

# Effects of enriched auditory experience on infants' speech perception during the first year of life

T. Christina Zhao<sup>1</sup> · Patricia K. Kuhl<sup>1</sup>

© The Author(s) 2017. This article is an open access publication

**Abstract** Infants rapidly learn language in their home environments. Between 6 and 12 months of age, infants' ability to process the building blocks of speech (i.e., phonetic information) develops quickly, and this ability predicts later language development. Typically, developing infants in a monolingual language environment rapidly tune in to the phonetic information of their native language, while their sensitivity to nonnative phonetic information starts to decrease. Yet, enriched experience to a new language during this time significantly improves infants' sensitivity to the sound contrasts used in that language when compared to a control group without exposure to the new language. More recently, a new study examined another type of enriched auditory experience—musical experience—to determine its effect not only on music processing but also on phonetic processing. Results showed that a 1-month laboratory music intervention focusing on rhythm learning enhanced 9-month-old infants' neural processing not only for music but also for speech. Together, these results suggest that these enriched auditory experiences in infancy may improve infants' general auditory pattern-detection skills and their sensitivity to phonetic information.

**Keywords** Infants · Speech perception · Language experience · Music intervention · Learning

In many species, the young are particularly sensitive to environmental inputs at certain periods during development. The barn owl's ability to localize prey is calibrated by auditory-visual input during an early sensitive period in development; wearing prisms (or

---

✉ T. Christina Zhao  
zhaotc@uw.edu

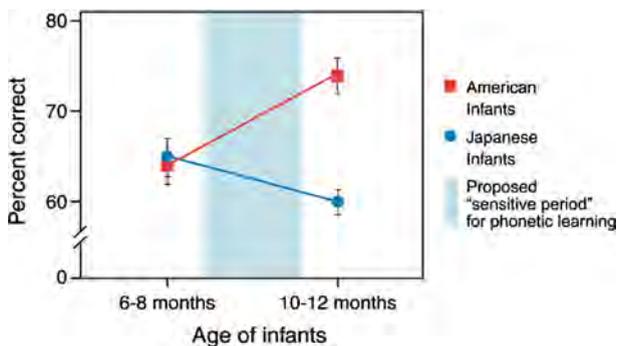
Patricia K. Kuhl  
pkkuhl@uw.edu

<sup>1</sup> Institute for Learning & Brain Sciences, University of Washington, Portage Bay Building, Box 357988, Seattle, WA 98195, USA

ear plugs) alters the mapping during this period (Knudsen 2002). Binocular fusion is dependent on binocular visual input during a critical period early in development; rearing cats with one occluded eye irreversibly alters binocular representation in the visual centers of the cortex (Hubel and Wiesel 1977; Shatz and Stryker 1978). In songbirds, learning their species-typical song depends on experience during a critical temporal window; presentation of conspecific song during that time is essential for normal development (Konishi 1985; Marler 1970). A recent theoretical paper (Werker and Hensch 2015) discusses the nature of the “critical” periods, especially the biological factors that “open” and “close” them. Here, we review work from our laboratory that focuses on one specific time period for human infants’ learning; namely, the “sensitive period” for phonetic learning and the experiential factors that may influence this learning process. We first discuss the developmental trajectory of infants’ abilities to discriminate native and nonnative phonetic contrasts between 6 and 12 months of age, and then several experiential factors we have observed in laboratory studies that influence infants’ ability to discriminate speech sounds during this sensitive period. Lastly, we discuss future directions for research that will help elucidate the mechanisms through which these experiential factors exert their influences.

## Early phonetic learning

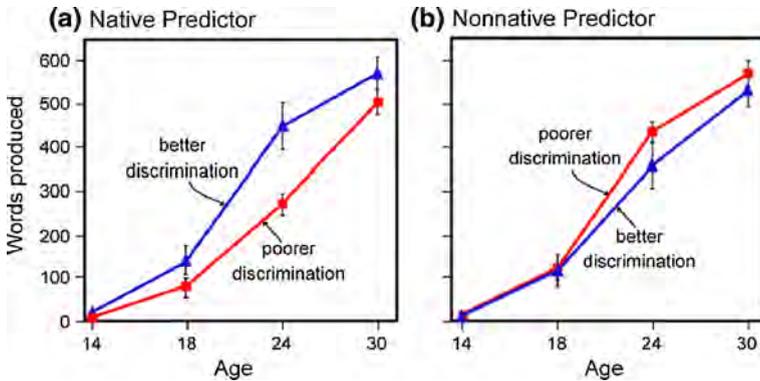
Infants’ language learning starts early in development. Infants’ speech perception skills show a dual change toward the end of the first year of life (Figure 1). Not only does nonnative speech perception decline (Best and McRoberts 2003; Werker and Tees 1984), but, also, native-language speech perception skills show improvement, reflecting a facilitative effect of experience with native language (Kuhl et al. 2006; Tsao, Liu, and Kuhl 2006). The mechanism underlying change during this sensitive period in development, and the relationship between the change in native and nonnative speech perception, is of theoretical interest. Data show that at the cusp of this developmental change, infants’ native and nonnative phonetic perception skills predict later language ability, but in opposite directions (Figure 2) (Kuhl et al. 2008). Better native phonetic discrimination at 7.5 months predicts faster native-language advancement; whereas better nonnative phonetic discrimination predicts slower native-language advancement. We have argued that



**Figure 1** Effects of age on discrimination of the American English /ra/-/la/ phonetic contrast by American and Japanese infants at 6–8 and 10–12 months of age

Note: Mean percent correct scores are shown with standard errors indicated.

