Clare & Lewandowsky, 2004) suggests that verbalization causes a conservative response bias, making the witness less likely to believe that the target is present in the line-up. Finally, the content account (Schooler & Engstler-Schooler, 1990) proposes that individuals whose level of perception exceeds their level of verbal abilities should be especially prone to verbal overshadowing. These individuals would be able to perceive various aspects of a face, but unable to appropriately put these perceptions into words. Accordingly, the detrimental effect of verbalization occurs because people rely on the impoverished verbal details of their memory at the expense of their superior perceptual memories. It is noteworthy that there is evidence for all three approaches (Chin & Schooler, 2008).

This study concentrates on the content account, as it holds the most interesting implications for preschool children due to their still limited level of verbal abilities. So far, research on the influence of verbal face description on face memory has focused mainly on adults. Among the few studies that examined children, Memon and Rose (2002) found no verbal overshadowing effect in 8- and 9-year olds. Children witnessed a live event in their classroom: A stranger entered the room, presented a picture of a dog, asked the children whether they had seen this dog, and left. Twenty-four hours later, children were assigned to a verbalization and non-verbalization group and afterwards had to complete a face recognition task in order to identify the man. No significant difference in accuracy was revealed between children in the verbalization and non-verbalization group. However, this study contains some methodological shortcomings, that is a 24-h delay between the event and verbalization, which might have been too long to elicit a verbal overshadowing effect (cf. Meissner & Brigham, 2001). Moreover, some children might have been more interested in the picture of the dog rather than the stranger's face which might have caused individual differences in the time attending and processing the stranger's face. Dehon, Vanootighem, and Brédart (2013) conducted another study, not only covering different age groups (7- to 8-year olds, 9- to 10-year olds, 13- to 14-year olds), but also taking the shortcomings of Memon and Roses' study (2002) into account. By presenting a video of a person located in the centre of the screen without any other interfering characters, such as another person or a dog, a verbal overshadowing effect was revealed for all age groups, irrespectively of whether children verbalized immediately after the presentation or delayed by 24 h, and also irrespectively of whether or not the recognition task was performed immediately after the verbalization or 24 h later. These results indicate that the verbal overshadowing effect is not restricted to adults, but occurs in children, too. However, no research on verbal overshadowing of face memory has been conducted so far with children younger than seven years old. Given that also preschool children become witnesses of criminal events, the question arises whether their memory is affected by verbalization, too.

The aim of this study was to investigate (a) whether the verbal overshadowing effect also occurs in 4- to 6-year olds, (b) whether the content account (Schooler & Engstler-Schooler, 1990) can make an explanatory contribution of this effect in young children and (c) whether visualization (i.e., drawing the seen face before recognizing it) affects their face memory.

First, we expected in line with the content account (Schooler & Engstler-Schooler, 1990) that young children should be especially prone to verbal overshadowing: After verbalization, children should be less accurate in face recognition than children in the non-verbalization group as young children's vocabulary is still limited compared to that of older children and adults (cf. Brandone, Salkind, Golinkoff, & Hirsh-Pasek, 2006). Presented with a face, 4- to 6-year olds should be able to perceive various aspects visually, but might have difficulties putting all these aspects into appropriate words. Thus, subsequent verbalization would not fully match the perceptual experience and result in poor verbal representations which overshadow the original visual representations (Schooler & Engstler-Schooler, 1990). Hence, children would judge based on the impoverished details of the verbal representations which should impair face recognition.

Second, we accessed individual differences in verbal abilities. According to the content account, subjects with poorer verbal abilities should suffer more from verbalization (i.e., exhibit lower accuracy in face recognition) than subjects with higher verbal abilities (Ryan & Schooler, 1998). Moreover, if there was a relationship between verbal abilities and accuracy in the non-verbalizing control group, too, one might assume that it is not verbal intelligence in particular but rather general intelligence (often correlated with verbal abilities), which contributes to better recognition performance.

Third, we examined whether face memory of 4- to 6-year olds might be affected not only by verbalization but also by visualization, that is drawing the face from memory prior to the recognition phase. Analogously to the content account focusing on verbal abilities, it might be possible that the perceptual abilities of children also exceed their drawing abilities, and thus give rise to a "visual overshadowing effect."

