



Published in final edited form as:

*Dev Psychobiol.* 2014 February ; 56(2): 292–315. doi:10.1002/dev.21197.

## Early Experience & Multisensory Perceptual Narrowing

**David J. Lewkowicz**

Florida Atlantic University, Boca Raton FL 33431

### Abstract

Perceptual narrowing is a reflection of early experience and contributes in key ways to perceptual and cognitive development. In general, findings have shown that unisensory perceptual sensitivity in early infancy is broadly tuned such that young infants respond to, and discriminate, native as well as non-native sensory inputs, whereas older infants only respond to native inputs. Recently, my colleagues and I discovered that perceptual narrowing occurs at the multisensory processing level as well. The present article reviews this new evidence and puts it in the larger context of multisensory perceptual development and the role that perceptual experience plays in it. Together, the evidence on unisensory and multisensory narrowing shows that early experience shapes the emergence of perceptual specialization and expertise.

---

Our world is specified by a plethora of physical attributes. When those physical attributes are detected by our sensory systems, they are perceived as belonging to perceptually coherent and meaningful objects and events rather than as collections of unrelated sensations (Gibson, 1966; Maier & Schneirla, 1964; Marks, 1978; Ryan, 1940; Stein & Meredith, 1993; Werner, 1973). This raises the obvious question: how does this ability develop? The answer must take into account the basic fact that humans, as well as many altricial species, are born structurally and functionally immature and relatively naïve because of only limited prenatal sensory experience. This means that multisensory perceptual mechanisms must emerge during development. In the case of human infants, multisensory perceptual mechanisms are fundamental to object and event perception, speech and language perception and production, and social responsiveness (Gibson, 1969; Piaget, 1952; Thelen & Smith, 1994). As a result, an understanding of perceptual, cognitive, and social development requires that we have a clear understanding of multisensory perceptual development as well.

This article reviews the current state of knowledge on the development of multisensory perception with a focus on multisensory perceptual narrowing (MPN), a newly discovered and seemingly paradoxical process. In essence, MPN contributes to multisensory perceptual development by gradually reducing the perceptual salience of some multisensory categories of information thereby narrowing response options. The paradoxical aspect of perceptual narrowing, including MPN, is that it reflects what Schneirla (1966) referred to as the non-obvious trace effects of the developing organism's typical ecological setting. Under normal conditions, developing infants are exposed to a wide array of sensory/perceptual experiences but, crucially, those experiences are usually restricted to only those attributes that are associated with that particular ecological setting. As a result, the perceptual expertise that ultimately emerges from this process mirrors the effects of that early selective experience.

This article is an update to a previous review of MPN by Lewkowicz & Ghazanfar (2009). It considers: (1) multisensory development and some of the theoretical issues related to it, (2) the progressive role of prenatal and postnatal experience in multisensory development, (3)

the concept of narrowing and its relationship to the earlier concept of canalization, (4) empirical findings on unisensory narrowing to set the stage for a discussion of MPN, (5) empirical findings to date on MPN, and (6) the theoretical implications of MPN.

## Development of Multisensory Perception

Infants enter the world with the ability to perceive certain forms of multisensory coherence. For example, newborn infants can learn to associate arbitrary objects and sounds (Slater, Brown, & Badenoeh, 1997), can perceive audio-visual equivalence based on intensity (Lewkowicz & Turkewitz, 1980), and respond differently to visual stimulation depending on whether auditory stimulation precedes it or not (Lewkowicz & Turkewitz, 1981). In addition, newborns can learn their mother's face when it is accompanied by her voice (Sai, 2005) and can perceive face-voice associations on the basis of their temporal co-occurrence (Lewkowicz, Leo, & Simion, 2010).

At birth, infants are relatively perceptually naïve and neurally and functionally immature. As a result, newborns possess rudimentary multisensory perceptual mechanisms that only enable them to perceive multisensory coherence based on relatively low-level perceptual cues. The role of two such cues has so far been studied in newborn multisensory perception: intensity and temporal audio-visual (A-V) synchrony. The latter cue is particularly powerful because it permits newborns to detect the temporal co-occurrence of any number of multisensory inputs which, in turn, enables them to bind such inputs and to perceive them as belonging to coherent multisensory objects or events. Crucially, newborns' ability to bind multisensory inputs does not depend on recognizing their identity. This is illustrated by findings from a study by Lewkowicz, Leo, & Simion (2010) of newborn infants' ability to match monkey visible and audible calls. Even though the newborns successfully matched the natural visible and audible calls, they also matched natural visible calls with broadband complex tones that had the same duration as the audible calls but whose envelope no longer had the temporal modulation of the natural audible calls. This indicates that the newborns' matching was based on the temporally synchronous onsets and offsets of the matching visible and audible stimuli and not on their identity.

The neural mechanisms that are essential for this sort of low-level responsiveness are known to exist. For example, studies have found that synchronous sounds can enhance the detection of partly occluded objects through the initial amplification of responses in primary as well as higher-order visual and auditory cortices and that they can do so regardless of task-context (Lewis & Noppeney, 2010). In other words, responsiveness to the synchronous occurrence of sights and sounds, regardless of their context, takes place at early stages of auditory and visual cortical processing. Assuming that newborns possess such early-processing cortical or, at a minimum, subcortical mechanisms (for evidence of the latter, see discussion below) then - despite their relative perceptual naïveté and neural immaturity - they can begin to construct a coherent conception of their multisensory world by binding audible and visible object and event attributes. This scenario is consistent with findings that newborns match monkey faces and vocalizations even in the absence of specific identity information (Lewkowicz et al., 2010).

There is no doubt that a perceptual mechanism that detects temporal co-occurrence is powerful and that it can bootstrap the development of multisensory coherence and, thus, of the concept of coherent multisensory objects and events. At the same time, however, there is little doubt that this sort of mechanism is quite limited. Indeed, given a perceptually rich and complex world and a rapidly growing nervous system that adapts itself to this world, infants can begin to accumulate perceptual experience and discover increasingly more complex types of intersensory relations. This, in turn, means that reliance on temporal A-V

synchrony cues can gradually begin to decline to a point where it is no longer the dominant intersensory relational cue. This kind of developmental process is illustrated by various findings. For example, young (3–4 month-old) infants can perceive the synchronous relations between moving objects and the sounds that they produce (Bahrick, 1988; Lewkowicz, 1992a, 1992b, 1994b, 1996) and, in addition, they begin to perceive the specific composition of the objects that make up such audiovisual events (Bahrick, 1988). Nonetheless, young infants rely to a great extent on temporal A–V synchrony and the redundancy that it creates for discovering higher-level multisensory invariant properties (Bahrick & Lickliter, 2000). This is evident in findings showing that 4-month-old infants only can perceive affect when it is represented by temporally synchronous auditory and visual perceptual attributes but that 5-month-old infants can perceive affect when it is represented by auditory-only attributes and that 7-month-old infants can perceive affect when it is represented by visual-only attributes (Flom & Bahrick, 2007).

Even though the perceptual importance of low-level intersensory relational cues declines, such cues continue to play an important role in perception throughout life. This is clear from findings that Chandrasekaran infants, children, adolescents, and adults all detect violations of temporal A–V synchrony relations and, most importantly, that they continue to rely on them for the perception of the multisensory coherence of a broad variety of multisensory events including light flashes and beeping sounds, moving and impacting objects, and audiovisual speech (Dixon & Spitz, 1980; Hillock-Dunn & Wallace, 2012; Lewkowicz, 2010; Lewkowicz & Flom, 2013). Thus, what declines across development is the degree of reliance on low-level cues for perceiving everyday multisensory coherence and what increases is the ability to detect audiovisual coherence on the basis of such higher-level intersensory relational perceptual cues as affect, gender, person identity, language identity, etc.

The general developmental pattern of decreasing reliance on low-level perceptual cues and the concurrent emergence of the ability to detect more complex intersensory relational perceptual cues is supported by various findings from studies of infant responsiveness to audiovisual inputs (Bremner, Lewkowicz, & Spence, 2012; Lewkowicz, 2000a; Lewkowicz & Ghazanfar, 2009; Lewkowicz & Lickliter, 1994; Lickliter & Bahrick, 2000; Walker-Andrews, 1997). For example, as early as two months of age, infants can match facially and vocally produced native speech syllables even when the audible syllable is synchronized with both visual syllables (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999, 2002, 2003; Walton & Bower, 1993). In other words, infants of this age can perceive the audiovisual coherence of speech syllables even in the absence of temporal A–V synchrony cues. Similarly, by three months of age infants match facial and vocal affect expressed by familiar people in the absence of synchrony (Kahana-Kalman & Walker-Andrews, 2001) but it is not until seven months of age that infants can match facial and vocal affect when synchrony cues are reduced and when the affect is expressed by strangers (Walker-Andrews, 1986). Likewise, it is not until six-to-eight months that infants begin to match audible and visible gender in the absence of synchrony (Patterson & Werker, 2002; Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991).

A similar developmental pattern holds for more basic types of intersensory relational cues like duration and spatiotemporal synchrony. For instance, it is not until six months that infants perceive duration-based A–V equivalence but even at this age infants only perceive it when concurrent synchrony cues specify it (Lewkowicz, 1986). Also, it is not until six months of age that infants begin to perceive “illusory” spatiotemporally based audiovisual relations such as when two objects moving through each other seemingly bounce against each other when a sound occurs during their coincidence (Scheier, Lewkowicz, & Shimojo, 2003). Finally, infants as young as two months of age can localize combined as opposed to

separate auditory and visual cues more rapidly but it is not until 8 months that they can integrate them in an adult-like non-linear manner (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006).

As indicated earlier, the spatiotemporal multisensory coherence that normally specifies our everyday ecological setting is a fundamental feature of our perceptual world. Given that infants take advantage of it to learn about their world, is early experience essential for the emergence of the behavioral and neural mechanisms that make it possible for infants to perceive temporal and spatial synchrony? First, as discussed in detail in a subsequent section on the effects of prenatal experience, fetuses have ample opportunity to experience spatiotemporal synchrony and, as suggested there, it is likely that such experience lays the foundation for responsiveness to spatiotemporal synchrony. Second, findings from studies manipulating prenatal and early postnatal experience have shown how important coherent audiovisual stimulation is for subsequent responsiveness. For example, studies of bobwhite quail embryos who were exposed to a spatially contiguous and temporally synchronous audiovisual maternal call exhibited a preference for a spatially contiguous maternal hen and call over the hen presented alone or the call presented alone (Jaime & Lickliter, 2006). Similarly, it has been found that bobwhite quail chicks' behavioral response to their mother's call depends on their having experience with spatiotemporally coordinated audible and visible maternal calls right after hatching (Lickliter, Lewkowicz, & Columbus, 1996). At the neural level, studies in cats also have found that the ability to integrate auditory and visual inputs depends on a developmental history of experience with spatiotemporally coincident auditory and visual inputs. Cats raised with auditory and visual stimuli presented randomly in space and time exhibit no multisensory integration in their superior colliculus neurons whereas cats raised with spatiotemporally coincident auditory and visual stimuli exhibit integration in these neurons (Xu, Yu, Rowland, Stanford, & Stein, 2012). Crucially, even though the cats in these experiments were raised with coincident simple flashes of light and simple bursts of broadband noise, the effects of such "simple" experience were found to generalize to responsiveness to other types of multisensory stimulus combinations. Together, these findings indicate that the developing avian and mammalian nervous system depends on exposure to the typical spatiotemporal concordance of the multisensory world for subsequent responsiveness to it as such.

The specific developmental timing of the relative importance of temporal synchrony cues versus higher-level cues is likely to depend on the amount of experience with the particular perceptual attribute in question. For example, as already mentioned, young infants can match facial and vocal representations of single phonemes in the absence of synchrony cues. Interestingly, however, young infants do not make such matches when the audible stimulus is a single tone or a three-tone non-speech analogue of an audible phoneme (Kuhl, Williams, & Meltzoff, 1991). Kuhl et al. (1991) interpreted the latter finding as evidence that infants require the full audible speech signal to make the match. Furthermore, they proposed that infants do not progress from an initial time when they relate faces and voices on the basis of some simple feature and then gradually build up a connection between the face and the full speech signal. Obviously, the Kuhl et al. (1991) proposal is at odds with the current view that reliance on low-level A-V synchrony relations bootstraps the development of responsiveness to higher-level multisensory cues. From the perspective of the current view, the Kuhl et al. (1991) proposal and is problematic for two reasons. First, there have been no published demonstrations to date that newborns match visual speech syllables with audible speech but that they do not with tones. Until such evidence is obtained, the Kuhl et al. hypothesis remains untested. Second, by the time infants are four months of age, they have had massive exposure to human faces producing visible and audible speech. As a result, it is reasonable to suppose that they develop two expectations. One is that human voices belong with human faces and, two, that whenever they see faces and lips moving, they expect to

hear and see speech, not tones. Indeed, consistent with this hypothesis, it has been found that 5-month-old infants look longer at human faces when they are paired with human voices than when they are paired with non-human vocalizations (Vouloumanos, Druhen, Hauser, & Huizink, 2009).

Aside from the issues raised above, the finding that infants can match audible and visible vowels without the aid of synchrony cues is interesting but it does not reflect the everyday world of infants. The audiovisual speech that infants usually experience is always co-specified by temporal and spatial synchrony cues. That is, whenever infants see and hear people talking, they hear and see the speech at the same time and coming from the same place. In other words, spatiotemporal audiovisual cues are an integral part of everyday audiovisual speech. Thus, until the proper experimental studies are done to investigate the separate and joint influence of phonetic and synchrony cues, it is premature to draw definitive conclusions regarding infants' perception of audiovisual speech. In addition, it should be noted that nearly all extant studies of infant matching of audible and visible speech have only investigated infant response to isolated phonemes. With one exception (e.g., (Dodd, 1979)), we know virtually nothing about infants' ability to perceive the relations between auditory and visual attributes of fluent speech. Unlike isolated phonemes, fluent audiovisual speech provides multiple and concurrent perceptual cues that specify phonetics, phonotactics, temporal and spatial synchrony, duration, tempo, rhythm/prosody, semantics, gender, affect, and identity. Obviously, infants can take advantage of one or more of these cues and in any combination depending on the degree to which they have acquired the ability to perceive each of them. This makes it clear that the question of when, whether, and how infants perceive the various perceptual cues associated with audiovisual speech is still an open one.