Source: Kuhl (2004)



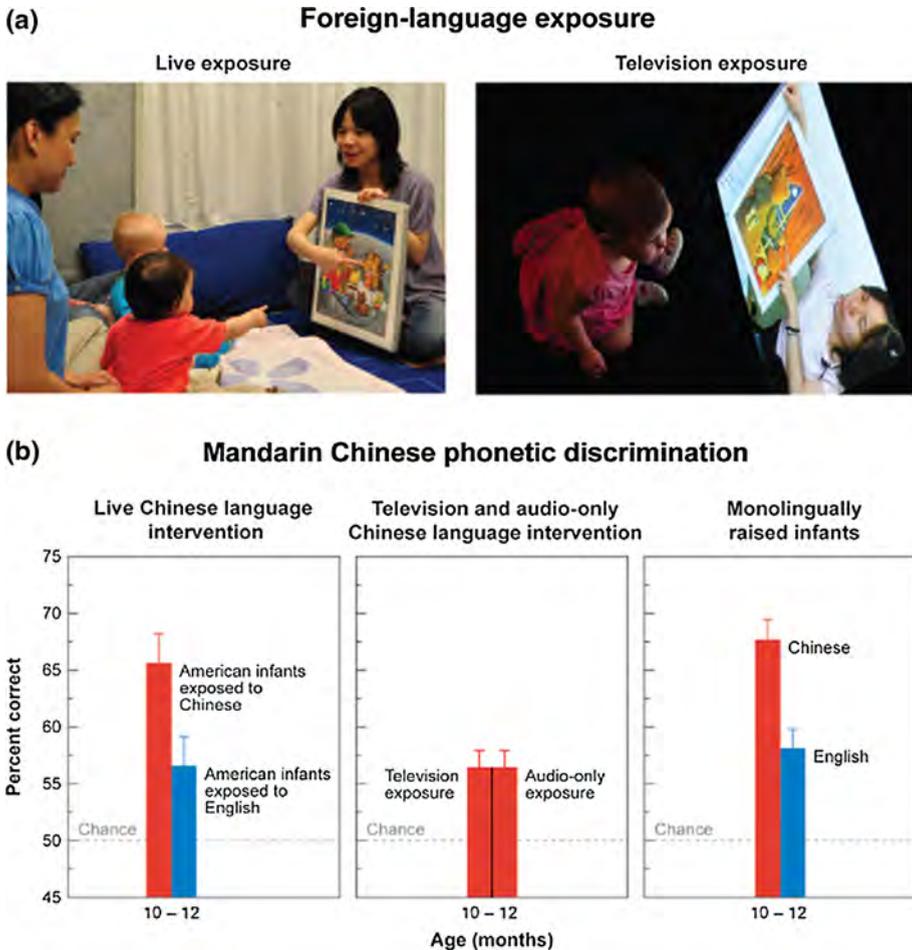
**Figure 2** A median split of infants whose MMNs indicate better versus poorer discrimination of (a) native and (b) nonnative phonetic contrasts is shown, along with their corresponding longitudinal growth curve functions for the number of words produced between 14 and 30 months of age

Source: Kuhl et al. (2008)

this pattern of results is indicative of “neural commitment” to the native language and reflects infant attention to the acoustic cues made available by language input, especially language input in the form of “motherese” (more recently termed “parentese” because both mothers and fathers use it) in one-on-one social contexts (Ramirez-Esparza, Garcia-Sierra, and Kuhl 2014, 2016). That is, better skills in native phonetic discrimination support neural network development, which allows more efficient processing of native speech sounds. Alternatively, better skills in nonnative phonetic perception reveal uncommitted neural circuitry that is less efficient for processing native speech sounds (see Kuhl 2004, Kuhl et al. 2008, for elaboration).