Method

Sample

Participants were 56 preschool children aged 4–6 years, recruited from local daycare centres after their parents signed a consent form. Children were randomly assigned to one of the three experimental conditions: the verbalization group (18 children, mean age: M = 5.1 years, SD = .8; 7 males, 11 females), visualization group (18 children, mean age: M = 5.0 years, SD = .8; 9 males, 9 females), and control group (20 children, mean age: M = 5.3 years, SD = .7; 11 males, 9 females). All children were native speakers, had normal or corrected-to normal vision and took part voluntarily.

Stimuli and procedure

There were four tasks, always presented in the same order.[2] Each task included the presentation of a video of a robbery on a laptop (screen size 15.6 inch) by means of headphones, which was followed by the experimental intervention and then by a recognition test, in which the face of the robber had to be identified out of five photographs of faces, each measuring 16 × 12 cm, including the target and four distractors. To ensure similarity between target faces and distractor faces, the photographs were rated in advance by ten students and the four most similar faces to the target were chosen. Children were tested individually in a separate room of their institutions. The experiment was introduced as a game in which children would play the role of a detective who has to hunt robbers. Each task started with a scene showing a person who presented her favourite object (e.g., a teddy bear). The person told that her favourite object had been stolen, and even though the object was back, unfortunately the robber has not been caught yet. Children were asked for help by memorizing the face of the robber in the upcoming scene. The next scene showed a person in neutral black clothing, taking away the favourite object from a table. At the end of the robbery scene, the camera focused on the still face of the robber for 5 s. Subsequently, children completed a filler task to generate a delay of 3.5 min that included doing handicrafts by richly decorating a certificate with stickers and paintings. Thereafter, the experimental intervention followed. Children of the verbalization group had to verbally describe the robber's face for about 1.5 min. To facilitate this, the experimenter used eight open-ended questions that were posed to the child asking for details and properties of the face (e.g., "Was there something in the robber's face that you noticed in particular?", see Appendix A). Children of the visualization group were asked to draw the robber's face for about 1.5 min, whereas children of the control group continued

[2] Except for the control group which was presented with tasks in random order, however, this did not influence performance as proactive interference could be ruled out in all experimental groups (see results).

with the filler task for another 1.5 min. Thereafter, the face recognition test was conducted. Children saw five monochrome portrait photographs placed on a table in front of them (two at the top and three at the bottom line). The position of the target face was randomized in each task for each child. Each portrait was printed in the centre of a horizontal sheet, showing a frontal face with features like part of the hair and ears. The set of photographs included a picture of the robber as well as distractor faces. Children were asked whether they could recognize the robber and to point on the appropriate picture. After their choice, children were asked once whether this was their final decision, which was confirmed by most of them. Their choice was coded as correct or incorrect. After the face recognition test, children continued with the next task until all four tasks were completed.

In addition, children of the verbalization and control group were tested for their verbal abilities using the subtest “Analogies,” as part of the P-ITPA (Potsdam- Illinois Test for Psycholinguistic Abilities, German version: Esser, Wyschkon, Ballaschk, & Hänsch, 2010). Children had to complete given sentences, read by the experimenter, in a logical way (e.g., “A giant is big, a dwarf is …”). As stated in the manual, the subtest “Analogies” assesses verbal intelligence, which was considered as an indicator for verbal ability as it also taps vocabulary. At the very end, each child was told that all robberies seen in the previous videos were simulated in purpose of the study and that no real person was harmed.

Results

The recognition score per child could range between 0 and 4. For the mean number of correctly recognized faces in each experimental condition, see Figure 1. It was first checked whether the mean recognition score was above chance (i.e., 20%). Thus, one-sample *t*-tests were conducted separately for each experimental condition with an adjusted alpha level set at .016. The recognition score of the verbalization group was significantly above chance level, $p < .001$, which was also the case for the control group, $p = .012$. The children of the visualization group, however, only tended to perform above chance level (*p* = .04, no more significant after correction of the alpha level).