Overall, the accumulated body of empirical findings on the development of multisensory perception shed new light on two classic theories of multisensory development. One, the developmental integration view, holds that infants are initially naïve with respect to multisensory relations and that their ability to perceive them emerges slowly during early development (Birch & Lefford, 1963; Piaget, 1952). The other, the developmental differentiation view, holds that infants can perceive certain forms of multisensory coherence from birth and that their multisensory perceptual abilities gradually improve as infants learn and perceptually differentiate increasing levels of perceptual specificity (Gibson, 1969). As has previously been noted (Botuck & Turkewitz, 1990), however, and as is clear from extant empirical findings, infants neither start out life completely naïve to the multisensory coherence of their world nor do they perceive all forms of multisensory coherence at birth.

Despite the fact that neither classic theoretical view fully represents the actual development of multisensory functions, together these two views have led to a general consensus that multisensory perceptual abilities improve and broaden in scope with age (Bremner et al., 2012; Lewkowicz, 1994a, 2000a; Lickliter & Bahrick, 2000; Spector & Maurer, 2009; Walker-Andrews, 1997). Indeed, the extant empirical findings support this conclusion. Together, the two theoretical views also have led to the assumption that experience plays a key role in the development of multisensory functions albeit in different ways. According to the developmental integration view, infants gradually discover multisensory relations through their activity-dependent interactions with their world and resulting experience. According to the developmental differentiation view, infants gradually discover increasingly more complex amodal invariants through perceptual learning and differentiation and through the discovery of increasingly more complex action-perception links.

In the current context, the most important assumption made by both classic theoretical views is that the role of experience is a positive and progressive one because it leads to an

improvement and broadening of multisensory perceptual capacity. Recent empirical findings have shown, however, that this assumption is incomplete. These findings have shown that the two classic theoretical views overlook the fact that experience also may have regressive - though not maladaptive - effects and that this can lead to the narrowing of certain forms of multisensory responsiveness in early development. In essence, the recent findings have indicated that MPN is part-and-parcel of multisensory perceptual development and have suggested that experience mediates it. This is not surprising because experience is known to contribute in critical ways to the progressive development of multisensory perceptual abilities. The next section considers how prenatal and postnatal progressive experience influences affect multisensory development.

## **Progressive Role of Experience in the Development of Multisensory Functioning**

The likelihood that prenatal experience lays the foundation for the emergence of multisensory functions has been discussed in the past (Kenny & Turkewitz, 1986; Lickliter, 2011; Turkewitz, 1994; Turkewitz & Kenny, 1985). Despite this, empirical findings on the effects of prenatal experience in humans are sparse because it is difficult to conduct such studies. In contrast, studies of the effects of postnatal experience are more practically plausible and, as a consequence and as some of the findings reviewed above have shown, a sizeable literature has now accumulated indicating that experience plays a key progressive role in the growth of multisensory functions.

### **Prenatal experience & multisensory development**

The field of behavioral embryology has a long and distinguished history. Carmichael (1946) reviewed many of the findings from these studies many years ago and noted: "A knowledge of behavior in prenatal life throws light upon many traditional psychological problems." Indeed, behavioral embryology can teach us a great deal about the development of all sorts of functions (Smotherman & Robinson, 1990). This is certainly true of multisensory development. To understand how prenatal experience might contribute to multisensory development, it is crucial to recognize first that both in birds and mammals, all sensory modalities, except vision, begin functioning prior to birth and that their onset is sequential vis-à-vis one another (Bradley & Mistretta, 1975; Gottlieb, 1971). Specifically, sensory function emerges in the tactile, vestibular, chemical, and auditory modalities in that order prior to hatching or birth and in the visual modality after hatching or birth. As a result, human fetuses have ample opportunities to acquire multisensory experience. This can occur in a variety of ways. For example, once the tactile and vestibular modalities have their functional onset, a fetus can experience angular and linear acceleration along with the tactile consequences of such movement as it bumps up against the amniotic sac. Furthermore, fetuses can hear by the third trimester (DeCasper & Spence, 1986). If the mother happens to vocalize while she is moving, during the third trimester fetuses can experience tactile, vestibular, and auditory sensations either all at once and/or in close temporal proximity.

Another scenario involves thumb sucking and swallowing. Fetuses are known to suck their thumbs and swallow amniotic fluid. Supposing that a fetus begins to suck its thumb while the mother is moving and vocalizing. This creates an opportunity for interaction between the tactile consequences of thumb sucking and concurrent vestibular and auditory stimulation. Moreover, because fetuses are known to profit from olfactory experience (Marlier & Schaal, 2005), if they stop sucking and then swallow amniotic fluid, they have the opportunity to taste and perhaps also smell the various substances contained in the amniotic fluid while sensing movement and hearing sounds.

Importantly, the number and complexity of prenatal multisensory interactions is limited by the sequential onset of the different sensory modalities. As pointed out by Turkewitz (Turkewitz, 1994; Turkewitz & Kenny, 1982), this can actually be an advantage for the developing fetus because it can promote the orderly emergence of multisensory functions. For example, bobwhite quails learn their maternal call prior to hatching and, therefore, in the absence of visual stimulation. This visual “deprivation” is actually key because if bobwhite embryos are exposed to earlier-than-normal visual stimulation (by having their heads extruded from the egg prior to hatching) they fail to learn the maternal call (Lickliter & Hellewell, 1992). In other words, “imprinting” to the maternal call that normally occurs prior to hatching in bobwhite quail can only occur in the absence of visual input. Turkewitz suggests that the neural immaturity and relative lack of perceptual experience is advantageous because the neural substrate for each modality can become organized without competition for neural space from a subsequently emerging sensory modality. This idea of the positive effects of early developmental limitations follows from an earlier idea of ontogenetic adaptations (Oppenheim, 1981). According to this idea, each developmental stage can be considered to be an ontogenetic adaptation to the immediate exigencies of the organism’s current ecology rather than an immature phase in the organism’s drive to reach maturity.

In terms of multisensory development, the twin theoretical concepts of ontogenetic limitations and ontogenetic adaptations are, in turn, based on a developmental systems view of development (Lewkowicz, 2011). This view holds that each stage in the developmental emergence of any behavioral function reflects the co-acting influences of continually changing neural, behavioral, and experiential factors that, at each point in development, produce the most efficient adaptations of the organism to its needs and sensory challenges. For example, initial responsiveness to the temporal synchrony of auditory and visual sensory inputs can permit fetuses and infants to detect the co-occurrence of such inputs without necessarily enabling them to detect other correlations related to the more specific attributes of the stimuli. Therefore, even though initially they may not be able to detect the co-occurrence of perceptual attributes related to identity, it is enough for them at that point to detect simple co-occurrence so that they can eventually discover the higher-level correspondences.

In sum, there is little doubt that the nervous system becomes increasingly more multisensory as prenatal development progresses. Moreover, the rudimentary multisensory processing abilities that infants exhibit at birth most likely reflect the cumulative effects of prenatal multisensory experience. Of course, postnatal experience picks up where prenatal experience leaves off.

### **Postnatal experience & multisensory development**

As noted earlier, the conventional wisdom has been that perceptual experience leads to a broadening of multisensory perceptual abilities. Indeed, as previously shown, studies on the effects of postnatal experience have provided abundant evidence that early experience plays a key role in the development of multisensory functions in, both, animals and humans. There are many other examples besides those discussed earlier. For instance, studies of the neural map of auditory space in ferrets and barn owls have shown that the development of spatial tuning of these maps depends on concurrent visual input (King, Hutchings, Moore, & Blakemore, 1988; Knudsen & Brainard, 1991). Studies of the development of multisensory cells in the superior colliculus of cats and monkeys have found that these cells only begin emerging after birth, that they do not integrate multisensory signals when they emerge, and that their ability to integrate auditory, visual, and tactile localization cues depends on the specific spatial alignment of these cues early in life (Wallace & Stein, 1997, 2000, 2001, 2007). Finally, as indicated earlier, studies of the development of bobwhite quails’ typical

response to the mother hen and her calls have found that the bobwhite hatchlings' response to her audiovisual attributes depends on pre- and post-hatching experience. This experience must include concurrent auditory, tactile, and visual stimulation arising from prenatal self-produced and sibling-produced vocalizations, egg-egg interactions, and exposure to the visual attributes of the mother hen (Lickliter & Banker, 1994; Lickliter et al., 1996).

The various findings on the effects of early experience demonstrate unequivocally that the young nervous system is highly plastic and that it depends on exposure to temporally and spatially aligned multisensory inputs for the development of normal multisensory functions. Perhaps the most dramatic illustration of the rather extraordinary plasticity of the developing nervous system comes from experiments involving the re-routing of visual input into what ultimately becomes auditory cortex in neonatal ferrets (Sharma, Angelucci, & Sur, 2000; von Melchner, Pallas, & Sur, 2000). Specifically, when visual input is re-routed to primary auditory cortex, neurons become responsive to visual input and exhibit organized orientation-selective modules normally found in visual cortex. In addition, re-routed animals exhibit visually appropriate behavioral responsiveness. Thus, the developing nervous system is so plastic that, prior to specialization, cortical tissue that ultimately becomes specialized for responsiveness to sensory input in one modality has the capacity to respond to input from another modality.

Of course, as specialization proceeds, plasticity declines but, as studies have shown, the brain does retain some plasticity into adulthood (Amedi, Merabet, Bermpohl, & Pascual-Leone, 2005). For example, low-level primary visual cortex responds to auditory stimulation in adults (Romei, Murray, Cappe, & Thut, 2009) and congenitally blind adults exhibit language processing in the occipital cortex (Bedny, Pascual-Leone, Dodell-Feder, Fedorenko, & Saxe, 2011). The mechanisms that underlie these effects are currently not known, but it is known that blindness and deafness lead to an invasion of the cortical areas that normally respond to the missing input by the remaining intact modalities (Bavelier & Neville, 2002; Merabet & Pascual-Leone, 2010). In addition, it has been proposed that in the case of blindness, responsiveness to other than visual stimulation in visual cortex may be due to the disinhibition of existing cross-modal connections and/or the creation of new connectivity patterns (Amedi et al., 2005).

Regardless of what mechanisms are ultimately found to underlie human adult plasticity, it is clear from studies of the effects of early sensory deprivation in humans that appropriate early auditory and visual input is key to the development of normal multisensory functions. Thus, children who are born deaf and then have their hearing restored with cochlear implants prior to 2.5 years of age exhibit appropriate integration of audiovisual speech at a later age. If, however, the implants are inserted after 2.5 years of age they exhibit poor integration (Schorr, Fox, van Wassenhove, & Knudsen, 2005). In other words, even though young organisms possess greater functional and neural plasticity, normal access to typical auditory input is required during a prescribed early period for audiovisual integration to develop normally. Similar findings have been found in studies of adults who, as infants, were deprived of patterned visual input due to dense binocular congenital cataracts. These individuals exhibit deficits in certain forms of audiovisual integration even though their cataracts were removed during infancy. For example, one study of such early-deprived individuals assessed A-V interference effects and compared their performance to the performance of a control group of normally sighted adults (Putzar, Goerendt, Lange, Rösler, & Röder, 2007). One task required subjects to report which color in a series of rapidly changing colors had been presented when a target flash occurred. To test for multisensory interaction effects, an auditory distracter stimulus was presented either shortly before or after the flash (auditory capture condition) or simultaneously with it (baseline condition). Results indicated that the deprived adults exhibited less auditory capture effects than did the



non-deprived control subjects even though the performance of the deprived individuals on unisensory detection tasks was not impaired. In a second task, audiovisual speech integration was tested to see whether identification of single words would be enhanced by concurrent visual information (i.e., seeing the lips utter the word). The lip-read information enhanced identification in the sighted individuals but not in the deprived individuals.

Another study of individuals deprived of early patterned vision (also due to cataracts in infancy) examined responsiveness to concurrent auditory and visual speech syllables and, again, found deficits in these individuals (Putzar, Hötting, & Röder, 2010). Here, the task involved the well-known McGurk illusion where subjects are presented with incongruent auditory and visual speech syllables and, depending on the specific syllables presented, subjects either fuse them or combine them (McGurk & MacDonald, 1976). Fusion produces an illusory percept where the identity of the heard syllable is changed by the conflicting visual syllable through a process of audiovisual integration. The deprived and the sighted control subjects exhibited nearly perfect unisensory auditory performance but the deprived subjects exhibited poorer lip-reading and, as a result, also exhibited less McGurk illusions. This was the case even when the deprived individuals were equated to the control subjects in terms of their lip-reading ability. A brain imaging study comparing deprived and sighted individuals' performance on a lip-reading task of silently uttered monosyllabic words found that these two groups differed (Putzar, Goerendt, et al., 2010). Only the control subjects exhibited activations in the various cortical areas associated with lip-reading (e.g., the superior and middle temporal areas and the right parietal cortex). These results indicate that early visual deprivation has deleterious effects on the developmental organization of the neural substrate underlying lipreading and audiovisual speech integration.

Together, the Schorr et al. (2005) and the Putzar et al. (Putzar, Goerendt, et al., 2010; Putzar et al., 2007; Putzar, Hötting, et al., 2010) results, like the previously reviewed results from animal work, demonstrate unequivocally that early access to auditory and visual inputs is essential for the development of normal multisensory functions and that such access must occur during a sensitive period. This is supported by the fact that the deficit found in early-deprived adults was not reversed even though access to combined auditory and visual inputs was restored through cataract removal during infancy. Presumably, the sensitive period ended by the time combined audiovisual input was restored.

A particularly interesting aspect of the results from the auditory and visual deprivation studies is that they are consistent with findings from research on the development of audiovisual responsiveness in infancy. This research has shown that infants "expect" the auditory and visual attributes of speech and non-speech events to be synchronized (presumably because of their extensive experience with synchrony going back to prenatal development and with A-V synchrony from birth on). This expectation is evident in the findings reviewed earlier showing that infants are sensitive to A-V asynchrony and that they can learn synchronous audiovisual events but not asynchronous ones (Bahrick, 1988; Bahrick & Lickliter, 2000; Lewkowicz, 1992a, 1992b, 1996, 2000b, 2010; Lewkowicz et al., 2010; Scheier et al., 2003). Furthermore, this is evident in the fact that the redundancy created by synchronous auditory and visual speech begins to capture attention just as infants begin learning how to talk (Lewkowicz & Hansen-Tift, 2012).