## Social influences on phonetic learning during the sensitive period

During infants’ sensitive period for phonetic learning, between 6 and 12 months of age, studies show that infant perception of speech is highly malleable. During this time, laboratory experiments indicate that distributional and statistical learning can occur with just two minutes’ exposure to novel speech material (e.g., Maye, Werker, and Gerken 2002; Saffran, Aslin, and Newport 1996). However, studies have also shown strong social influences in their investigations of whether infants are capable of phonetic learning at 9 months of age from natural first-time exposure to a foreign language (Conboy and Kuhl 2011; Kuhl, Tsao, and Liu 2003). Kuhl and colleagues (Kuhl, Tsao, and Liu 2003), in a foreign-language intervention experiment, exposed 9-month-old infants to Mandarin Chinese, a language with prosodic and phonetic structure very different from English. Infants heard 4 native speakers of Mandarin (2 male, 2 female) during twelve 25-minute sessions of book reading and play during a 4–6 week period. A control group of infants also came into the laboratory for the same number and variety of reading and play sessions, but heard only English. On average, infants heard about 33,000 Mandarin syllables during the course of the 12 language-exposure sessions. Researchers tested two additional groups of infants; they exposed one group to Mandarin language material on a video screen, and presented the second group the exact same Mandarin material in the same room and on the same timetable but in an audio-only condition (Figure 3).



**Figure 3** Foreign-language learning experiments show the need for social interaction in language acquisition

*Notes:* Nine-month-old infants experienced 12 sessions of Mandarin Chinese through **a** natural interaction with a Chinese speaker (*left*) or the identical linguistic information delivered via television (*right*) or audiotape (not shown). **b** Natural interaction resulted in significant learning of Mandarin phonemes when compared with a control group who participated in interaction using English (*left*). No learning occurred from television or audiotaped presentations (*middle*). Data for age-matched Chinese and American infants learning their native languages are shown for comparison (*right*).

*Source:* Adapted from Kuhl, Tsao, and Liu (2003)

After exposure, researchers tested all 4 groups on Mandarin phonetic discrimination. The results from behavioral tests (conditioned head-turn, see Kuhl et al. 2006) on infants after exposure demonstrated that only the group exposed to Mandarin in a social context by live humans learned the Mandarin contrast. The data demonstrated two things: (a) phonetic learning from first-time exposure can occur at 9 months of age, and (b) phonetic learning from natural language exposure during the sensitive period requires social interaction. Similar second-language exposure experiments using Spanish explored both phonetic and word learning, as well as the degree to which social factors, such as visual attention, during

the exposure sessions predict individuals' learning. Using brain measures (event-related potential, ERP, measures; see Kuhl et al. 2008), the results with Spanish replicated previous findings using Mandarin; additionally, they show that English phonetic discrimination does not decline—in fact, it increases, as expected, as Spanish contrast learning increases (Conboy and Kuhl 2011). Moreover, analyses of the video records revealed a significant positive relationship between infants' social skills—which allowed them to shift gaze between the foreign-language tutor and the toys as the tutor held new toys and named them in the foreign language—and increased neural responsiveness to the Spanish contrast (Conboy, Brooks, Meltzoff, and Kuhl 2015). These correlations between social responses and brain measures of learning buttress the argument that infants' social skills are coupled to language learning.

The data on infant speech-perception reviewed above suggest that infants are very sensitive to social language input during the period between 6 and 12 months. Infants' sensitivity is so high that even a foreign language introduced for the first time at 9 months causes robust phonetic learning when it is delivered in a social context. This leads to the hypothesis that the mechanisms underlying infant speech-perception are somehow “tuned” to language input, delivered socially, during this time. The corollary hypothesis is that only language input can influence these mechanisms at this time.

A recent experiment suggests that the corollary hypothesis must be altered. In the next section, we review the results of an experiment that exposes infants to music in a way that is similar to previous experiments using foreign-language interventions during the sensitive period (Conboy and Kuhl 2011; Kuhl, Tsao, and Liu 2003). In the music intervention, researchers exposed infants to a particular rhythmical structure in music, the triple meter (the waltz), for 12 sessions in a social context, using a randomized control design. The control group experienced similar activities in a social setting, but no music. After 12 sessions, the research team tested both intervention and control infants with violations of rhythmic structure in both music and speech. The results show effects on both music and speech, and reveal activation in the infants' auditory-sensory and prefrontal cortices. In the remaining sections, we detail these findings and discuss their implications.

## Effects of music intervention on infants' phonetic learning

During the last decade, music training that starts early in development has received increasing attention in the science community as an important early experience, given the growing amount of evidence suggesting the robust and extensive training-related benefits in auditory, language, and cognitive abilities (Kraus and Chandrasekaran 2010; Shahin 2011; Zatorre 2013). Previous studies—using various methodologies, including behavioral, electrophysiological, and neural imaging methods—have demonstrated repeatedly that musically trained adults and children exhibit enhanced processing of musical information (e.g., musical pitch and meter) in comparison to nontrained groups (Fujioka, Ross, Kakigi, Pantev, and Trainor 2006; Geiser, Sandmann, Jancke, and Meyer 2010; Habibi, Cahn, Damasio, and Damasio 2016; Koelsch, Schroger, and Tervaniemi 1999; Pantev et al. 1998; Vuust et al. 2005; Zhao and Kuhl 2015a, b).