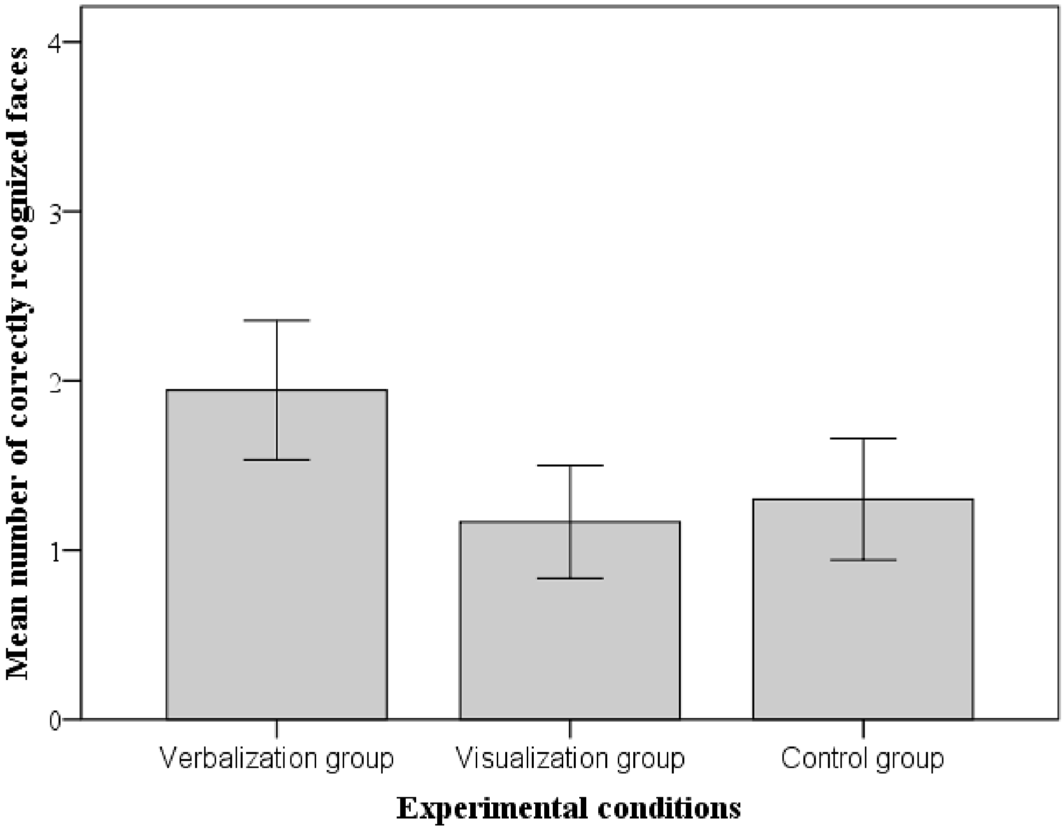

Figure 1. Mean number of correctly recognized faces in each experimental condition (max.: 4). *Note.* Error bars indicate standard deviations.

Additionally, a Chi-square analysis was conducted for each group separately (adjusted alpha .016) to check for proactive interference (i.e., best performance in the first task and progressive impairment across the following tasks). As no significant differences were revealed between successive tasks, a proactive interference could be ruled out. In order to test our main question whether the accuracy in face recognition significantly differed between the experimental conditions and whether a verbal or visual overshadowing effect occurred, the recognition scores were entered into an ANOVA. A main effect of condition was revealed, $F(2, 53) = 4.95$, $MSE = .40$, $p = .011$, $\eta^2 = .16$. Post hoc tests (Tukey) indicated that the accuracy in face recognition was significantly larger in the verbalization group compared to the visualization group, p = .014, and also larger compared to the control group, p = .042, suggesting a verbal facilitation effect rather than a verbal overshadowing effect. No significant differences were revealed between the control group and visualization group, $p > .05$, suggesting no visual overshadowing effect. In order to examine whether children exhibiting a lower level of verbal intelligence performed less accurately in face recognition than children with a higher level, a regression analysis was conducted separately for the verbalization and control group. The regression model in the verbalization group was significant, $F(1, 17) = 6.68$, MSE = 3.81, p = .020, R^2 = .30, with verbal intelligence predicting recognition accuracy, t = 2.59, p = .020, β = .54 (see Figure 2). No predictive value of verbal intelligence could be revealed in the control group, t = 1.67, $p > .05$, β = .3