In sum, two developmental principles can be distilled from the foregoing section. First, appropriate prenatal and postnatal experience is essential for the emergence of normal multisensory function. Second, experience is required during a sensitive period – a delimited window of time during early development - for normal multisensory functions to emerge.

## Experience & Canalization

The critical effects of prenatal and postnatal experience discussed so far all involve a progressive broadening of perceptual capacity. This fits with the reasonable and conventional view that early multisensory experience has positive developmental effects. It is theoretically possible, however, that in some cases experience may narrow an initially broadly tuned multisensory perceptual system and, thus, lead to a regression of perceptual function. Indeed, the general idea that developmental experience can sometimes lead to functional regression is not new. Many years ago, Holt (1931) put forth the concept of behavioral canalization to account for the developmental emergence of organized motor activity patterns from the initially diffuse motor patterns that are characteristic of early embryonic development in chicks. Holt proposed that the initial and diffuse motor patterns are canalized into organized motor patterns via behavioral conditioning. Later, Kuo (1967) expanded Holt's concept of behavioral canalization by asserting that conditioning cannot be solely responsible for narrowing of behavioral potential. He suggested that canalization is also due to the individual's developmental history, context, and experience. Finally, Gottlieb (1991a) put Kuo's concept of behavioral canalization to empirical test in his work on the development of perceptual tuning in birds.

Working with mallard ducks, Gottlieb (1991a) investigated the developmental basis of imprinting in birds by studying the role of prenatal factors in the usual preference that hatchlings exhibit for the species-specific call of their mother. Gottlieb found that mallards acquire their preference for the species-specific maternal call prior to hatching as a result of exposure to the vocalizations of their siblings as well as self-produced vocalizations. Especially interesting was the finding that in the absence of such early experience hatchlings were more broadly tuned and, as a result, responded to the maternal calls of other species (i.e., chickens) as well. In other words, the hatchlings' ultimate preference for the species-specific call reflected a narrowing of an initially broader sensitivity and a buffering against the learning of another species' call.

When Gottlieb (1998) considered the implications of his findings, he drew a key distinction between two historical meanings of the concept of canalization. One of these meanings is the one discussed here where developmental experience - broadly construed to include all obvious and non-obvious stimulative influences (Lehrman, 1953, 1970; Schneirla, 1966) - helps construct developmental outcomes. The other meaning is that proposed by Waddington (1957) who held that canalization is a genetically controlled process that restricts the range of developmental outcomes. Gottlieb (1998) included Waddington's concept in his expanded version of it but noted that developmental canalization not only emanates from genetic influences but also from normally occurring developmental experiences which, in turn, can serve as signals for gene activation. This expanded version of the concept of canalization not only recognizes the crucial contribution that intrinsic biological factors make to development but also calls attention to the equally important contribution that all stimulative factors make to developmental outcomes and, thus, requires that we investigate them.

The work on canalization in birds was taking place at the same time as work that has been termed "narrowing" was taking place at the human level. The latter work has led to discoveries showing that the regressive effects of early experience lead to the tuning of the human infant's perceptual system and that this ultimately leads to a match between the infant's perceptual system and the exigencies of its everyday ecological niche. The next section reviews this work by discussing studies on narrowing of speech, face, and music perception.

## Unisensory Perceptual Narrowing

Evidence that narrowing has profound effects on the development of perceptual functions began to appear in the early 1980s and by now has grown into a large and impressive body of evidence. Together, this evidence has made it clear that perceptual narrowing is a domain-general process in that the perception of speech, music, and faces narrows during the early months of life. In general, like pre-hatching duck embryos, human infants are initially broadly tuned and respond to native as well as non-native perceptual inputs. Then, through selective exposure to native-only inputs, perceptual tuning narrows over the first few months of life.

### Narrowing of speech perception

Some of the earliest evidence of narrowing came from studies of infant response to phonetic distinctions. In the first of these studies, Werker and Tees (1984) found that 6–8 month-old English-learning infants discriminated non-native consonants (i.e., the Hindi retroflex /Da/ versus the dental /da/ and the Thompson glottalized velar /k'i/ versus the uvular /q'i/) but that 10–12 month-old infants no longer discriminated them. Subsequent studies found that such cross-linguistic narrowing occurs in response to other consonant and vowel pairs (Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Cheour et al., 1998; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992).

Subsequent studies have shown that the decline reflects the effects of language-specific experience. One of these studies has shown that native-language experience not only leads to the narrowing of perceptual sensitivity to a non-native language but that it also facilitates the discrimination of native-language phonetic contrasts between six and 12 months of age (Kuhl et al., 2006). A second study (Kuhl, Tsao, & Liu, 2003) also reported that language-specific experience maintains sensitivity to the phonetics of that particular language as well as that social interaction is essential for this to occur. In this study, it was found that English-learning infants who were exposed to a Mandarin Chinese speaker during multiple play sessions between 9 and 10 months of age were better able to discriminate a Mandarin Chinese phonetic contrast (which does not occur in English) than were infants who were not exposed to the Mandarin Chinese. Subsequent studies have found, however, that social interaction may not be necessary for maintaining discrimination. For example, Yeung & Werker (2009) found that maintenance (or reactivation of sensitivity) to non-native distinctions does not require that infants engage in contingent interactions but, simply, that non-native sounds be paired with distinct objects. Similarly, Yoshida, Pons, Maye, & Werker (2010) found that simple distributional learning is effective at reactivating sensitivity to non-native distinctions at 10 months, an age when such sensitivity would otherwise be in decline. Regardless of whether social interaction is essential or not for maintenance of non-native phonetic discrimination, it is important to note that once narrowing has occurred during infancy, native-language phonetic discrimination abilities continue to improve throughout childhood (Sundara, Polka, & Genesee, 2006). Whether this improvement depends specifically on social interaction or not is currently not known.

One of the most interesting facts that has emerged from studies of narrowing in the speech domain has been that native-language experience also narrows infant response to silent visual speech (Weikum et al., 2007). Specifically, it has been found that monolingual English-learning 4- and 6-month-old infants can discriminate silently articulated English as well as French syllables but that monolingual English-learning 8-month-old infants exhibit no evidence of non-native language discrimination. In contrast, bilingual 8-month-old infants continue to discriminate both languages. It has also been reported that bilingual experience in the target languages may not be necessary for maintaining the initial broad perceptual tuning. Sebastián-Gallés et al. (2012) reported that 8-month-old infants who grow

up in a bilingual environment where Spanish and Catalan are spoken also can discriminate English and French silent visible syllables but that infants who are only exposed to Spanish or Catalan do not make such discriminations. Unfortunately, interpretation of these findings is complicated by the lack of data from younger (i.e., 4-month-old) monolingual infants' response in this task. Based on a perceptual narrowing account, it is theoretically possible that younger monolingual infants, who are known to have broad perceptual sensitivity, also should be able to discriminate between two unfamiliar visual-only languages other than the one that they normally experience. If that turns out to be the case then this would call into question the conclusion that bilingual experience *per se* contributes to greater sensitivity to visible speech contrasts in any language.

### Narrowing of face perception

Studies have found evidence of narrowing in infant response to other-species and other-race faces too. Pascalis, Haan, and Nelson (2002) first showed that 6-month-old infants can recognize and discriminate monkey faces but that 9-month-old infants no longer do. Subsequent studies found that experience underlies this kind of narrowing in that infants who are exposed to monkey faces at home between six and nine months of age continue to exhibit discrimination of monkey faces at nine months of age (Pascalis et al., 2005; Scott & Monesson, 2009). Studies of infant response to other-race faces have yielded similar evidence of narrowing. They have found that 3-month-old infants discriminate the faces of other races, that this ability declines by nine months of age (Kelly et al., 2007), that it is independent of culture (Kelly et al., 2009), and that narrowing is the result of selective experience with same-race faces (Sangrigoli & de Schonen, 2004). Furthermore, studies have found that the perceptual system continues to be relatively plastic during the decline as well as right after the initial decline is completed. That is, infants who are given extra exposure to other-race faces while narrowing is in progress continue to discriminate other-race faces (Anzures et al., 2012). Similarly, infants who have narrowed but who are given additional exposure and testing time during an experiment testing for discrimination of other-species faces exhibit successful discrimination of such faces (Fair, Flom, Jones, & Martin, 2012). Finally, it appears that selective experience with human faces tunes the visual system to prototypical face attributes; whereas 6-month-old infants prefer faces with atypically large eyes over faces with typically sized eyes, 12-month-old infants prefer faces with typically sized eyes (Lewkowicz & Ghazanfar, 2012). Why 6-month-old infants prefer large eyes is not clear, but the reversal of preference by 12 months suggests that as infants encode the human face prototype, their perceptual tuning for prototypically human eyes narrows.

### Narrowing of music perception

Like evidence from speech and face perception, evidence from studies of music perception indicates that infants are initially broadly tuned and that the tuning narrows as infants are exposed to the dominant musical rhythms and meters of their culture. For example, it has been found that 6-month-old North American infants can detect violations of simple meters (2:1 ratios of inter-onset interval durations) as well as complex musical meters (3:2 ratios characteristic of non-Western music) but that 12-month-old North American infants no longer detect violations of complex meters (Hannon & Trehub, 2005b). As in speech and face perception, findings indicate that narrowing of perceptual sensitivity to musical meter is due to selective experience with native input. Two weeks of exposure to non-native meters at 12 months of age is sufficient to restore discrimination of non-native rhythms in infants but not in adults (Hannon & Trehub, 2005a, 2005b). This indicates that the sensitive period for tuning the auditory system to culturally-specific musical meters closes sometime between 12 months of age and adulthood.

## General Principles

Three general principles emerge from the work on unisensory perceptual narrowing. First, narrowing is a relative phenomenon in that, as noted by Werker and Tees (2005), it represents a re-organization of perceptual sensitivity, not a loss of discriminability. Second, narrowing/re-organization is made possible by early plasticity which, as indicated earlier, does not end in infancy. That is, even though plasticity declines rapidly during infancy, some plasticity is retained into later development as illustrated by the fact that the other-race effect can be reversed during childhood (Sangrigoli, Pallier, Argenti, Ventureyra, & de Schonen, 2005). Finally, even though narrowing leads to a decline in sensitivity to non-native perceptual inputs, it also marks the beginning of specialization and initial expertise which then grows continually into the adult years.

## Multisensory Perceptual Narrowing

The foregoing indicates that perceptual narrowing is a domain-general process. This raises the theoretical possibility that perceptual narrowing is a pan-sensory process as well. What follows is a review of recent studies in which my colleagues and I investigated this possibility

### MPN of the perception of other-species faces & vocalizations in human infants

It is known that infants as young as two months of age and as old as 18 months of age can perceive the coherence of human faces and human vocalizations as evident in their ability to match them (Kahana-Kalman & Walker-Andrews, 2001; Kuhl & Meltzoff, 1982; Patterson & Werker, 1999, 2002, 2003; Poulin-Dubois, Serbin, Kenyon, & Derbyshire, 1994; Walker, 1982; Walker-Andrews, 1986) and that this ability improves with age (Bahrick, Hernandez-Reif, & Flom, 2005). It was not known until recently, however, whether this ability extends to the multisensory perception of other-species faces and vocalizations. Lewkowicz and Ghazanfar (2006) hypothesized that it probably does and also proposed that this ability probably narrows as infants age and as they acquire increasingly greater experience with human faces and vocalizations.

To test their prediction, Lewkowicz and Ghazanfar measured visual preferences in 4-, 6-, 8- and 10-month old infants while they watched side-by-side movies depicting the same rhesus monkey repeatedly making a coo call on one side and a grunt call on the other side (call onsets were simultaneous). The visible calls were presented in silence during the first two trials and together with one or the other audible vocalizations during the second two trials. As predicted, the 4- and 6-month-old infants looked longer at a visible call in the presence of its matching vocalization than in its absence whereas 8- and 10-month-old infants did not.

The monkey coo is longer than the grunt. This means that during the trials when the audible call was presented, its temporal onsets and offsets corresponded to the onsets and offsets of the matching visible call but only its onsets corresponded to the onsets of the non-matching visible call. As a result, infants may have based their matching on A-V temporal synchrony and/or duration. Lewkowicz, Sowinski, and Place (2008) tested this possibility by repeating the Lewkowicz & Ghazanfar (2006) study but this time with the visible and audible calls desynchronized (but still corresponding in terms of duration). This time, the younger infants no longer exhibited face-voice matching indicating that synchrony, not duration, drives matching. Lewkowicz, Sowinski, and Place (2008) also examined whether the older infants' failure to make matches in the original study was due to narrowing of unisensory responsiveness and whether the decline in matching persists beyond 10 months of age. They found that unisensory narrowing does not account for the results in the older infants because 8–10 month-old infants exhibited differential looking at the silent visual calls and because they discriminated the audible-only coos and grunts. Finally, Lewkowicz et al. (2008) found

that older, 12- and 18-month-old, infants did not match the monkey faces and vocalizations (even though the visible and audible calls were synchronized). This indicates that the effects of MPN persist into the second year of life.

This initial set of studies documenting MPN in infancy were followed up by subsequent studies designed to investigate the generality of this process. These studies asked three specific questions. First, does the broad multisensory perceptual tuning found in young infants and their ability to match other-species faces and vocalizations represent the initial developmental condition in humans (i.e., is it present at birth)? Second, does this broad multisensory perceptual tuning extend to other domains such as, for example, audiovisual speech? Finally, what might be the evolutionary roots of MPN?

### **Newborn matching of other-species faces and vocalizations**

Lewkowicz, Leo, and Simion (2010) tested newborn infants to determine whether they, like 4- and 6-month-old infants, might be broadly tuned. Thus, the newborns were tested with the identical stimuli and procedures used in the Lewkowicz & Ghazanfar (2006) study. Like the 4- and 6-month-old infants, newborns matched the monkey visible and audible calls. In a follow-up experiment, Lewkowicz, Leo, and Simion (2010) examined the possibility that newborns may base their matching on low-level features of the visible and audible calls rather than the identity information inherent in them. Given that 4- and 6-month-old infants base their matching on onset and offset synchrony (Lewkowicz et al., 2008), it was hypothesized that newborns also may not only match on the basis of synchrony but that they may do it simply on the basis of the energy onsets and offsets of the matching calls. To test this possibility, Lewkowicz, Leo, and Simion (2010) repeated the first experiment except that this time they presented a complex tone instead of the natural call so as to remove identity information from the audible call. Despite the lack of identity information, the newborns still matched.