More importantly, prior studies have also demonstrated generalization effects in the trained individuals from their early musical experience to other domains, one of the most studied being speech processing. The ability to accurately and efficiently process complex speech sounds is critical in language development as speech processing in infants can

robustly predict language abilities in early childhood (see “[Early phonetic learning](#)” section); and, at the same time, studies have shown that developmental language disorders (e.g., dyslexia, specific language impairment) have origins in auditory processing deficits (Goswami 2011; Tallal and Gaab 2006). So far, researchers have found that musically trained adults and children can better encode the acoustic details in speech at the level of the brainstem, especially when speech is embedded in noise (Bidelman, Weiss, Moreno, and Alain 2014; Parbery-Clark, Skoe, Lam, and Kraus Parbery-Clark et al. 2009; Parbery-Clark, Tierney, Strait, and Kraus 2012; Strait, Parbery-Clark, O’Connell, and Kraus 2013). At the cortical level, researchers observed musically trained individuals to better process pitch information in both native and foreign speech compared to nonmusicians; one study focusing on the temporal information in speech demonstrated that adult musicians could track syllable structures in words better as well (Magne, Schon, and Besson 2006; Marie, Magne, and Besson 2011; Marques, Moreno, Castro, and Besson 2007; Wong, Skoe, Russo, Dees, and Kraus 2007). These cross-domain effects from early music training to speech perception raise theoretically interesting and important questions about different levels of processing (e.g., lower-level acoustic processing vs. higher-level cognitive skills) affected by early experience and how they can support these observed generalization effects (Kraus and Chandrasekaran 2010).

Following this growing literature, we examined the rich experience of music training in an even earlier developmental stage (9 months of age) for both theoretical and methodological reasons (Zhao and Kuhl 2016). Theoretically, this approach allowed us to compare the effects of music experience during the sensitive period of phonetic learning to other previously studied experiences, such as experience of a foreign language (Kuhl, Tsao, and Liu 2003). Methodologically, (1) we were able to randomly assign infants at this age to complete either a structured laboratory-controlled music intervention (Intervention) or control activities (Control). This approach allowed controlling for effects related to predispositions (e.g., genetics), prior music experience, and the variability in individuals’ music training (e.g., onset, nature, and duration of the music experience); (2) we focused on temporal information processing, which has less experimental data regarding effects derived from early music training. In this study, the Intervention targeted infants’ learning of a specific meter (triple meter—e.g., waltz) and we tested the effects of the Intervention on both music (metrical structure) and speech (syllable structure); (3) we used neural responses, measured by magnetoencephalography (MEG), as outcome measures to compare Intervention and Control infants in the spatial and temporal aspects of their cortical responses.

We predicted enhancement in both music and speech domains, following the rationale that the Intervention—targeting infants’ learning of a specific meter—exerts influence at a higher level of processing. We argued that the Intervention infants would become better at extracting the temporal pattern of complex sounds over time, leading to their ability to make more robust predictions about the timing of future stimuli based on the extracted temporal structure—an ability that would affect both music and speech processing.

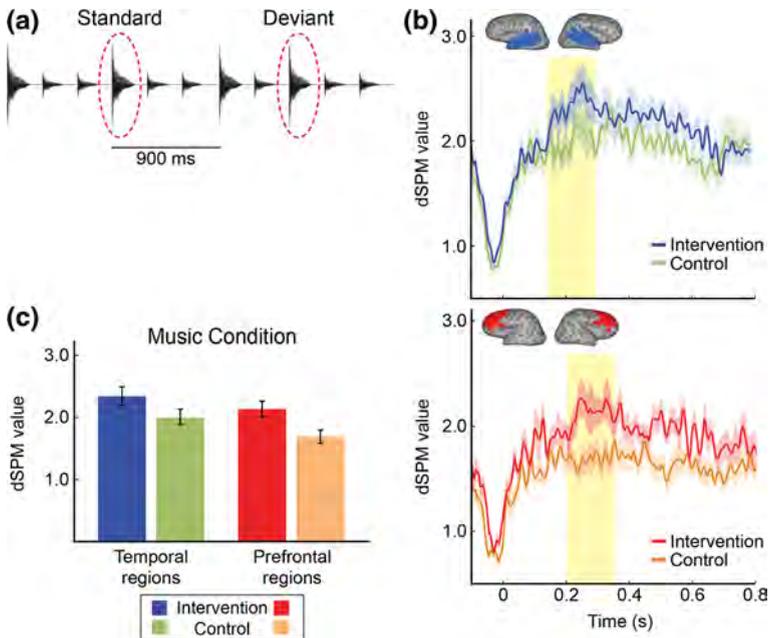
The design of the Intervention/Control sessions paralleled our prior studies in the laboratory on infant speech learning at 8–10 months of age (see “[Social influences](#)”). Specifically, we recruited 9-month-old infants raised in monolingual English-speaking environments with comparable prior and concurrent music listening experiences at home, whose parents were not performing musicians. We randomly assigned infants to the Intervention or Control group for 12 sessions (15 minutes each), over a 4-week period, of corresponding activity in the laboratory.