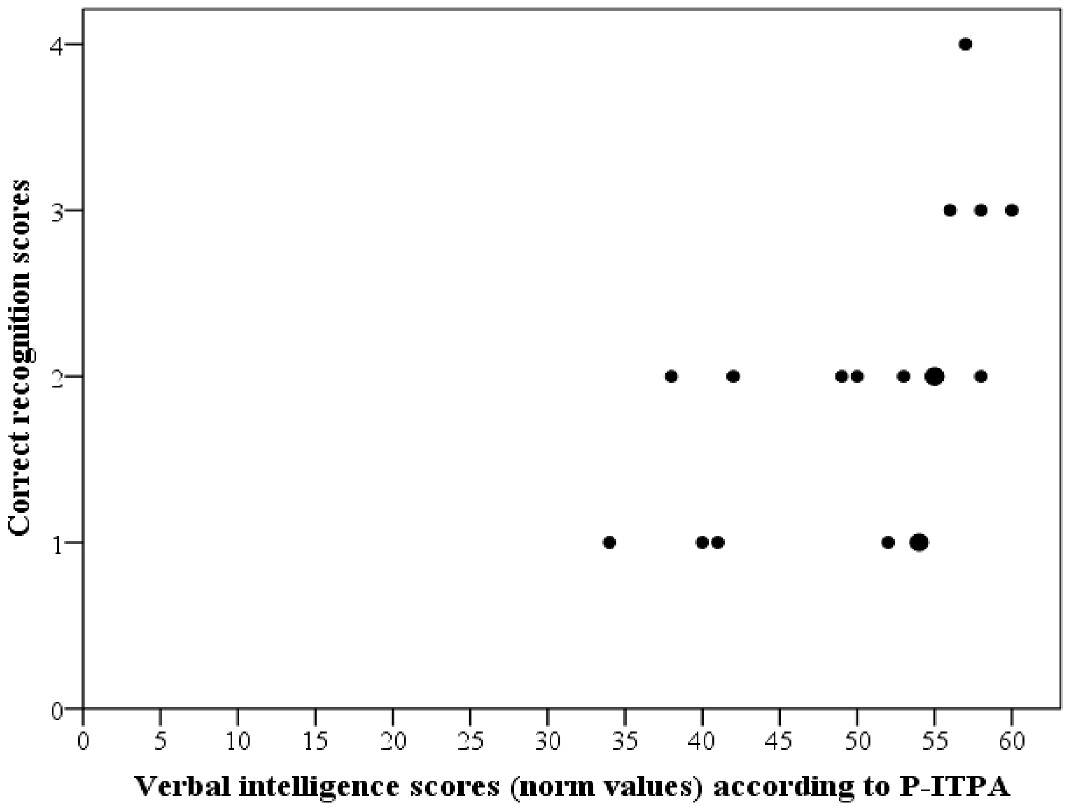

Figure 2. Scatterplot of recognition accuracy and verbal intelligence in the verbalization group. *Note.* Bold dots: Two participants with the same scores.

Discussion

The main questions of this study were (a) whether the verbal overshadowing effect occurs in 4- to 6-year olds, too, (b) whether the content account (Schooler & Engstler-Schooler, 1990) can make an explanatory contribution of this effect in young children and (c) whether visualization affects their face memory in a detrimental way. Contrary to our expectations, no verbal overshadowing effect was revealed. Instead, children seemed to benefit from verbalization compared to performance in the control and visualization group, suggesting a verbal facilitation effect. Verbal intelligence was a significant predictor for recognition accuracy in the verbalization group: Children with better verbal abilities profited more from verbalization than children with verbal poorer abilities, which is in line with the content account, but refers now to a verbal facilitation effect instead of a verbal overshadowing effect. Moreover, verbal intelligence did not predict recognition accuracy in the control group. Thus, we conclude that the beneficial effects of verbalization in part can be explained by the level of verbal intelligence rather than general intelligence (often correlated with verbal abilities). The latter could have been supported, if there had been a relationship between verbal intelligence and accuracy in the control group, too, which was not the case.

At least, three explanatory approaches can account for the absence of verbal overshadowing in 4- to 6-year olds. Previous studies suggest that face encoding strategies of younger children differ from the strategies of older children and adults. Whereas the latter encode faces using a holistic strategy based on the face's permanent configural information (i.e., the spatial relations among the face's features), children under 10 years of age use a feature-based strategy, that is they focus on distinctive features such as the nose rather than spatial configurations (Carey & Diamond, 1977; Meinhardt-Injac et al., 2011; Schwarzer, 2000). For example, Carey and Diamond (1977) demonstrated that children younger than 10 years of age rely more on paraphernalia (i.e., hats and eyeglasses) than older children to distinguish unfamiliar faces, and are easily fooled in recognition by the manipulation of these features (but see Baenninger, 1994). Carey and Diamond (1977) concluded that older children are more able to abstract permanent spatial relations in face (i.e., holistic strategy), and thus perform better in a recognition task when distinctive features are manipulated. Different face encoding strategies might have interesting implications regarding the susceptibility for verbal overshadowing. According to the processing account of verbal overshadowing (Schooler, 2002), verbalization triggers the use of feature-based strategy instead of holistic processing, which in turn impedes face recognition. Given a more feature-based strategy rather than a holistic encoding strategy in young children, it might be possible that the detrimental effects of verbalization do not apply in this age group: Due to a default feature-based encoding strategy, children are less disrupted by verbalization and thus, no verbal overshadowing effect occurs.