These findings indicated that newborns do not rely on the dynamic correlations that are typically available in audiovisual vocalizations (Chandrasekaran, Trubanova, Stillitano, Caplier, & Ghazanfar, 2009) for making face-vocalization matches. Rather, it appears that they match on the basis of the synchronous onsets and offsets of visual and auditory energy. This is consistent with an ontogenetic-adaptation interpretation in that it shows that the newborn multisensory response system is primarily sensitive to the synchronous onsets and offsets of audiovisual stimulation. Importantly, the Lewkowicz, Leo, and Simion (2010) findings demonstrate that temporal synchrony is sufficient for matching but not that it is necessary. Other findings from newborn studies do, however, provide evidence that temporal synchrony is necessary for matching. For example, Slater, Quinn, Brown, and Hayes (1999) found that newborn infants who hear a sound only when they look at an object learn this arbitrary object-sound pairing but that newborns who hear the sound regardless of whether they look at the object or not do not learn the pairing.

Together, these findings make it clear that a system that is sensitive primarily to intersensory temporal synchrony is quite powerful even if it is rather rudimentary. Its power derives from the fact that it can set in motion the gradual discovery of a more complex and coherent multisensory world based on higher-level perceptual cues.

### **MPN of responsiveness to non-native audiovisual speech syllables**

If the decline in multisensory matching of other-species faces and vocalizations reflects a general feature of multisensory perceptual development then it should be reflected in other domains. Pons, Lewkowicz, Soto-Faraco, and Sebastián-Gallés (2009) hypothesized that it might also be reflected in the development of audiovisual speech perception. To investigate

this possibility, they tested 6- and 11-month old English- and Spanish-learning infants' response to two different visible syllables following familiarization to an audible version of one of these two syllables. Specifically, Pons et al. (2009) first presented side-by-side videos of the same woman repeatedly uttering a silent /ba/ on one screen and a silent /va/ on another screen and recorded looking preferences to establish a baseline preference. Following this, they familiarized the infants to the audible version of one of the syllables (in the absence of the visual syllables) and then repeated the presentation of the visible-only syllables and measured visual preferences again. If infants could match the correct visible syllable to the one that they had just heard then they were expected to look longer at it relative to the amount of time they looked at it prior to familiarization. Crucially, because infants never heard and saw the syllables at the same time, the familiarization/test design ensured that they had to extract the higher-level audible and visible syllable features to match them rather than simply detect their synchronous occurrence.

Based on the possibility of perceptual narrowing, it was expected that the English-learning infants would match the audible and visible syllables at both ages but that the Spanish-learning infants would only match at six months of age. The reason that the older Spanish-learning infants' were not expected to match was because /b/ and /v/ are homophones in Spanish and, as a result, the phonetic distinction between a /ba/ and /va/ does not exist in Spanish. As predicted and as Figure 1 shows, the English-learning infants matched the audible and visible syllables at both ages whereas the Spanish-learning infants only matched them at six months of age. These results provided the first evidence that MPN is involved in the development of native audiovisual speech perception in infancy. They also raised the question of whether the effects of this type of audiovisual speech narrowing persist into adulthood. To answer this question, Pons et al. (2009) also tested monolingual English- and Spanish-speaking adults with a similar procedure except that the adults were asked to indicate which of the two visible syllables corresponded to an immediately preceding presentation of one or the other audible syllable. As Figure 2 shows, the English-learning adults made correct matches on over 90% of the trials whereas the Spanish-speaking adults made random choices. Thus, once MPN of audiovisual speech occurs, its effects appear to persist into adulthood (although future studies will need to determine what happens in the intervening years).

### **MPN & the ability to recognize the amodal identity of one's native language**

If perception of audiovisual speech at the syllabic level undergoes narrowing then this is likely to affect infant perception of fluent audiovisual speech too. One way to test this possibility is to ask whether infants can recognize the amodal identity of their native speech. It is known that infants can distinguish between languages on the basis of their prosody starting at birth (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998) and that this early and broad sensitivity becomes refined and attuned to the specific prosodic characteristics of their native language (Jusczyk, Cutler, & Redanz, 1993; Nazzi, Jusczyk, & Johnson, 2000; Pons & Bosch, 2010). As a result, it is likely that this prosodic sensitivity is pan-sensory and, thus, that as infants learn the various attributes related to speech and language - including such higher-level segmental features as stress patterns (Jusczyk et al., 1993), languages-specific combinations (Jusczyk, Luce, & Charles-Luce, 1994), and familiar word forms (Swingley, 2005; Vihman, Nakai, DePaolis, & Hallé, 2004) - they begin to recognize their native audiovisual speech as familiar and can distinguish it from non-native audiovisual speech. To examine this possibility, Lewkowicz & Pons (2013) used the Pons et al. (2009) familiarization/test procedure to test 6–8 and 8–10 month-old, monolingual, English-learning infants' ability to recognize the amodal identity of English and Spanish utterances. First, infants were familiarized with a continuous-speech utterance either in English or in Spanish and then tested for recognition of the visible form of the utterance by

seeing side-by-side videos of a bilingual female uttering the utterance in English on one side and in Spanish on the other side.

Results indicated that the younger infants did not exhibit any differential responsiveness to the videos of the two languages. In contrast, the older infants did. Specifically, those older infants who were familiarized with the English audible utterance looked longer at the Spanish-speaking face following familiarization whereas those who were familiarized with the Spanish audible utterance did not exhibit differential looking. This novelty effect is similar to previously reported novelty effects in infancy in studies of multisensory perception (Gottfried, Rose, & Bridger, 1977) and of visual perception (Pascalis et al., 2002).

Lewkowicz & Pons (2013) interpreted the preference for the novel visible utterance as indicating that older infants recognized the correspondence between the previously heard utterance in their native and, thus, familiar language and a face that could be seen speaking in the same language. Crucially, because the audible and visible information was presented at different times, infants had to extract, remember, and match the common spatiotemporally correlated patterns of optic and acoustic prosodic information to recognize the amodal identity of the fluent audiovisual speech. It should be noted that alone this result does not provide clear evidence of MPN. It does, however, provide evidence of MPN when it is considered in the context of the failure to find an effect following familiarization with a non-native utterance. That is, if infants become specialized for their native language by the end of the first year of life then it should be more difficult for them to extract the patterns of non-native optic and acoustic prosody. English and Spanish belong to different rhythmic classes - the former is stress-timed and the latter is syllable-timed. Presumably, because of MPN, the prosodic characteristics of the native language were more familiar to the older infants and this permitted them to match whereas the prosodic characteristics of a non-native language were too unfamiliar to permit a match.

### **Developmental changes in multisensory selective attention & MPN**

Usually, when we interact with other people, we can hear them talking as well as see their lips moving. Seeing as well as hearing speech is known to increase its salience and comprehension (Rosenblum, Johnson, & Saldana, 1996; Sumby & Pollack, 1954; Summerfield, 1979). Therefore, it would be adaptive if infants could take advantage of the increased salience of audiovisual speech and begin to lipread, especially when they begin to produce their first speech sounds at the onset of canonical babbling. Moreover, it is possible that if they do begin to lipread then the degree to which they do so may be modulated by MPN.

To determine whether infants begin lipreading when they begin babbling, Lewkowicz and Hansen-Tift (2012) investigated selective attention to talking faces in 4–12 month-old infants. Prior studies have investigated infants' selective attention to static faces, dynamic silent faces, or talking faces but all of them focused on the period prior to the onset of babbling (Cassia, Turati, & Simion, 2004; Haith, Bergman, & Moore, 1977; Hunnius & Geuze, 2004; Merin, Young, Ozonoff, & Rogers, 2007). As a result, Lewkowicz and Hansen-Tift (2012) included older infants to determine whether selective attention might shift to a talker's mouth once infants enter the canonical babbling stage during the second half of the first year of life. This shift was expected for several reasons. First, the speaker's mouth is where the tightly coupled and highly redundant patterns of auditory and visual speech information that imbue audiovisual speech with its greater salience are located (Chandrasekaran, Trubanova, Stillitano, Caplier, & Ghazanfar, 2009; Munhall & Vatikiotis-Bateson, 2004). As a result, the mouth is likely to attract a good deal of attention. Second, it is known that the development of speech-production capacity in infancy is facilitated by



imitation, language-specific experience, and social contingency (de Boysson-Bardies, Hallé, Sagart, & Durand, 1989; Goldstein & Schwade, 2008; Kuhl & Meltzoff, 1996). If infants begin lipreading as they are learning how to speak then they can more accurately imitate and respond to the communication signals of others. Finally, endogenous attentional mechanisms, which enable infants to voluntarily direct their attention to events of interest, begin emerging around six months of age (Colombo, 2001). If infants become interested in producing and imitating speech sounds when they begin babbling then the newly emerging endogenous attentional mechanisms enable them to shift their attention to their interlocutor's mouth.

In addition to the *a priori* theoretical predictions about the emergence of lipreading, there are good theoretical reasons to expect that early experience, via MPN, is likely to play a key role in this process. That is, as infants acquire increasingly greater experience with their native language and as they become increasingly more specialized, they are likely to change the way they attend to people speaking in different languages. Specifically, infants may begin attending less to the source of native audiovisual speech (i.e., the mouth) once perceptual narrowing to their native language has occurred but they may continue attending more to the mouth in response to non-native audiovisual speech.

Lewkowicz & Hansen-Tift (2012) conducted two experiments to determine whether infants become lipreaders as they enter the canonical babbling stage and whether their reliance on lipreading is affected by MPN. Using an eye tracker, they recorded the point of visual gaze while infants watched and listened either to a 50 s of a female reciting a monologue in her native English (Experiment 1) or a video of another female reciting the same monologue in her native Spanish (Experiment 2). To determine when the predicted developmental changes in selective attention may occur, Lewkowicz & Hansen-Tift (2012) tested separate cohorts of 4-, 6-, 8-, 10- and 12-month-old monolingual English-learning infants and a group of monolingual English-speaking adults in each experiment, respectively.

As Figure 3 shows, the overall pattern of results across the different ages and across the two experiments was consistent with predictions. That is, as reported in prior studies (Haith et al., 1977; Merin et al., 2007), 4-month-olds looked more at the eyes than the mouth. In contrast, by six months of age, infants began to exhibit initial evidence of the expected attentional shift to the mouth in that by now they looked equally long at the eyes and mouth. By eight months of age, the shift now appeared to have occurred fully in that by now 8-month-old infants directed a significant proportion of their looking at the talker's mouth. The same was the case at 10 months of age. Finally, an attentional shift away from the mouth began to appear at 12 months in response to native speech in that infants no longer looked more at the mouth than at the eyes. This second shift was further confirmed by the findings from adults showing that they looked more at the eyes. Combined with the infant data at 12 months, the adult data suggest that the initial shift back to the eyes at 12 months of age in response to native speech is completed sometime after 12 months. As Figure 3 shows, the overall pattern of results in response to non-native speech was the same except that, as predicted, the 12-month-olds continued to look longer at the mouth.

Together, these data indicate clearly that infants become lipreaders when they begin learning how to talk and that MPN plays a key role in how they allocate their selective attention to the eyes and mouth of a talker. The clearest evidence of the effects of early experience and MPN can be seen at 12 months. Specifically, when infants were exposed to native audiovisual speech they no longer looked longer at the mouth but when they were exposed to non-native audiovisual speech they continued to look longer at the mouth. Presumably, they continued do so in this instance because the non-native speech was now unfamiliar to them and they were attempting to disambiguate it by taking advantage of the greater salience

of audiovisual speech at its source. This difference at 12 months of age is evidence of the differential effects of MPN (see Figure 3).

### Evolutionary roots of MPN

Given that MPN plays such a key role in human multisensory perceptual development, Zangenehpour, Ghazanfar, Lewkowicz, & Zatorre (2009) investigated whether this process might operate in the early perceptual development of other primate species as well. There were two reasons to expect that it might operate in other species. First, as noted earlier, perceptual narrowing plays a role in unisensory (i.e., auditory) responsiveness in birds (Gottlieb, 1991a, 1991b). Second, non-human primates, such as Macaque monkeys (*Macaca mulatta*, *Macaca fuscata*), capuchins (*Cebus apella*) and chimpanzees (*Pan troglodytes*) all can perceive the correspondence of facial and vocal expressions during communicative encounters (Adachi, Kuwahata, Fujita, Tomonaga, & Matsuzawa, 2006; Ghazanfar et al., 2007; Ghazanfar & Logothetis, 2003; Izumi & Kojima, 2004; Jordan, Brannon, Logothetis, & Ghazanfar, 2005; Parr, 2004). Although there is no direct evidence of perceptual narrowing in these species, the fact that they perceive face-voice relations raises the possibility that MPN might contribute to the emergence of this ability in non-human primates as well.

To test this possibility, Zangenehpour, Ghazanfar, Lewkowicz, & Zatorre (2009) used the multisensory paired-preference procedure and the same stimulus materials used by Lewkowicz & Ghazanfar (2006) to determine whether young vervet monkeys (*Chlorocebus pygerythrus*), who had no prior experience with rhesus monkeys, also might exhibit MPN in response to rhesus faces and vocalizations. The vervets ranged in age between 23 and 65 weeks (~6 to 16 months) and, thus, were well beyond the age when MPN would have been expected to be complete. Despite this, and in contrast to human infants, results revealed that the vervets matched the rhesus monkey faces and vocalizations. Unlike human infants, however, matching was characterized by greater looking at the non-matching face. Follow-up experiments revealed that this was due to the fact that the matching face-vocalization combination was more perceptually salient than the non-matching one, that this induced a fear reaction, and that because of this the animals looked more at the non-matching face-vocalization combination. This conclusion was supported by two findings. First, when the affective value of the audible call was eliminated by replacing the natural vocalization with a complex tone whose onsets and offsets corresponded to those of the matching visible call, the vervets now looked more at the matching visible call. Second, an analysis of pupillary responses (a measure of affective responsiveness) revealed that the pupils of vervets who were exposed to the matching natural face-vocalization combination were more dilated than they were when the vervets were exposed to the face-tone combination.