In the Intervention/Control sessions, we incorporated several key components to maximize infants' learning specific to the Intervention while reflecting naturalistic infant music classes: (1) Intervention infants experienced various infant tunes and songs only in triple meter (e.g., waltz). We selected triple meter as the target temporal structure because studies have shown that it is a more difficult temporal structure in Western music for infants to process at this age than duple meter (e.g., marching music) (Bergeson and Trehub 2006), yet infants can rapidly learn temporal patterns in the music of their culture (Gerry, Faux, and Trainor 2010; Hannon and Trehub 2005a, b); (2) Intervention infants, with the aid of caregivers, tapped out the musical beats with maracas or their feet, and their caregivers often bounced them in synchronization to the musical beats—activities that are common in infant music classes and effective in infants' learning of temporal structure (Phillips-Silver and Trainor 2005); (3) the Control sessions offered comparable visits to a laboratory, familiarity with the laboratory environment, levels of social interaction with other infants and caregivers, and levels of motor activity and engagement, but without music. For example, infants, aided by their parents, played with toy cars, blocks, and other objects that required coordinated movements, such as moving and stacking; (4) in both the Intervention and Control sessions, researchers engaged infants in a social setting with 1–2 other infants and their caregivers, a setting demonstrated in previous work to be effective when infants are exposed to a foreign language (Kuhl, Tsao, and Liu 2003). Experimenters facilitated each session by engaging the infants and their caregivers in the activities to a comparable degree.

To examine whether the intervention enhanced infants' general ability to extract temporal structure and generate more robust predictions about future stimuli in complex auditory sounds, we examined Intervention infants' neural responses to temporal structure violations in both music and speech in temporal (auditory) and prefrontal cortical regions, in comparison to their Control group counterparts. We quantified the neural responses by a specific neural response, namely the mismatch response (MMR), traditionally measured by an oddball paradigm. In this paradigm, a standard stimulus is presented on approximately 85% of the trials to establish a temporal structure; on the remaining 15% of the trials, a deviant stimulus that violates this temporal structure is randomly presented on the remaining 15% of the trials (Figures 4a, 5a). The magnitude of the MMR, which peaks around 150–350 ms after the violation onset, thus reflects neural sensitivity to the violation of temporal structure—and thus the tracking and learning of that temporal structure (Bekinschtein et al. 2009; Schwartze and Kotz 2013; Winkler, Denham, and Nelken 2009). We recorded neural responses to all stimuli using magnetoencephalography (MEG), which has excellent temporal resolution and good spatial resolution, allowing the examination of MMR in the specific time windows of interests (i.e., around 150–350ms post violation) and in target cortical regions (i.e., temporal and prefrontal regions).

Our results supported our hypotheses and answered our specific questions, demonstrating that: (1) the Intervention group exhibited a larger MMR response to violations in temporal structure for music (i.e., triple meter) when compared to the Control group; (2) the effects were observed in both temporal (auditory) and prefrontal regions of the cortex (Figure 4b, c); (3) the enhancement in temporal structure processing generalized to the speech domain, reflected by a larger MMR in temporal and prefrontal cortical regions in response to violations of a foreign temporal structure in the Intervention group (Figure 5b, c).

We therefore demonstrated that a short-term laboratory-controlled music intervention at 9 months of age that reflects naturalistic infant music classes affects not only infants' functional processing of temporal structure in music but also—more importantly—infants'



**Figure 4** Music condition (MEG)

*Notes:* **a** Schematics of stimuli; standard and deviant sounds are acoustically identical, and deviants violate the standard temporal structure. **b** *Top:* The group average of the difference waves for the temporal regions of the cortex for the Intervention group and the Control group. Shaded region indicates the selected time window for the MMR. Time 0 marks the onset of the strong beat. *Bottom:* The group average of the difference waves for the prefrontal regions of the cortex for the Intervention group and the Control group. **c** Mean MMR values within the target time window by region (temporal region vs. prefrontal region) and group (Intervention vs. Control).

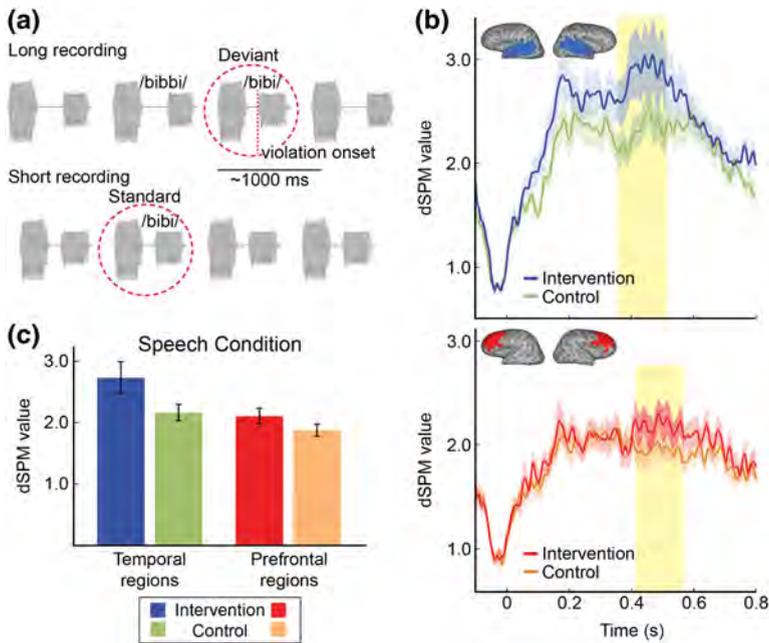
*Source:* Adapted from Zhao and Kuhl (2016)