Second, it might be possible that young children are less prone to verbal overshadowing than older children and adults due to their lower level of vocabulary. This might be in line with the findings of Meissner and Brigham (2001) and Meissner, Brigham, and Kelley (2001): Experiments that fostered a detailed elaboration were more likely to reveal a verbal overshadowing effect in adults than experiments that only employed a free recall. One explanation is that deeper elaborations also increase the probability of self-generated erroneous details which subsequently impair face recognition (Meissner & Brigham, 2001). With respect to the generally limited vocabulary in 4- to 6-year olds, of course we tried to deepen the verbalization process by using a questionnaire, but maybe the degree of verbal elaboration was still not sufficient to trigger the generation of incorrect details that could have elicited a verbal overshadowing effect.

Finally, the comparison of the visualization with the verbalization group yielded significant differences in recognition accuracy, too. In line with that research on sensory dominance in childhood reports an auditory dominance in 4- to 6-year olds rather than a visual dominance (cf. Nava & Pavani, 2012; Sloutsky & Napolitano, 2003). This could be taken as a hint that younger children profit more from language-based processing rather than from visualization, which again would contradict the existence of a verbal overshadowing effect in this age group.

In conclusion, no verbal overshadowing effect could be demonstrated in young children. Rather, a verbal facilitation effect could be revealed. Verbal intelligence, however, was a significant predictor for recognition accuracy in the verbalization group, indicating that the beneficial effects of verbalization are mediated by the level of verbal abilities.

References

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahnik, S., Birch, S., Birt, A. R., & Zwaan, R. A. (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science, 9*, 556–578.

Baenninger, M. (1994). The development of face recognition: featural or configurational processing? *Journal of Experimental Child Psychology, 57*, 377-396.

Brandone, A. C., Salkind, S. J., Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Language development. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: development, prevention, and intervention* (pp. 499-514). Bethesda, MD: National Association of School Psychologists.

Brown, C., & Lloyd-Jones, T. J. (2005). Verbal facilitation of face recognition. *Memory & Cognition, 33,* 1442-1456.

Carey , S., & Diamond , R. (1977). From piecemeal to configurational representation of faces. *Science, 195*, 312-314.

Chin, J., & Schooler, J.W. (2008). Why do words hurt? Content, process, and criterion shift accounts of verbal overshadowing. *The European Journal of Cognitive Psychology, 20*, 396-413.

Clare, J. & Lewandowsky, S. (2004). Verbalizing facial memory: criterion effects in verbal overshadowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 739-755.

Dehon, H., Vanootighem, V., & Brédart, S. (2013). Verbal overshadowing of face memory does occur in children too! *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00970

Esser, G., Wyschkon, A., Ballaschk, K., & Hänsch, S. (2010). *Potsdam-Illinois Test für Psycholinguistische Fähigkeiten (P-ITPA)*. Göttingen: Hogrefe.

Francis, G. (2012). Too good to be true: publication bias in two prominent studies from experimental psychology. *Psychonomic Bulletin & Review, 19*, 151-156.

Meinhardt-Injac, B., Zöllig, J., Persike, M., Martin, M., Seiffge-Krenke, I., & Meinhardt, G. (2011). Developmental changes in the microgenesis of face perception revealed by effects of context and inversion. *Vision Research, 51*, 1338-1350.

Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology, 15*, 603-616.

Meissner, C. A., Brigham, J. C., & Kelley, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition, 29*, 176-186.

Memon, A., & Rose, R. (2002). Identification abilities of children: Does verbalisation impair face and dog recognition? *Psychology, Crime and Law, 8*, 229-242.

Nava, E., & Pavani, F. (2012). Changes in sensory dominance during childhood: converging evidence from the colavita effect and the sound-induced flash illusion. *Child Development, 84*, 604-116.

Ryan, R. S., & Schooler, J. W. (1998).Whom do words hurt? Individual differences in susceptibility to verbal overshadowing. *Applied Cognitive Psychology, 12*, 105-125.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology, 16*, 989–997.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology, 22*, 36-71.