Overall, the findings from this study indicated that young vervet monkeys exhibit cross-species multisensory matching later in development than human infants do. This absence of MPN in vervets may be partly due to the relatively greater neural maturity of monkeys compared to humans. Monkeys possess approximately 65% of their adult brain size at birth whereas human infants only possess around 25% of their adult brain size at birth (Malkova, Heuer, & Saunders, 2006; Sacher & Staffeldt, 1974). Also, the fiber tracts in monkey brains are more myelinated than in human brains at the same postnatal age (Gibson, 1991; Malkova et al., 2006). Because of this greater neural maturity, non-human primates either may not be as open to the effects of early sensory experience and may need more time to incorporate the effects of experience, or they may be closed to the effects of experience altogether. The former is more likely because prior studies have found that experience does affect the development of unisensory and multisensory responsiveness in vervets and in other Old World monkeys (Lewkowicz & Ghazanfar, 2009). Therefore, MPN may operate in Old

World monkeys but on a longer time scale than in humans. This conjecture remains to be tested.

## Conclusions

Perceptual narrowing is a ubiquitous developmental process that reflects the effects of early experience and contributes to the development of perceptual specialization and expertise in early life. The initial discovery of a decline in responsiveness to non-native phonemes in the early 1980's demonstrated that narrowing contributes to perceptual development in human infants. This led to a burgeoning of interest in this topic and has now produced a substantial body of findings. They have shown that narrowing occurs in the auditory and visual modalities during infancy and that this is a domain-general process that leads to a decline in responsiveness to non-native phonemes, faces of other species and other races, and non-native musical meters. Our recent discovery of MPN and the various findings documenting its existence have shown that narrowing is a pan-sensory process as well.

To date, our findings have shown that MPN contributes to the development of native multisensory categories that include vocalizing faces, audiovisual speech, and audiovisual language identity and that it plays an important role in the allocation of selective attention in older infants' response to audiovisual speech. In addition, our findings from studies of developing vervet monkeys suggest that MPN may be limited to humans, although this is still very much an open question because there are good reasons to suspect that MPN may simply take longer to manifest itself in non-human primates (Lewkowicz & Ghazanfar, 2009)

Overall, it is now clear that early experience plays a key role in the developmental broadening of multisensory perceptual functions and that it continues to play an important role in developmental broadening of some multisensory functions past infancy. So far, it appears that experience with native inputs also contributes to MPN, but unlike the case of developmental broadening of multisensory functions, it seems that experience exerts its principal effects on MPN during infancy. The latter conclusion should be treated with some caution, however, until additional studies are carried out. For example, studies in which infants are exposed to non-native inputs (e.g., monkey faces and vocalizations or of people producing non-native audiovisual speech) during the period of narrowing are needed to determine whether MPN can be delayed. If it can be delayed, then this would provide direct evidence that experience plays a role in MPN. In addition, studies in which experience with non-native inputs is provided after narrowing has been completed would provide information on whether MPN is restricted to a sensitive period. If post-narrowing experience with non-native inputs can no longer restore the broad perceptual tuning normally observed earlier in development then this would be evidence that MPN is restricted to a sensitive period. Currently, no data on whether MPN is restricted to a sensitive period are available.

The findings to date also make it clear that perceptual broadening and narrowing work hand-in-hand. Perceptual broadening strengthens sensitivity to various perceptual attributes but only those that represent the native ecology. That is, as infants acquire increasing experience with their native world and as their nervous system grows, they gradually discover increasingly greater and more differentiated perceptual structure that is limited to their typical ecological niche. The developmental "problem" is that the nervous system is immature and relatively inexperienced at birth. Because of this, young infants are so broadly perceptually tuned that they tend to respond indiscriminately to all perceptual inputs regardless of whether they represent their ecological niche or not. In the case of multisensory perception, the problem is that young infants perceive very broad and poorly defined multisensory categories. This causes them to bind faces and voices regardless of

their species and to bind human faces and voices uttering native as well as non-native speech. MPN helps to further fine-tune the perceptual system by incorporating the effects of nearly exclusive exposure to native stimulus categories. In other words, perceptual narrowing is a process that solves the problem of excessively broad tuning due to initial structural and functional immaturity. Ultimately, the combined effect of broadening and narrowing processes is that infants' perceptual sensitivity and responsiveness increases to native categories while it decreases to non-native categories. This, in turn, leads to the emergence of initial perceptual specialization and expertise which enables infants to process native perceptual information in the most efficient and adaptive manner. A currently open question is whether the maintenance of the sort of multisensory tuning acquired during infancy depends on continued exposure to the native environment after the first year of life. The most likely answer is that it does given that experience continues to play an important role throughout childhood.

Figure 4 illustrates the operation of broadening and narrowing processes as well as the period when MPN occurs. The stippled cone that grows from birth into adulthood represents increasing specialization and expertise while the embedded shapes represent the approximate temporal location of MPN during development and the subsequent growth in multisensory processing capacity. Importantly, because MPN is an extension of unisensory narrowing, the embedded shapes can just as well represent unisensory narrowing. This, in turn, can be used to illustrate the fact that the length of the sensitive periods for responsiveness to different types of information differ. For example, as the inset in Figure 4 shows, the sensitive period for responsiveness to vowels closes earlier (Kuhl et al., 1992) than for consonants (Polka & Werker, 1994). This makes it possible that the sensitive periods for MPN differ across different tasks and domains but, as indicated earlier, it is currently not known whether MPN is restricted by one or more sensitive periods. Finally, it is currently not known whether MPN reflects unisensory narrowing or whether it reflects a unique multisensory process. So far, however, at least in the case of moving/vocalizing faces from other species, MPN has been found in spite of the absence of unisensory narrowing (Lewkowicz et al., 2008).

It is important to emphasize that characterization of the process underlying perceptual narrowing is key to understanding it. Some have characterized the outcome of narrowing as a loss of perceptual function (for examples see (Fair et al., 2012)) and the explicit assumption that lack of relevant experience leads to the loss. This is incorrect on several grounds. First, absence of relevant experience does not necessarily lead to a decline in discriminative ability (Sundara et al., 2006). Second, increased study time during a non-native face discrimination task can facilitate discrimination in infants who otherwise have narrowed (Fair et al., 2012). Third, auditory and visual narrowing can be prevented and/or reversed through "training" with non-native inputs during the sensitive period (Anzures et al., 2012; Hannon & Trehub, 2005b; Kuhl et al., 2003; Pascalis et al., 2005; Scott & Monesson, 2009). Thus, the perceptual system continues to be plastic even after narrowing has initially occurred. Of course, this is not to say that plasticity does not decrease as development progresses. Plasticity does decrease and, once it has, it is more difficult to re-establish responsiveness to non-native inputs. For example, 12-month-olds' responsiveness to non-Western musical rhythms can be restored but adults' responsiveness cannot be restored by the same interventions (Hannon & Trehub, 2005b). Similarly, adults who learn a foreign language late in life speak with an accent which is very difficult to undo (Flege, 1999). Overall, then, sensitivity to those categories of perceptual information that are either never or rarely experienced does not narrow to zero. Instead, daily experience with typically experienced perceptual categories increases discriminative capacity for those categories whereas discriminative capacity for inexperienced categories decreases (Lee, Anzures, Quinn, Pascalis, & Slater, 2011). Therefore, narrowing represents developmental regression

of an initially broad but poorly defined perceptual category due to a re-organization of underlying mechanisms rather than to a loss of perceptual sensitivity (Werker & Tees, 2005).

If narrowing is viewed as regression of an initially broadly tuned perceptual system then how might the nervous system reflect such regression? The dominant view has been that narrowing is due to the pruning of “exuberant” neural connections which are typically found in the young developing nervous system and which represent the initial growth of relatively diffuse and global neural networks (Cowan, Fawcett, O’Leary, & Stanfield, 1984; Low & Cheng, 2006). Presumably, as those connections are eliminated through a process of experience-dependent pruning the remaining networks become more modularized and are used to mediate responsiveness to categories of information that represent developing infants’ current environment (Scott, Pascalis, & Nelson, 2007; Spector & Maurer, 2009). As noted by Lewkowicz and Ghazanfar (2009), however, this view is problematic because extra synapses are not widespread throughout the developing brain early in life (Purves, White, & Riddle, 1996; Quartz & Sejnowski, 1997). Moreover, even though neural pruning does occur in early life, the dominant process is neural growth and proliferation resulting in an explosive growth of new synaptic connections throughout the nervous system. As this occurs, those connections are continually exposed to the effects of early experience and are modified by it. This is true in the developing visual system (Li, Van Hooser, Mazurek, White, & Fitzpatrick, 2008) as it is in the developing auditory system (Roberts, Tschida, Klein, & Mooney, 2010). For example, in the latter case, the effects of learning the species-specific song in juvenile zebra finches are reflected at the neural level in a decrease in the rate of turnover of dendritic spines – the main site of excitatory synaptic transmission in vertebrate brains - and in the strengthening rather than elimination of synapses. In other words, a constructive neural process reflects the changes taking place during periods of neural and behavioral plasticity.

Evidence also shows that constructive processes are involved in narrowing *per se*. One study (Scott & Monesson, 2010) compared 9-month-old infants’ event-related potentials (ERPs) to upright and inverted monkey faces following 3-month individualized training with monkey faces (i.e., each of six monkey faces had a specific name), category-level training (all faces were called monkey), or exposure training (no label was given to the faces). Only infants who received individualized training exhibited evidence of holistic processing by showing differential responsiveness to upright versus inverted monkey faces. Crucially, none of the groups exhibited an inversion effect prior to training at six months of age indicating that there is no “loss” of discriminative ability during the period when narrowing of responsiveness to non-native faces normally occurs. As a result, Scott and Monesson concluded that narrowing of responsiveness to non-native faces is most likely due to an increase in synaptic efficacy rather than synaptogenesis and pruning. A second study (Grossmann, Missana, Friederici, & Ghazanfar, 2012) presented congruent and incongruent auditory and visual monkey vocalizations (e.g., auditory coo+visual coo vs. auditory coo +visual grunt) and human versions of same where humans mimicked the monkey vocalizations to 4- and 8-month-old infants. Consistent with the narrowing effects found for monkey faces and voices found by Lewkowicz & Ghazanfar (2006), the ERP results showed that the 4-month-olds distinguished between the congruent and the incongruent faces and voices regardless of species, whereas the 8-month-olds only responded to the congruency of human faces and voices. Crucially, occipital and frontal brain processes and their functional connectivity became more sensitive to the congruence of human faces and voices relative to monkey faces and voices. In other words, a constructive/progressive neurodevelopmental event, rather than a destructive one, was associated with narrowing.

The re-organization that occurs as a result of MPN is, in part, the result of a decline in reliance on the relatively primitive, low-level, response system that detects A–V temporal synchrony relations. Crucially, this low-level system enables younger infants to respond to a wider range of stimulation from a greater set of categories only because it cannot distinguish between these categories. Nonetheless, the relatively immature system is ontogenetically best adapted to the exigencies of its environment given its structural and functional capacity. The reason that the eventual re-organization occurs is because the low-level perceptual mechanism is inadequate for the recognition and integration of higher-level attributes of multisensory objects and events. The developmental push for the re-organization results from a combination of factors including neural growth, differentiation, early plasticity, and the perceptually rich everyday environment that continually challenges infants to progress to the next level of processing. In this way, as re-organization takes place, infants begin to detect higher-level features of multisensory objects and events and, in so doing, begin to perceive their semantic, indexical, and social attributes (among others). Importantly, as in the case of unisensory narrowing, the earlier mechanism is not lost but, rather, it continues to play a role in responsiveness except that it is no longer the primary mechanism.

In conclusion, there is little doubt that infancy is a time of tremendous change. This change is due to the growth of a highly plastic nervous system that is open to perceptual experience and the concurrent growth of new behavioral capacities. As shown here, experience plays a central role in the changes observed during infancy. Furthermore, it is now clear that the conventional effects of experience (i.e., broadening) as well as the paradoxical effects of experience (i.e., narrowing) contribute to perceptual and cognitive development. Although the latter effects are currently less understood, once they are it will be possible to determine how broadening and narrowing contribute to the emergence of unisensory and multisensory perceptual expertise and the growth of general knowledge.

## Acknowledgments

The author was supported by grant R01HD057116 from the Eunice Kennedy Shriver National Institute Of Child Health & Human Development. The content is solely the responsibility of the author and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute Of Child Health & Human Development or the National Institutes of Health.