processing of syllable structure in speech. We based our prediction of the generalization effects from the Intervention to speech on the rationale that infants would learn to better attend to and extract auditory patterns in the temporal domain, allowing them to generate—from learned patterns—more robust predictions about the timing of future events. Our results thus strongly supported the idea that such enriched music intervention experience may support the development of a broader set of perceptual skills.

The design of the Intervention, as well as the use of foreign syllable structure, in the MEG testing in this study allows us to compare the current results to our previous experiments examining the effects of foreign-language intervention during this sensitive period of phonetic learning. In the next section, we discuss in more detail the implications of the result showing enhanced sensitivity to foreign syllable structure contrasts.

## Summary and discussion

In this article, we have introduced the concept of what we term a “sensitive period” for infants’ phonetic learning between the age of 6 and 12 months (Kuhl 2004). Decades of research have demonstrated that infants’ ability to discriminate native speech contrasts



**Figure 5** Speech condition (MEG)

*Notes:* **a** Schematics of stimuli; *deviants* /bibi/ violate the syllable structure of /bibbi/. In a separate recording (lower panel), /bibbi/ served as *standards* in a constant stream. **b** *Top*: The group average of the difference waves for the temporal regions of the cortex for the Intervention group and the Control group. Shaded region indicates the selected time window for the MMR, shifted accordingly with the onset of violation (210ms after the onset of the nonword /bibbi/, marked by Time 0). *Bottom*: The group average of the difference waves for the prefrontal regions of the cortex for the Intervention group and the Control group. **c** Mean MMR values within the target time window by region (temporal region vs. prefrontal region) and group (Intervention vs. Control).

*Source:* Adapted from Zhao and Kuhl (2016)

improves, in contrast to their ability to discriminate nonnative speech contrasts that decreases during this period (Kuhl et al. 2006; Werker and Tees 1984). Further, we discussed that infants' phonetic learning during this sensitive period is highly malleable, depending on the auditory input infants receive at that time. The skill to discriminate nonnative speech contrasts provides a window for us to study how inputs during the sensitive period can affect infants' phonetic learning. In a series of studies, we demonstrated that experience with a foreign language could enhance infants' ability to discriminate the nonnative speech contrasts in that language. More importantly, language experience during this time needs to be social in nature—the same input delivered through a TV screen did not result in learning (Conboy and Kuhl 2011; Kuhl, Tsao, and Liu 2003). Yet, in our most recent study, we show that a music intervention targeting rhythm learning during this sensitive period also enhanced infants' ability to discriminate a nonnative speech contrast that is based on syllable structure differences.

How does the enriched auditory experience of foreign language and music exert its influence on infants' phonetic learning during the sensitive period for phonetic learning? Previous research has demonstrated the influences of cognitive skills on speech perception in this period; 11-month old monolingual infants show a strong negative correlation

between specific cognitive controls skills (inhibitory control) and nonnative speech discrimination (Conboy, Sommerville, and Kuhl 2008; Diamond, Werker, and Lalonde 1994; Lalonde and Werker 1995). The authors' interpretation is that infants with good inhibitory control skills are better able to ignore speech sounds that are irrelevant to their native language, and, therefore, that they exhibit lower nonnative speech discrimination skills, which has been shown to correlate with faster native-language growth (Figure 2; Kuhl et al. 2008). On the other hand, literature on infants and children raised in bilingual language environments demonstrate enhanced cognitive flexibility compared to their monolingual counterparts (Bialystok and Craik 2010; Kovács and Mehler 2009a, b). We, therefore, speculate that an enriched auditory experience (i.e., foreign language and music) provides complex yet patterned auditory input; when delivered in a social setting, it allows infants to develop enhanced cognitive abilities to switch between inputs and attune their attentional resources to the relevant and important auditory information.