Schwarzer, G. (2000). Development of face processing: the effect of face inversion. *Child Development, 71*, 391–401.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833.

Sporer, S. L., Malpass, R. S., & Koehnken, G. (Eds.) (1996). *Psychological issues in eyewitness identification.* Mahwah, NJ: Lawrence Erlbaum.

Appendix A

Questions posed to the children in the verbalization group:

Now I will ask you some questions about the robber that you saw in the video. It is important that you concentrate and try to remember the face.

• Was the face familiar to you?

• Was it a friendly or unfriendly face?

• Was there something in the face you noticed in particular?

• Was it a narrow or wide face?

• Can you say something about the eyes?

• Can you say something about the nose?

• Can you say something about the mouth?

• Can you say something about the hair style?

CHAPTER 5

Summary

The aim of this thesis was to examine the processing of audio-visual information in children and adults. In particular, I was interested in the developmental trajectory of sensory dominance. A great body of research has reported visual dominance in adults (e.g., Colavita visual dominance effect: Colavita 1976; Mc Gurk effect: McGurk & MacDonald, 1976; Ventriloquist effect: Howard & Templeton, 1966). So far, only few studies have addressed sensory dominance in children. For infants and 4-year-olds, auditory dominance has been reported (infants: Lewkowicz, 1988a, 1988b; Robinson & Sloutsky, 2004; 4-year-olds: Massaro, 1984; Sloutsky & Napolitano, 2003; Thompson & Massaro, 1994), with 4-year-olds exhibiting modality flexibility depending on stimulus familiarity (Napolitano & Sloutsky, 2004). Moreover, it was revealed that auditory dominance persists in 6-year-olds, and that adult-like visual dominance starts to emerge around the age of 9 years (Nava & Pavani, 2012; Thompson & Massaro, 1994). To further examine the developmental trajectory of sensory dominance two studies were conducted with 6-year-olds, 9-year-olds, and adults.

In Study 1 children and adults performed a Colavita task. Stimuli with semantic content were presented to test whether sensory dominance in children persists with more complex and meaningful stimuli in comparison to the simple stimuli that have been used in previous research (Nava & Pavani, 2012). A further question was whether semantic congruency might be a modulating factor for sensory dominance. Previous research in adults has shown that the Colavita visual dominance effect is modulated by temporal and spatial congruency (Koppen & Spence, 2007), but not by semantic congruency (Koppen et al., 2008). However, as for children, none of these factors (i.e., the influence of temporal, spatial and semantic congruency on sensory dominance) have been examined in a Colavita task. In addition to semantic congruency, the influence of stimulus frequency and color information on sensory dominance was also considered. In line with the study of Nava and Pavani (2012), 6-year-olds exhibited auditory dominance whereas 9-year-olds and adults exhibited visual dominance. Semantic congruency was a modulating factor of visual dominance in 9-year-olds, that is visual dominnace occurred only in semantically congruent trials. In contrast, sensory dominance in 6-year-olds and adults was not affected by semantic congruency. Additionally, the absence of color information did not affect auditory dominance in 6-year-olds, hardly affected visual dominance in adults, but terminated visual dominance in 9-year-olds. Stimulus frequency had no effect on sensory dominance in any age group. The results suggest that a) sensory dominance in children and adults is not restricted to simple lights and sounds but also occur with meaningful stimuli and that b) sensory dominance in 9-year-olds is more prone to interference than sensory dominance in 6-year-olds and adults, which might indicate a transitional stage at around the age of 9 years.

In Study 2, it was examined whether auditory dominance in 6-year-olds and visual dominance in older children and adults as reported by Nava and Pavani (2012) also occurs for a spatial task. Moreover, previous research (e.g., Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003) has made a distinction between sensory preference (i.e., the privileged use of one modality) and sensory dominance (i.e., disrupted encoding of non-privileged modality input). Therefore, I was interested in whether participants exhibited dominance or just a modality preference and whether there was a difference between age groups regarding sensory preference or dominance. Among 9-year-olds there was a strong trend towards privileged auditory processing. In contrast, 6-year-olds and adults showed no privileged modality processing. Nevertheless, 6-year-olds exhibited nearly twice as much auditory than visual responses on the group level and there were twice as many auditory responders than visual responders on the individual level. In turn, on a descriptive level, there was a trend towards more visual than auditory responses in the group of adults, and twice as many visual than auditory responders were revealed on the individual level. Considering the small sample sizes, it is possible that the lack of statistical significance on the group and the individual level among 6-year-olds and adults results from a lack of power. On the question of whether participants exhibited sensory preference or dominance, sensory preference was revealed among adults, whereas 6-year-olds exhibited sensory dominance. A considerable part of 9-year-olds exhibited sensory preference (albeit not on a statistically significant level), which might indicate increasing attentional resources in this age group that improve the ability to encode two modalities simultaneously.