## References

- Adachi I, Kuwahata H, Fujita K, Tomonaga M, Matsuzawa T. Japanese macaques form a cross-modal representation of their own species in their first year of life. *Primates*. 2006; 47:350–354. [PubMed: 16636747]
- Amedi A, Merabet LB, Bermpohl F, Pascual-Leone A. The occipital cortex in the blind: lessons about plasticity and vision. *Current Directions in Psychological Science*. 2005; 14:306–311.
- Anzures G, Wheeler A, Quinn PC, Pascalis O, Slater AM, Heron-Delaney M, Lee K. Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology*. 2012; 112:484–495. [PubMed: 22625845]
- Bahrick LE. Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child development*. 1988; 59:197–209. [PubMed: 3342712]
- Bahrick LE, Hernandez-Reif M, Flom R. The Development of Infant Learning About Specific Face-Voice Relations. *Developmental Psychology*. 2005; 41:541–552. [PubMed: 15910161]
- Bahrick LE, Lickliter R. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*. 2000; 36:190–201. [PubMed: 10749076]
- Bavelier D, Neville HJ. Cross-modal plasticity: where and how? *Nat Rev Neurosci*. 2002; 3:443–452. doi: 10.1038/nrn848 nrn848 [pii]. [PubMed: 12042879]

- Bedny M, Pascual-Leone A, Dodell-Feder D, Fedorenko E, Saxe R. Language processing in the occipital cortex of congenitally blind adults. *Proceedings of the National Academy of Sciences*. 2011; 108:4429–4434.
- Best CT, McRoberts GW, LaFleur R, Silver-Isenstadt J. Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior & Development*. 1995; 18:339–350.
- Birch HG, Lefford A. Intersensory development in children. *Monographs of the Society for Research in Child Development*. 1963; 25:1–48. [PubMed: 14058560]
- Botuck S, Turkewitz G. Intersensory functioning: Auditory-visual pattern equivalence in younger and older children. *Developmental Psychology*. 1990; 26:115.
- Bradley R, Mistretta C. Fetal sensory receptors. *Physiological Reviews*. 1975; 55:352–382. [PubMed: 1098067]
- Bremner, AJ.; Lewkowicz, DJ.; Spence, C. *Multisensory Development*. Oxford: Oxford University Press; 2012.
- Carmichael, L. The onset and early development of behavior. In: Carmichael, L., editor. *Manual of child psychology*. New York: John Wiley & Sons, Inc.; 1946. p. 43-166.
- Cassia VM, Turati C, Simion F. Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science*. 2004; 15:379–383. [PubMed: 15147490]
- Chandrasekaran C, Trubanova A, Stillitano S, Caplier A, Ghazanfar AA. The natural statistics of audiovisual speech. *PLoS Computational Biology*. 2009; 5 doi: e1000436.
- Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Näätänen R. Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*. 1998; 1:351–353.
- Colombo J. The development of visual attention in infancy. *Annual Review of Psychology*. 2001; 52:337–367.
- Cowan WM, Fawcett JW, O'Leary DD, Stanfield BB. Regressive events in neurogenesis. *Science*. 1984; 225:1258–1265. [PubMed: 6474175]
- de Boysson-Bardies B, Hallé P, Sagart L, Durand C. A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*. 1989; 16:1–17. [PubMed: 2925806]
- DeCasper AJ, Spence MJ. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development*. 1986; 9:133–150.
- Dixon NF, Spitz LT. The detection of auditory visual desynchrony. *Perception*. 1980; 9:719–721. [PubMed: 7220244]
- Dodd B. Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*. 1979; 11:478–484. [PubMed: 487747]
- Fair J, Flom R, Jones J, Martin J. Perceptual Learning: 12-Month-Olds' Discrimination of Monkey Faces. *Child development*. 2012; 83:1996–2006. [PubMed: 22803582]
- Flege, JE. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 1999. Age of learning and second language *speech Second language acquisition and the Critical Period Hypothesis*; p. 101-131.
- Flom R, Bahrick LE. The Development of Infant Discrimination of Affect in Multimodal and Unimodal Stimulation: The Role of Intersensory Redundancy. *Developmental Psychology*. 2007; 43:238–252. [PubMed: 17201522]
- Ghazanfar A, Tureson H, Maier J, van Dinther R, Patterson R, Logothetis N. Vocal-tract resonances as indexical cues in rhesus monkeys. *Current Biology*. 2007; 17:425–430. [PubMed: 17320389]
- Ghazanfar AA, Logothetis NK. Facial expressions linked to monkey calls. *Nature*. 2003; 423:937–938. [PubMed: 12827188]
- Gibson, EJ. *Principles of perceptual learning and development*. New York: Appleton; 1969.
- Gibson, JJ. *The senses considered as perceptual systems*. Boston: Houghton-Mifflin; 1966.
- Gibson, KR. Myelination and behavioral development: A comparative perspective on questions of neoteny, altriciality and intelligence. In: Gibson, KR.; Petersen, AC., editors. *Brain maturation and cognitive development: comparative and cross-cultural perspectives*. New York: Aldine de Gruyter; 1991. p. 29-63.

- Goldstein MH, Schwade JA. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*. 2008; 19:515. [PubMed: 18466414]
- Gottfried AW, Rose SA, Bridger WH. Cross-modal transfer in human infants. *Child development*. 1977; 48:118–123. [PubMed: 844349]
- Gottlieb G. Ontogenesis of sensory function in birds and mammals. In: Tobach, E.; Aronson, LR.; Shaw, E., editors. *The biopsychology of development*. New York: Academic Press; 1971. p. 67-128.
- Gottlieb G. Experiential canalization of behavioral development: Results. *Developmental Psychology*. 1991a; 27:35–39.
- Gottlieb G. Experiential canalization of behavioral development: Theory. *Developmental Psychology*. 1991b; 27:4–13.
- Gottlieb G. Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesis. *Psychological Review*. 1998; 105:792–802. [PubMed: 9830380]
- Grossmann T, Missana M, Friederici AD, Ghazanfar AA. Neural correlates of perceptual narrowing in cross-species face-voice matching. *Developmental Science*. 2012; 15:830–839. [PubMed: 23106737]
- Haith MM, Bergman T, Moore MJ. Eye contact and face scanning in early infancy. *Science*. 1977; 198:853–855. [PubMed: 918670]
- Hannon EE, Trehub SE. Metrical categories in infancy and adulthood. *Psychological Science*. 2005a; 16:48–55. [PubMed: 15660851]
- Hannon EE, Trehub SE. Tuning in to musical rhythms: infants learn more readily than adults. *Proceedings of the National Academy of Sciences USA*. 2005b; 102:12639–12643.
- Hillock-Dunn A, Wallace MT. Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*. 2012; 15:688–696. [PubMed: 22925516]
- Holt, EB. *Animal drive and the learning process*. Vol. 1. New York: Holt; 1931.
- Hunnis S, Geuze RH. Developmental Changes in Visual Scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*. 2004; 6:231–255.
- Izumi A, Kojima S. Matching vocalizations to vocalizing faces in a chimpanzee (*Pan troglodytes*). *Animal Cognition*. 2004; 7:179–184. [PubMed: 15015035]
- Jaime M, Lickliter R. Prenatal exposure to temporal and spatial stimulus properties affects postnatal responsiveness to spatial contiguity in bobwhite quail chicks. *Developmental Psychobiology*. 2006; 48:233. [PubMed: 16568412]
- Jordan K, Brannon E, Logothetis N, Ghazanfar AA. Monkeys match the number of voices they hear to the number of faces they see. *Current Biology*. 2005; 15:1034–1038. [PubMed: 15936274]
- Jusczyk PW, Cutler A, Redanz NJ. Infants' preference for the predominant stress patterns of English words. *Child development*. 1993; 64:675–687. [PubMed: 8339688]
- Jusczyk PW, Luce PA, Charles-Luce J. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory & Language*. 1994; 33:630–645.
- Kahana-Kalman R, Walker-Andrews AS. The role of person familiarity in young infants' perception of emotional expressions. *Child development*. 2001; 72:352–369. [PubMed: 11333071]
- Kelly DJ, Liu S, Lee K, Quinn PC, Pascalis O, Slater AM, Ge L. Development of the other-race effect during infancy: evidence toward universality? *Journal of Experimental Child Psychology*. 2009; 104:105–114. [PubMed: 19269649]
- Kelly DJ, Quinn PC, Slater A, Lee K, Ge L, Pascalis O. The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*. 2007; 18:1084–1089. [PubMed: 18031416]
- Kenny PA, Turkewitz G. Effects of unusually early visual stimulation on the development of homing behavior in the rat pup. *Dev Psychobiol*. 1986; 19:57–66. [PubMed: 3699252]
- King AJ, Hutchings ME, Moore DR, Blakemore C. Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature*. 1988; 332:73–76. [PubMed: 3347247]



- Knudsen EI, Brainard MS. Visual instruction of the neural map of auditory space in the developing optic tectum. *Science*. 1991; 253:85–87. [PubMed: 2063209]
- Kuhl PK, Meltzoff AN. The bimodal perception of speech in infancy. *Science*. 1982; 218:1138–1141. [PubMed: 7146899]
- Kuhl PK, Meltzoff AN. Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*. 1996; 100:2425–2438. [PubMed: 8865648]
- Kuhl PK, Stevens E, Hayashi A, Deguchi T, Kiritani S, Iverson P, Liu HM. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*. 2006; 9:F13–F21. [PubMed: 16472309]
- Kuhl PK, Tsao FM, Liu HM. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences USA*. 2003; 100:9096–9101.
- Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*. 1992; 255:606–608. [PubMed: 1736364]
- Kuhl PK, Williams KA, Meltzoff AN. Cross-modal speech perception in adults and infants using nonspeech auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*. 1991; 17:829–840. [PubMed: 1834794]
- Kuo, ZY. *The dynamics of behavior development: An epigenetic view*. New York: Plenum; 1967.
- Lee, K.; Anzures, G.; Quinn, PC.; Pascalis, O.; Slater, A. Development of face processing expertise. In: Rhodes, G.; Calder, AJ.; Johnson, MH.; Haxby, JV., editors. *Oxford handbook of face perception*. Oxford University Press; 2011. p. 753-778.
- Lehrman DS. A critique of Konrad Lorenz's theory of instinctive behavior. *Quarterly Review of Biology*. 1953; 28:337–363. [PubMed: 13121237]
- Lehrman, DS. Semantic and conceptual issues in the nature-nurture problem. In: Aronson, LR.; Lehrman, DS.; Tobach, E.; Rosenblatt, JS., editors. *Development and evolution of behavior*. San Francisco: W. H. Freeman; 1970. p. 17-52.
- Lewis R, Noppeney U. Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *The Journal of Neuroscience*. 2010; 30:12329–12339. [PubMed: 20844129]
- Lewkowicz DJ. Developmental changes in infants' bisensory response to synchronous durations. *Infant Behavior & Development*. 1986; 9:335–353.
- Lewkowicz DJ. Infants' response to temporally based intersensory equivalence: The effect of synchronous sounds on visual preferences for moving stimuli. *Infant Behavior & Development*. 1992a; 15:297–324.
- Lewkowicz DJ. Infants' responsiveness to the auditory and visual attributes of a sounding/moving stimulus. *Perception & Psychophysics*. 1992b; 52:519–528. [PubMed: 1437484]
- Lewkowicz, DJ. Development of intersensory perception in human infants. In: Lewkowicz, DJ.; Lickliter, R., editors. *The development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1994a. p. 165-203.
- Lewkowicz DJ. Limitations on infants' response to rate-based auditory-visual relations. *Developmental Psychology*. 1994b; 30:880–892.
- Lewkowicz DJ. Perception of auditory-visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception & Performance*. 1996; 22:1094–1106. [PubMed: 8865617]
- Lewkowicz DJ. The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*. 2000a; 126:281–308. [PubMed: 10748644]
- Lewkowicz DJ. Infants' perception of the audible, visible and bimodal attributes of multimodal syllables. *Child development*. 2000b; 71:1241–1257. [PubMed: 11108094]
- Lewkowicz DJ. Infant perception of audio-visual speech synchrony. *Developmental Psychology*. 2010; 46:66–77. [PubMed: 20053007]
- Lewkowicz DJ. The biological implausibility of the nature-nurture dichotomy and what it means for the study of infancy. *Infancy*. 2011; 16:331–367. [PubMed: 21709807]