Summarized, the results of Study 1 and Study 2 might support default auditory dominance in young children as proposed by Napolitano and Sloutsky (2004) that persists up to 6 years of age. However, in the study of Napolitano and Sloutsky (2004) 4-year-olds exhibited modality flexibility depending on stimulus familiarity. Visual dominance was revealed when visual stimuli were more familiar than auditory stimuli. Auditory dominance occurred when auditory stimuli were more familiar than visual stimuli and also when stimuli of both modalities were unfamiliar. Given that auditory dominance is the default whereas visual dominance is a result of learning (i.e., moderated by stimulus familiarity), visual dominance in 6-year-olds may be elicited easier than in 4-year-olds, as 6-year-olds may be familiar with a broader range of visual objects and scenes than 4-year-olds.

For 9-year-olds results on privileged modality processing were inconsistent. Whereas visual dominance was revealed in Study 1, privileged auditory processing was revealed in Study 2. Moreover, sensory dominance depended on stimulus characteristics such as color information (Study 1) and was less robust compared to sensory dominance in 6-year-olds and adults. Finally, privileged auditory processing in Study 2 contradicts the Modality Appropriateness Hypotheses (Welch & Warren, 1986), proposing that spatial tasks correspond with privileged visual processing. It thus might be possible that in 9-year-olds stimulus characteristics rather than task demands, such as working on a spatial versus temporal task, are more relevant for sensory dominance.

Among adults a visual dominance was observed in Study 1, which has also been demonstrated in preceding studies of the Colavita effect (for a review see Spence, Parise, & Chen, 2012). No sensory dominance was revealed in Study 2, which also contradicts the Modality Appropriateness Hypotheses. However, as mentioned above, there was a trend towards privileged visual processing on a descriptive level. Thus, lack of statistical significance regarding privileged visual processing should be interpreted with caution due to potential problems with the statistical power in this study.

Study 3 referred to verbal and visual overshadowing effects in 4- to 6-year-olds. The aim was to examine whether verbalization (i.e., verbally describing a previously seen face), or visualization (i.e., drawing the seen face) might affect later face recognition. No effect of visualization on recognition accuracy was revealed. However, there might have been a methodological shortcoming: note that children in the verbalization group were asked specific questions about the features (i.e., forced recall instruction), whereas children of the visualization group spent 1.5 minutes with free drawing (i.e., free recall instruction). As recognition accuracy can be affected by the manner in which participants are asked to recall the target (see Meissner, 2002 for “instructional bias effects” on verbalization) it might be worth to also test differential effects of forced versus free recall on visualization in future studies.

As opposed to a verbal overshadowing effect, a verbal facilitation effect was revealed. Moreover verbal intelligence was a significant predictor for recognition accuracy in the verbalization group but not in the control group. This suggests that strengthening verbal intelligence in children can pay off in non-verbal domains as well which might have educational implications. However, this requires further research.

References

Colavita, F. B. (1974). Human sensory dominance. *Perception and Psychophysics, 16*, 409-412.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. New York: Wiley.

Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the colavita visual dominance effect. *Experimental Brain Research, 184*, 533-546.

Koppen, C., & Spence, C. (2007). Audiovisual asynchrony modulates the colavita visual dominance effect. *Brain Research, 1186*, 224-232.

Lewkowicz, D. J. (1988a). Sensory dominance in infants: I. six-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 155-171.

Lewkowicz, D. J. (1988b). Sensory dominance in infants: II. ten-month-old infants' response to auditory-visual compounds. *Developmental Psychology, 24*, 172-182.

Massaro, D. W. (1984). Children´s perception of visual and auditory speech. *Child Development, 55*, 1777-1788.

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

Meissner, C. A. (2002). Applied aspects of the instructional bias effect in verbal over-shadowing. *Applied Cognitive Psychology*, 16, 911–928.

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*, 1850-1870.