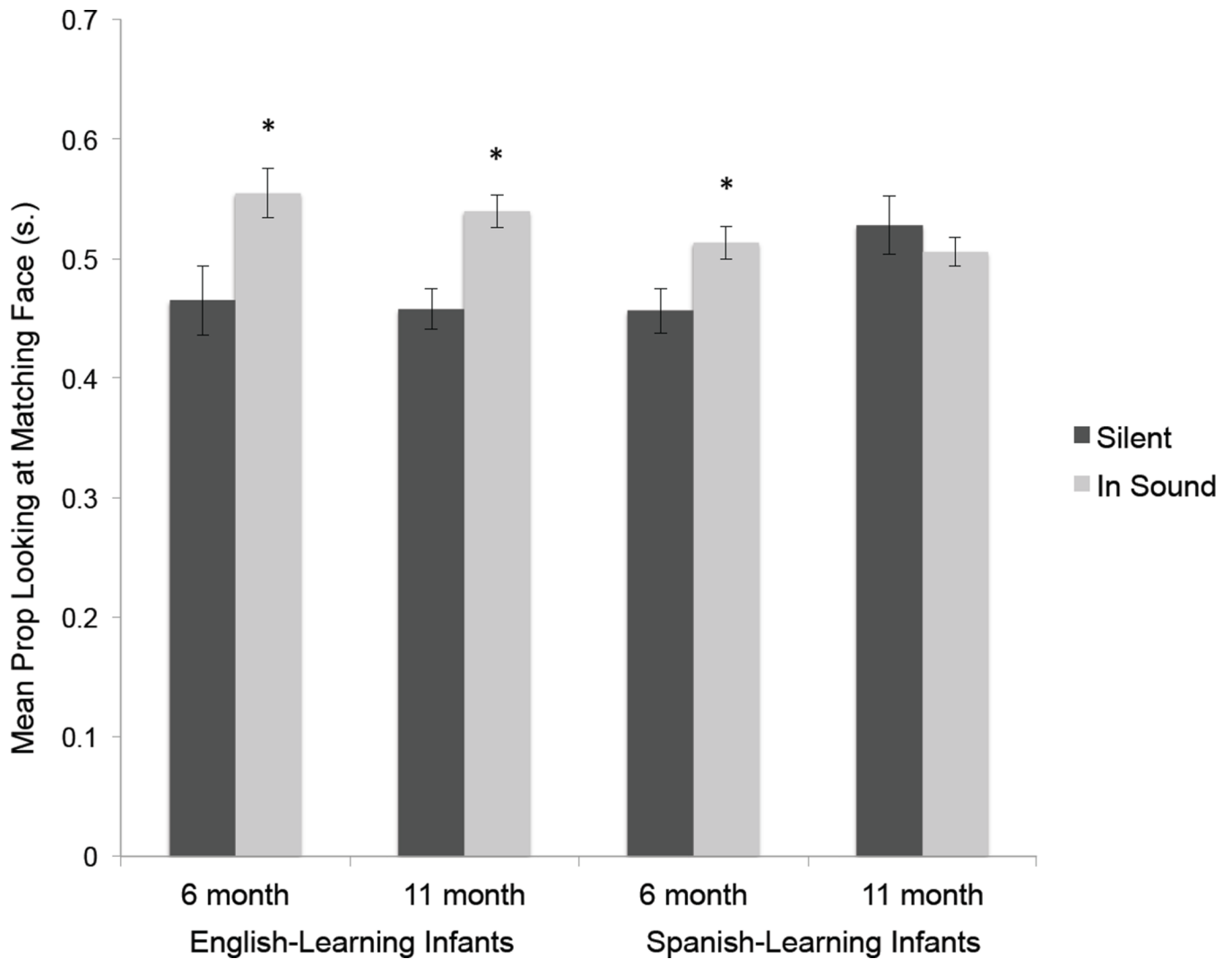
- Lewkowicz DJ, Flom R. The audio-visual temporal binding window narrows in early childhood. *Child development*. 2013
- Lewkowicz DJ, Ghazanfar AA. The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy Sciences USA*. 2006; 103:6771–6774.
- Lewkowicz DJ, Ghazanfar AA. The emergence of multisensory systems through perceptual narrowing. *Trends in cognitive sciences*. 2009; 13:470–478. [PubMed: 19748305]
- Lewkowicz DJ, Ghazanfar AA. The development of the uncanny valley in infants. *Developmental Psychobiology*. 2012; 54:124–132. [PubMed: 21761407]
- Lewkowicz DJ, Hansen-Tift AM. Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*. 2012; 109:1431–1436.
- Lewkowicz DJ, Leo I, Simion F. Intersensory perception at birth: Newborns match non-human primate faces & voices. *Infancy*. 2010; 15:46–60.
- Lewkowicz, DJ.; Lickliter, R., editors. *The development of intersensory perception: Comparative perspectives*. Hillsdale, N.J: Lawrence Erlbaum Associates, Inc.; 1994.
- Lewkowicz DJ, Pons F. Recognition of a modal language identity emerges in infancy. *International Journal of Behavioral Development*. 2013; 37(2):90–94.
- Lewkowicz DJ, Sowinski R, Place S. The decline of cross-species intersensory perception in human infants: Underlying mechanisms and its developmental persistence. *Brain Research*. 2008; 1242:291–302. [PubMed: 18486112]
- Lewkowicz DJ, Turkewitz G. Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Developmental Psychology*. 1980; 16:597–607.
- Lewkowicz DJ, Turkewitz G. Intersensory interaction in newborns: modification of visual preferences following exposure to sound. *Child Dev*. 1981; 52:827–832. [PubMed: 7285654]
- Li Y, Van Hooser SD, Mazurek M, White LE, Fitzpatrick D. Experience with moving visual stimuli drives the early development of cortical direction selectivity. *Nature*. 2008; 456:952–956. [PubMed: 18946471]
- Lickliter R. The integrated development of sensory organization. *Clinics in perinatology*. 2011; 38:591. [PubMed: 22107892]
- Lickliter R, Bahrick LE. The development of infant intersensory perception: advantages of a comparative convergent-operations approach. *Psychological Bulletin*. 2000; 126:260–280. [PubMed: 10748643]
- Lickliter, R.; Banker, H. Prenatal components of intersensory development in precocial birds. In: Lewkowicz, DJ.; Lickliter, R., editors. *Development of intersensory perception: Comparative perspectives*. Norwood N.J: Lawrence Erlbaum Associates, Inc.; 1994. p. 59-80.
- Lickliter R, Hellewell TB. Contextual determinants of auditory learning in bobwhite quail embryos and hatchlings. *Developmental Psychobiology*. 1992; 25:17–31. [PubMed: 1740227]
- Lickliter R, Lewkowicz DJ, Columbus RF. Intersensory experience and early perceptual development: The role of spatial contiguity in bobwhite quail chicks' responsiveness to multimodal maternal cues. *Developmental Psychobiology*. 1996; 29:403–416. [PubMed: 8809492]
- Low LK, Cheng HJ. Axon pruning: an essential step underlying the developmental plasticity of neuronal connections. *Philos Trans R Soc Lond B Biol Sci*. 2006; 361:1531–1544. [PubMed: 16939973]
- Maier, NRF.; Schneirla, TC. *Principles of animal psychology*. New York: Dover Publications; 1964.
- Malkova L, Heuer E, Saunders RC. Longitudinal magnetic resonance imaging study of rhesus monkey brain development. *European Journal of Neuroscience*. 2006; 24:3204–3212. [PubMed: 17156381]
- Marks, L. *The unity of the senses*. New York: Academic Press; 1978.
- Marlier L, Schaal B. Human newborns prefer human milk: conspecific milk odor is attractive without postnatal exposure. *Child development*. 2005; 76:155–168. [PubMed: 15693764]
- McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature*. 1976; 264:229–239.
- Mehler J, Juszyk PW, Lambertz G, Halsted N, Bertoincini J, Amiel-Tison C. A precursor of language acquisition in young infants. *Cognition*. 1988; 29:143–178. [PubMed: 3168420]

- Merabet LB, Pascual-Leone A. Neural reorganization following sensory loss: the opportunity of change. *Nat Rev Neurosci*. 2010; 11:44–52. doi: nrm2758 [pii] 10.1038/nrm2758. [PubMed: 19935836]
- Merin N, Young GS, Ozonoff S, Rogers SJ. Visual Fixation Patterns during Reciprocal Social Interaction Distinguish a Subgroup of 6-Month-Old Infants At-Risk for Autism from Comparison Infants. *J Autism Dev Disord*. 2007; 37:108–121. [PubMed: 17191096]
- Munhall, KG.; Vatikiotis-Bateson, E. Spatial and Temporal Constraints on Audiovisual Speech Perception. In: Calvert, GA.; Spence, C.; Stein, BE., editors. *The handbook of multisensory processes*. Cambridge, MA: MIT Press; 2004. p. 177-188.
- Nazzi T, Bertoncini J, Mehler J. Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception & Performance*. 1998; 24:756–766. [PubMed: 9627414]
- Nazzi T, Jusczyk PW, Johnson EK. Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory & Language*. 2000; 43:1–19.
- Neil PA, Chee-Ruiter C, Scheier C, Lewkowicz DJ, Shimojo S. Development of multisensory spatial integration and perception in humans. *Developmental Science*. 2006; 9:454–464. [PubMed: 16911447]
- Oppenheim, RW. Ontogenetic adaptations and retrogressive processes in the development of the nervous system and behavior: A neuroembryological perspective. In: Connolly, KJ.; Prechtl, HFR., editors. *Maturation and Development: Biological and Psychological Perspectives*. Lippincott: Philadelphia; 1981. p. 73-109.
- Parr LA. Perceptual biases for multimodal cues in chimpanzee (*Pan troglodytes*) affect recognition. *Animal Cognition*. 2004; 7:171–178. [PubMed: 14997361]
- Pascalis O, Haan Md, Nelson CA. Is face processing species-specific during the first year of life? *Science*. 2002; 296:1321–1323. [PubMed: 12016317]
- Pascalis O, Scott LS, Kelly DJ, Shannon RW, Nicholson E, Coleman M, Nelson CA. Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences USA*. 2005; 102:5297–5300.
- Patterson ML, Werker JF. Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior & Development*. 1999; 22:237–247.
- Patterson ML, Werker JF. Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental Child Psychology*. 2002; 81:93–115. [PubMed: 11741376]
- Patterson ML, Werker JF. Two-month-old infants match phonetic information in lips and voice. *Developmental Science*. 2003; 6:191–196.
- Piaget, J. *The origins of intelligence in children*. New York: International Universities Press; 1952.
- Polka L, Werker JF. Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*. 1994; 20:421–435. [PubMed: 8189202]
- Pons F, Bosch L. Stress pattern preference in Spanish-learning infants: The role of syllable weight. *Infancy*. 2010; 15:223–245.
- Pons F, Lewkowicz DJ, Soto-Faraco S, Sebastián-Gallés N. Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences USA*. 2009; 106:10598–10602.
- Poulin-Dubois D, Serbin LA, Kenyon B, Derbyshire A. Infants' intermodal knowledge about gender. *Developmental Psychology*. 1994; 30:436–442.
- Purves D, White LE, Riddle DR. Is neural development darwinian? *Trends in Neurosciences*. 1996; 19:460–464. [PubMed: 8931267]
- Putzar L, Goerendt I, Heed T, Richard G, Büchel C, Röder B. The neural basis of lip-reading capabilities is altered by early visual deprivation. *Neuropsychologia*. 2010; 48:2158–2166. [PubMed: 20385153]
- Putzar L, Goerendt I, Lange K, Rösler F, Röder B. Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience*. 2007; 10:1243–1245.

- Putzar L, Hötting K, Röder B. Early visual deprivation affects the development of face recognition and of audio-visual speech perception. *Restorative Neurology and Neuroscience*. 2010; 28:251–257. [PubMed: 20404412]
- Quartz SR, Sejnowski TJ. The neural basis of cognitive development: A constructivist manifesto. *Behavioral and brain sciences*. 1997; 20:537–556. [PubMed: 10097006]
- Roberts TF, Tschida KA, Klein ME, Mooney R. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature*. 2010; 463:948–952. [PubMed: 20164928]
- Romei V, Murray MM, Cappe C, Thut G. Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology*. 2009; 19:1799–1805. doi: S0960-9822(09)01707-2 [pii] 10.1016/j.cub.2009.09.027. [PubMed: 19836243]
- Rosenblum LD, Johnson JA, Saldana HM. Point-light facial displays enhance comprehension of speech in noise. *Journal of Speech & Hearing Research*. 1996; 39:1159–1170. [PubMed: 8959601]
- Ryan TA. Interrelations of the sensory systems in perception. *Psychological Bulletin*. 1940; 37:659–698.
- Sacher GA, Staffeldt EF. Relation of gestation time to brain weight for placental mammals: implications for the theory of vertebrate growth. *American Naturalist*. 1974; 108:593–615.
- Sai FZ. The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development*. 2005; 14:29–50.
- Sangrigoli S, de Schonen S. Recognition of own-race and other-race faces by three-month-old infants. *Journal of Child Psychology and Psychiatry*. 2004; 45:1219–1227. [PubMed: 15335342]
- Sangrigoli S, Pallier C, Argenti AM, Ventureyra VAG, de Schonen S. Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*. 2005; 16:440–444. [PubMed: 15943669]
- Scheier C, Lewkowicz DJ, Shimojo S. Sound induces perceptual reorganization of an ambiguous motion display in human infants. *Developmental Science*. 2003; 6:233–244.
- Schneirla TC. Behavioral development and comparative psychology. *The Quarterly Review of Biology*. 1966; 41:283–302. [PubMed: 5975996]
- Schorr EA, Fox NA, van Wassenhove V, Knudsen EI. Auditory-visual fusion in speech perception in children with cochlear implants. *Proc Natl Acad Sci U S A*. 2005; 102:18748–18750. doi: 0508862102 [pii] 10.1073/pnas.0508862102. [PubMed: 16339316]
- Scott LS, Monesson A. The origin of biases in face perception. *Psychological Science*. 2009; 20:676–680. [PubMed: 19422630]
- Scott LS, Monesson A. Experience-dependent neural specialization during infancy. *Neuropsychologia*. 2010; 48:1857–1861. [PubMed: 20153343]
- Scott LS, Pascalis O, Nelson CA. A domain general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*. 2007; 16:197–201. [PubMed: 21132090]
- Sebastián-Gallés N, Albareda-Castellot B, Weikum WM, Werker JF. A bilingual advantage in visual language discrimination in infancy. *Psychological Science*. 2012; 23:994–999. [PubMed: 22810164]
- Sharma J, Angelucci A, Sur M. Induction of visual orientation modules in auditory cortex. *Nature*. 2000; 404:841–847. [PubMed: 10786784]
- Slater A, Brown E, Badenoch M. Intermodal perception at birth: Newborn infants' memory for arbitrary auditory-visual pairings. *Early Development & Parenting*. 1997; 6:99–104.
- Slater A, Quinn PC, Brown E, Hayes R. Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory-visual pairings. *Developmental Science*. 1999; 2:333–338.
- Smotherman WP, Robinson SR. The prenatal origins of behavioral organization. *Psychological Science*. 1990; 1:97–106.
- Spector F, Maurer D. Synesthesia: a new approach to understanding the development of perception. *Developmental Psychology*. 2009; 45:175–189. [PubMed: 19210000]
- Stein, BE.; Meredith, MA. *The merging of the senses*. Cambridge, MA: The MIT Press; 1993.