Nava, E., & Pavani, F. (2012). Changes in sensory dominance during childhood: converging evidence from the colavita effect and the sound-induced flash illusion. *Child Development, 84*, 604-116.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*, 1387-1401.

Spence, C., Parise, C., & Chen, Y. (2012). The colavita visual dominance effect. In M. Murray & M. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 529-550). Boca Raton: CRC Press.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822-833.

Thompson, L. A., & Massaro, D. W. (1994). Children's integration of speech and pointing gestures in comprehension. *Journal of Experimental Child Psychology, 57*, 327-354.

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions, in K. R. Boff, L. Kaufman, & J. P. Thomas (Eds), *Handbook of perception and human performance* (pp. 1-36). New York: Wiley.

Universität Kassel, Fachbereich Humanwissenschaften

Erklärung zur kumulativen Dissertationen im Promotionsfach Psychologie

Erklärung zum Eigenanteil an den veröffentlichten oder zur Veröffentlichung vorgesehenen wissenschaftlichen Schriften innerhalb meiner Dissertationsschrift gemäß § 5a Abs. 4 Satz 1 der Allgemeinen Bestimmungen für Promotionen an der Universität Kassel vom 13. Juni 2012:

I. Allgemeine Angaben

Wille, Claudia

Institut für Psychologie, Universität Kassel

Thema der Dissertation: “Audio-Visual Information Processing Across Different Age Groups”

II. Nummerierte Aufstellung der eingereichten Schriften:

1. Wille, C. & Ebersbach, M. (2016). Semantic congruency and the (reversed) Colavita effect in children and adults. *Journal of Experimental Child Psychology, 141*, 23-33.
2. Wille, C. & Ebersbach, M. (submitted). Sensory dominance across different age groups. *Journal of Experimental Child Psychology.*
3. Wille, C., Völker, F., Kühnel, J., & Ebersbach, M. (2016). Verbal facilitation effects instead of verbal overshadowing in face memory of 4- to 6-year olds. *European Journal of Developmental Psychology, 13,* 231-240.

III. Darlegung des eigenen Anteils an diesen Schriften:

Zu Nr. 1:

- Konzeption: mehrheitlich
- Literaturrecherche: mehrheitlich
- Methodenentwicklung: mehrheitlich
- Entwicklung des Versuchsdesigns: mehrheitlich
- Datenerhebung: mehrheitlich
- Datenauswertung: mehrheitlich
- Ergebnisdiskussion: mehrheitlich
- Erstellen des Manuskripts: mehrheitlich
- Überarbeitung nach dem 1. Review: mehrheitlich

Zu Nr. 2:

- Konzeption: mehrheitlich
- Literaturrecherche: mehrheitlich
- Methodenentwicklung: mehrheitlich
- Entwicklung des Versuchsdesigns: mehrheitlich
- Datenerhebung: mehrheitlich
- Datenauswertung: mehrheitlich
- Ergebnisdiskussion: mehrheitlich
- Erstellen des Manuskripts: mehrheitlich
- Überarbeitung nach dem 1. Review: mehrheitlich

Zu Nr. 3:

- Entwicklung der Konzeption: in Teilen
- Literaturrecherche: mehrheitlich
- Methodenentwicklung: in Teilen
- Entwicklung des Versuchsdesigns: in Teilen
- Datenerhebung: in Teilen
- Datenauswertung: in Teilen
- Ergebnisdiskussion: mehrheitlich
- Erstellen des Manuskripts: mehrheitlich
- Überarbeitung nach dem 1. Review: mehrheitlich

IV. Anschriften der Mitautoren

Mirjam Ebersbach: mirjam.ebersbach@uni-kassel.de

Franziska Völker: Franziska.Voelker@gmx.net

Jessica Kühnel: jessica.kuehnel@web.de

Unterschrift des Antragstellers:

Ort und Datum | Dipl.-Psych. Claudia Wille

Ich bestätige die von Frau Wille unter Punkt III abgegebene Erklärung:

1. Prof. Dr. Mirjam Ebersbach Unterschrift:……………………………

2. Franziska Völker Unterschrift:……………………………

3. Jessica Kühnel Unterschrift:……………………………

Eidesstattliche Versicherung und Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlich-materiellen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich hierfür nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Ort und Datum　　　　　　　　　　　　　　　　　　Dipl.-Psych. Claudia Wille

Acknowledgements

I thank my family and Marc, Mirjam and my colleagues, interns and student assistants for their support. I thank the day care centers, children and parents, and everyone attending the studies.