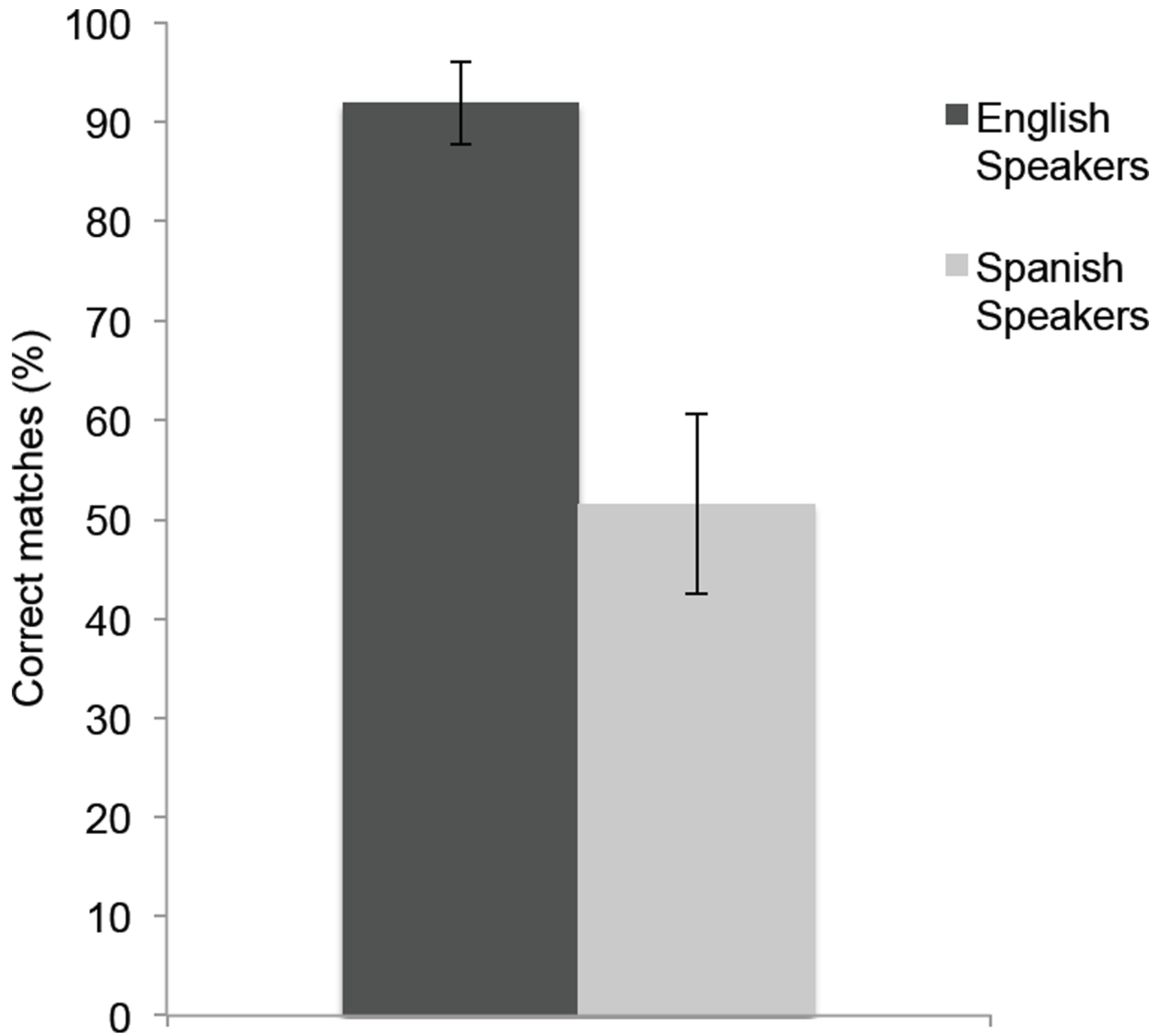
- Sumbly WH, Pollack I. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*. 1954; 26:212–215.
- Summerfield AQ. Use of visual information in phonetic perception. *Phonetica*. 1979; 36:314–331. [PubMed: 523520]
- Sundara M, Polka L, Genesee F. Language-experience facilitates discrimination of /d- ð/ in monolingual and bilingual acquisition of English. *Cognition*. 2006; 100:369–388. [PubMed: 16115614]
- Swingley D. 11-month-olds' knowledge of how familiar words sound. *Developmental Science*. 2005; 8:432–443. [PubMed: 16048516]
- Thelen, E.; Smith, LB. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press; 1994.
- Turkewitz, G. Sources of order for intersensory functioning. In: Lewkowicz, DJ.; Lickliter, R., editors. *The Development of Intersensory Perception Comparative Perspectives*. Hillsdale: L. Erlbaum; 1994. p. 3-17.
- Turkewitz G, Kenny PA. Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychobiology*. 1982; 15:357–368. [PubMed: 7106395]
- Turkewitz G, Kenny PA. The role of developmental limitations of sensory input on sensory/perceptual organization. *Journal of Developmental and Behavioral Pediatrics*. 1985; 6:302–306. [PubMed: 3905863]
- Vihman M, Nakai S, DePaolis R, Hallé P. The role of accentual pattern in early lexical representation. *Journal of Memory and Language*. 2004; 50:336–353.
- von Melchner L, Pallas SL, Sur M. Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*. 2000; 404:871–876. [PubMed: 10786793]
- Vouloumanos A, Druhen M, Hauser M, Huizink A. Five-month-old infants' identification of the sources of vocalizations. *Proceedings of the National Academy of Sciences*. 2009; 106:18867–18872.
- Waddington, CH. *The strategy of the genes*. London: Allen & Unwin; 1957.
- Walker AS. Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*. 1982; 33:514–535. [PubMed: 7097157]
- Walker-Andrews AS. Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology*. 1986; 22:373–377.
- Walker-Andrews AS. Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*. 1997; 121:437–456. [PubMed: 9136644]
- Walker-Andrews AS, Bahrack LE, Raglioni SS, Diaz I. Infants' bimodal perception of gender. *Ecological Psychology*. 1991; 3:55–75.
- Wallace MT, Stein BE. Development of multisensory neurons and multisensory integration in cat superior colliculus. *Journal of Neuroscience*. 1997; 17:2429–2444. [PubMed: 9065504]
- Wallace MT, Stein BE. Onset of cross-modal synthesis in the neonatal superior colliculus is gated by the development of cortical influences. *Journal of Neurophysiology*. 2000; 83:3578–3582. [PubMed: 10848574]
- Wallace MT, Stein BE. Sensory and multisensory responses in the newborn monkey superior colliculus. *Journal of Neuroscience*. 2001; 21:8886–8894. [PubMed: 11698600]
- Wallace MT, Stein BE. Early experience determines how the senses will interact. *Journal of Neurophysiology*. 2007; 97:921. [PubMed: 16914616]
- Walton GE, Bower TG. Amodal representations of speech in infants. *Infant Behavior & Development*. 1993; 16:233–243.
- Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastián-Gallés N, Werker JF. Visual language discrimination in infancy. *Science*. 2007; 316:1159. [PubMed: 17525331]
- Werker JF, Tees RC. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*. 1984; 7:49–63.

- Werker JF, Tees RC. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology Special Issue: Critical Periods Re-examined: Evidence from Human Sensory Development*. 2005; 46:233–234.
- Werner, H. *Comparative psychology of mental development*. New York: International Universities Press; 1973.
- Xu J, Yu L, Rowland BA, Stanford TR, Stein BE. Incorporating cross-modal statistics in the development and maintenance of multisensory integration. *The Journal of Neuroscience*. 2012; 32:2287–2298. [PubMed: 22396404]
- Yeung HH, Werker JF. Learning words' sounds before learning how words sound: 9-Month-olds use distinct objects as cues to categorize speech information. *Cognition*. 2009; 113:234–243. [PubMed: 19765698]
- Yoshida KA, Pons F, Maye J, Werker JF. Distributional phonetic learning at 10 months of age. *Infancy*. 2010; 15:420–433.
- Zangenehpour S, Ghazanfar AA, Lewkowicz DJ, Zatorre RJ. Heterochrony and cross-species intersensory matching by infant vervet monkeys. *PLoS ONE*. 2009; 4:e4302. [PubMed: 19172998]



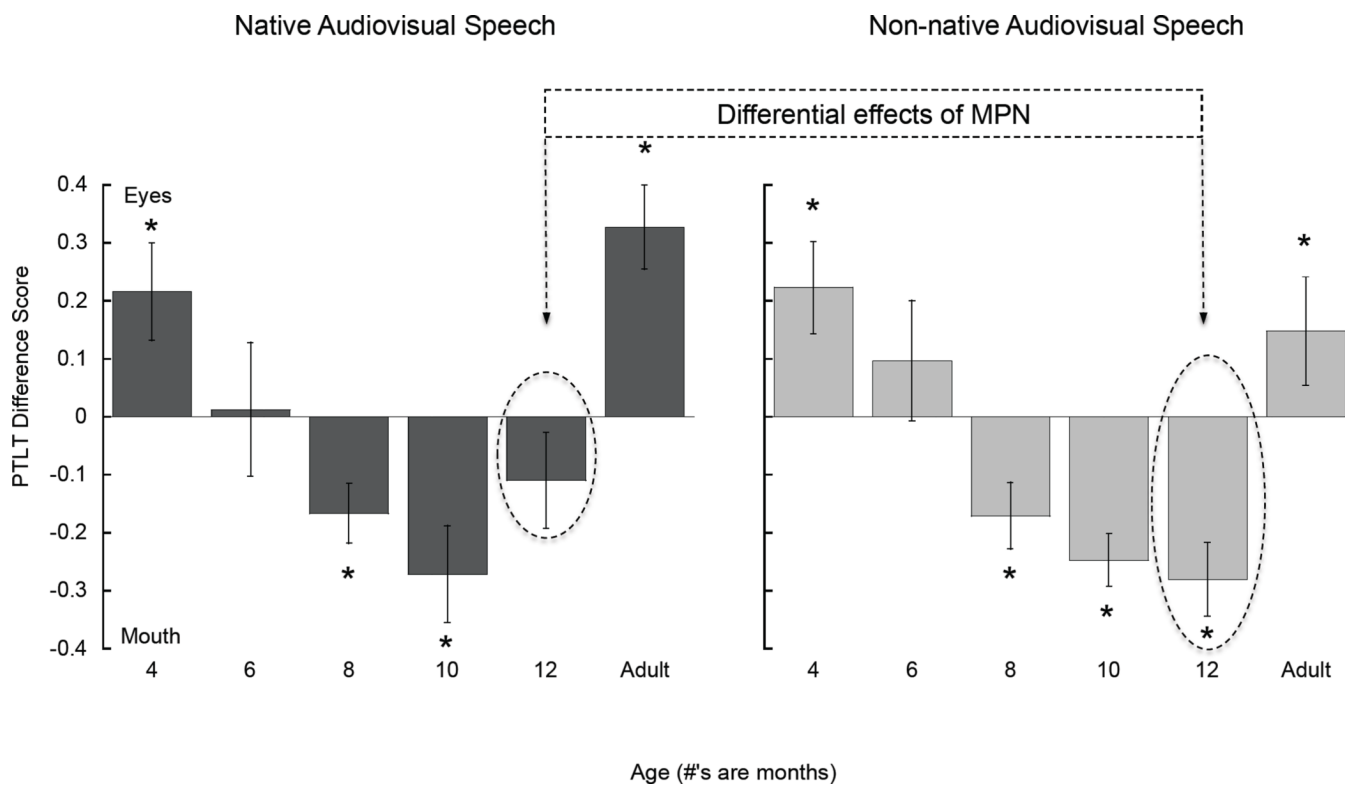
**Figure 1.**

Mean proportion of looking at the face that was seen silently producing the syllable /ba/ or /va/ prior to familiarization (Silent condition) with the corresponding audible /ba/ or /va/ and following it (In Sound condition). Asterisks indicate a statistically significant difference across the two conditions. Error bars represent the standard error of the mean.



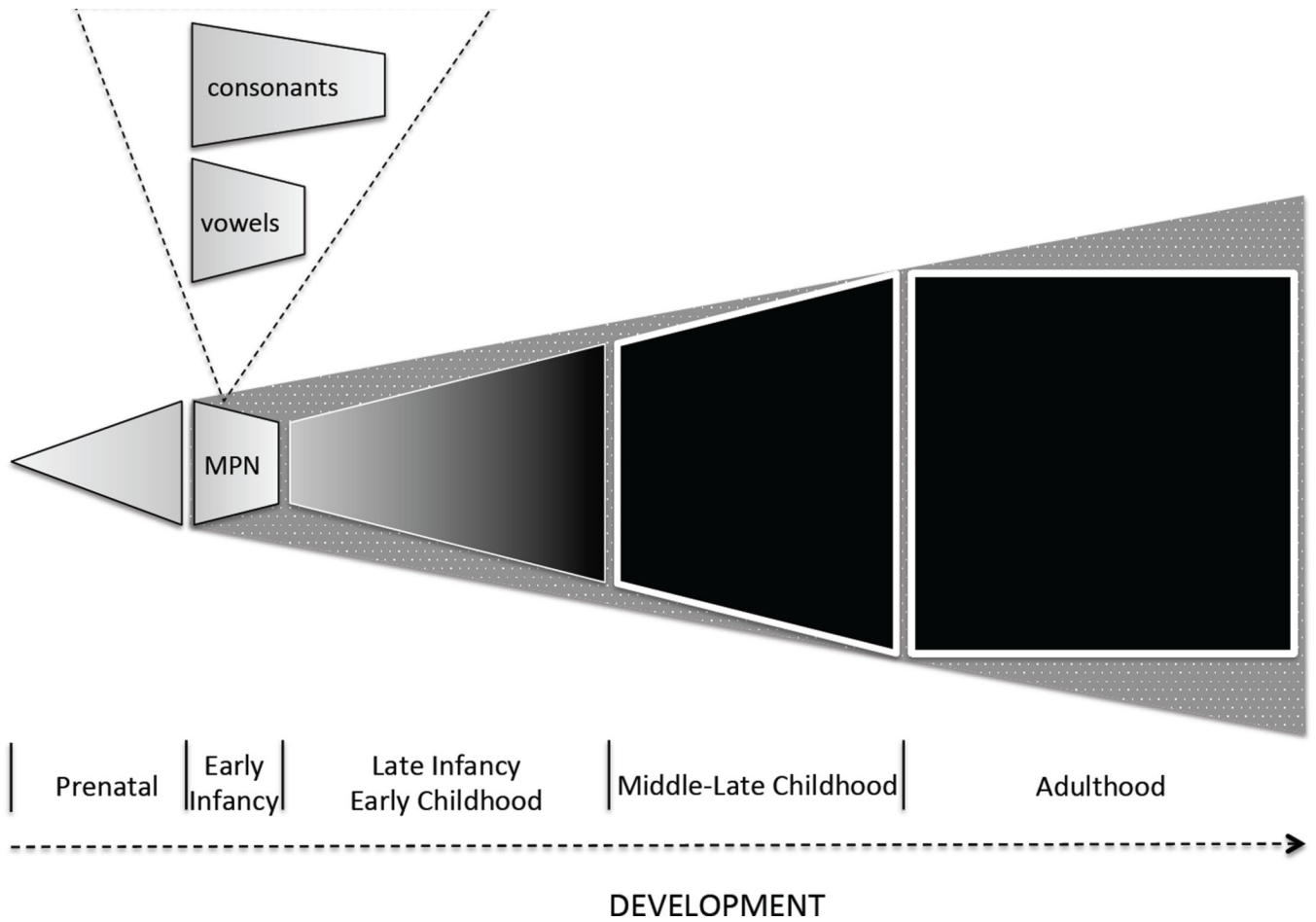
**Figure 2.** Percent correct matches that adult English- and Spanish-speaking adults made when asked whether a previously heard /ba/ or /va/ corresponded to a visible and silent /ba/ or /va/. Error bars represent the standard error of the mean.





**Figure 3.**

Proportion of total looking time (PTLT) difference scores, calculated as the difference between PTLT directed at the eyes versus the mouth, as a function of age in response to native (left) and non-native (right) audiovisual speech. Asterisks indicate a statistically significant difference in looking at the eyes and mouth. Error bars represent standard errors of the mean.



**Figure 4.** Schematic depiction of the development of multisensory perception, with a focus on multisensory perceptual narrowing (MPN) and its timing in early development. The inset corresponding to MPN is meant to illustrate the different timing of unisensory narrowing (see text for more details).