One specific mechanism by which infants can learn to effectively allocate attentional resources is predictive coding. By extracting the temporal pattern of input, the dynamic attending theory posits that attentional resources are allocated to time windows during which the brains predict that important information will occur (e.g., musical beats, syllables) (Jones and Boltz 1989). Investigators have demonstrated that infants as young as 3 months of age are able to extract temporal patterns and predict future stimuli based on the extracted information (Basirat, Dehaene, and Dehaene-Lambertz 2014; Emberson, Richards, and Aslin 2015). Our recent data using complex auditory stimuli suggest that a music intervention focusing on temporal information learning may have increased infants' ability to extract high-level temporal patterns and generate stronger predictions about future stimuli—a skill that they can apply both in music and in speech processing. Future research is warranted to, first, establish the relationships between different general cognitive skills (e.g., inhibition, flexibly switching attention) and infants' ability to discriminate native and nonnative speech sounds. Then, it will be critical to directly test whether short-term language or music experience, in comparison to no exposure, affects these cognitive skills—which can, in turn, affect phonetic learning during the “sensitive period”. In the longer term, researchers should dissect and systematically examine the various components of these enriched auditory experiences (e.g., social elements, multi-model elements) in order to evaluate the effectiveness of each element and the interactions among them. This will not only enhance our theoretical understanding of infant phonetic learning but will also inform the design of early-education interventions, especially for infants at risk for communication disorders.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition*, 132(2), 137–150. doi:10.1016/j.cognition.2014.03.013.
- Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F. O., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), 1672–1677. doi:10.1073/pnas.0809667106.

- Bergeson, T. R., & Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception*, 23(4), 345–360. doi:[10.1525/mp.2006.23.4.345](https://doi.org/10.1525/mp.2006.23.4.345).
- Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, 46, 183–216.
- Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science*, 19(1), 19–23. doi:[10.1177/0963721409358571](https://doi.org/10.1177/0963721409358571).
- Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience*, 40(4), 2662–2673. doi:[10.1111/ejn.12627](https://doi.org/10.1111/ejn.12627).
- Conboy, B. T., Brooks, R., Meltzoff, A. N., & Kuhl, P. K. (2015). Social interaction in infants' learning of second-language phonetics: An exploration of brain-behavior relations. *Developmental Neuropsychology*, 40(4), 216–229. doi:[10.1080/87565641.2015.1014487](https://doi.org/10.1080/87565641.2015.1014487).
- Conboy, B. T., & Kuhl, P. K. (2011). Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Developmental Science*, 14(2), 242–248. doi:[10.1111/j.1467-7687.2010.00973.x](https://doi.org/10.1111/j.1467-7687.2010.00973.x).
- Conboy, B. T., Sommerville, J. A., & Kuhl, P. K. (2008). Cognitive control factors in speech perception at 11 months. *Developmental Psychology*, 44(5), 1505–1512. doi:[10.1037/a0012975](https://doi.org/10.1037/a0012975).
- Diamond, A., Werker, J. F., & Lalonde, C. (1994). Toward understanding commonalities in the development of object search, detour navigation, categorization, and speech perception. In G. Dawson & K. W. Fischer (Eds.), *Human behavior and the developing brain*. New York, NY: Guilford Press.
- Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences of the United States of America*, 112(31), 9585–9590. doi:[10.1073/pnas.1510343112](https://doi.org/10.1073/pnas.1510343112).
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, 129, 2593–2608. doi:[10.1093/brain/aw1247](https://doi.org/10.1093/brain/aw1247).
- Geiser, E., Sandmann, P., Jancke, L., & Meyer, M. (2010). Refinement of metre-perception training increases hierarchical metre processing. *European Journal of Neuroscience*, 32(11), 1979–1985. doi:[10.1111/j.1460-9568.2010.07462.x](https://doi.org/10.1111/j.1460-9568.2010.07462.x).
- Gerry, D. W., Faux, A. L., & Trainor, L. J. (2010). Effects of Kindermusik training on infants' rhythmic enculturation. *Developmental Science*, 13(3), 545–551. doi:[10.1111/j.1467-7687.2009.00912.x](https://doi.org/10.1111/j.1467-7687.2009.00912.x).
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3–10. doi:[10.1016/j.tics.2010.10.001](https://doi.org/10.1016/j.tics.2010.10.001).
- Habibi, A., Cahn, B. R., Damasio, A., & Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Developmental Cognitive Neuroscience*, 21, 1–14. doi:[10.1016/j.dcn.2016.04.003](https://doi.org/10.1016/j.dcn.2016.04.003).
- Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, 16(1), 48–55. doi:[10.1111/j.0956-7976.2005.00779.x](https://doi.org/10.1111/j.0956-7976.2005.00779.x).
- Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, 102(35), 12639–12643. doi:[10.1073/pnas.0504254102](https://doi.org/10.1073/pnas.0504254102).
- Hubel, D. H., & Wiesel, T. N. (1977). Ferrier Lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London B: Biological Sciences*, 198(1130), 1–59. doi:[10.1098/rspb.1977.0085](https://doi.org/10.1098/rspb.1977.0085).
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491. doi:[10.1037//0033-295x.96.3.459](https://doi.org/10.1037//0033-295x.96.3.459).
- Knudsen, E. I. (2002). Instructed learning in the auditory localization pathway of the barn owl. *Nature*, 417(6886), 322–328. doi:[10.1038/417322a](https://doi.org/10.1038/417322a).
- Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6), 1309–1313.
- Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience*, 8, 125–170. doi:[10.1146/annurev.ne.08.030185.001013](https://doi.org/10.1146/annurev.ne.08.030185.001013).
- Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16), 6556–6560. doi:[10.1073/pnas.0811323106](https://doi.org/10.1073/pnas.0811323106).
- Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, 325(5940), 611–612. doi:[10.1126/science.1173947](https://doi.org/10.1126/science.1173947).
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599–605. doi:[10.1038/nrn2882](https://doi.org/10.1038/nrn2882).

- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. doi:[10.1038/nrn1533](https://doi.org/10.1038/nrn1533).
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B-Biological Sciences*, 363(1493), 979–1000. doi:[10.1098/rstb.2007.2154](https://doi.org/10.1098/rstb.2007.2154).
- Kuhl, P. K., Stevens, E., Hayachi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13–F21. doi:[10.1111/j.1467-7687.2006.00468.x](https://doi.org/10.1111/j.1467-7687.2006.00468.x).
- Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15), 9096–9101. doi:[10.1073/pnas.1532872100](https://doi.org/10.1073/pnas.1532872100).
- Lalonde, C. E., & Werker, J. F. (1995). Cognitive influences on cross-language speech perception in infancy. *Infant Behavior and Development*, 18(4), 459–475.
- Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211.
- Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, 23(2), 294–305. doi:[10.1162/jocn.2010.21413](https://doi.org/10.1162/jocn.2010.21413).
- Marler, P. (1970). Birdsong and speech development: Could there be parallels? *Am Sci*, 58(6), 669–673.
- Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, 19(9), 1453–1463.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111. doi:[10.1016/s0010-0277\(01\)00157-3](https://doi.org/10.1016/s0010-0277(01)00157-3).
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811–814. doi:[10.1038/33918](https://doi.org/10.1038/33918).
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, 30(6), 653–661.
- Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinctions of speech syllables. *Neuroscience*, 219, 111–119. doi:[10.1016/j.neuroscience.2012.05.042](https://doi.org/10.1016/j.neuroscience.2012.05.042).
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308(5727), 1430. doi:[10.1126/science.1110922](https://doi.org/10.1126/science.1110922).
- Ramirez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, 17(6), 880–891. doi:[10.1111/desc.12172](https://doi.org/10.1111/desc.12172).
- Ramirez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2016). The impact of early social interactions on later language development in Spanish-English bilingual infants. *Child Development*. doi:[10.1111/cdev.12648](https://doi.org/10.1111/cdev.12648).
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. doi:[10.1126/science.274.5294.1926](https://doi.org/10.1126/science.274.5294.1926).
- Schwartz, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience and Biobehavioral Reviews*, 37(10), 2587–2596. doi:[10.1016/j.neubiorev.2013.08.005](https://doi.org/10.1016/j.neubiorev.2013.08.005).
- Shahin, A. (2011). Neurophysiological influence of musical training on speech perception. *Frontiers in Psychology*. doi:[10.3389/fpsyg.2011.00126](https://doi.org/10.3389/fpsyg.2011.00126).
- Shatz, C. J., & Stryker, M. P. (1978). Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *The Journal of Physiology*, 281, 267–283.
- Strait, D. L., Parbery-Clark, A., O'Connell, S., & Kraus, N. (2013). Biological impact of preschool music classes on processing speech in noise. *Developmental Cognitive Neuroscience*, 6, 51–60. doi:[10.1016/j.dcn.2013.06.003](https://doi.org/10.1016/j.dcn.2013.06.003).
- Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, 29(7), 382–390. doi:[10.1016/j.tins.2006.06.003](https://doi.org/10.1016/j.tins.2006.06.003).
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America*, 120(4), 2285–2294.
- Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., et al. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *Neuroimage*, 24(2), 560–564. doi:[10.1016/j.neuroimage.2004.08.039](https://doi.org/10.1016/j.neuroimage.2004.08.039).
- Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, 66(66), 173–196. doi:[10.1146/annurev-psych-010814-015104](https://doi.org/10.1146/annurev-psych-010814-015104).

- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, 7, 49–63. doi:[10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3).
- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, 13(12), 532–540. doi:[10.1016/j.tics.2009.09.003](https://doi.org/10.1016/j.tics.2009.09.003).
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422. doi:[10.1038/nn1872](https://doi.org/10.1038/nn1872).
- Zatorre, R. J. (2013). Predispositions and plasticity in music and speech learning: Neural correlates and implications. *Science*, 342(6158), 585–589. doi:[10.1126/science.1238414](https://doi.org/10.1126/science.1238414).
- Zhao, T. C., & Kuhl, P. K. (2015a). Effect of musical experience on learning lexical tone categories. *Journal of the Acoustical Society of America*, 137(3), 1452–1463. doi:[10.1121/1.4913457](https://doi.org/10.1121/1.4913457).
- Zhao, T. C., & Kuhl, P. K. (2015b). Higher-level linguistic categories dominate lower-level acoustics in lexical tone processing. *Journal of the Acoustical Society of America*, 138(2), EL133–EL137. doi:[10.1121/1.4927632](https://doi.org/10.1121/1.4927632).
- Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences of the United States of America*, 113(19), 5212–5217. doi:[10.1073/pnas.1603984113](https://doi.org/10.1073/pnas.1603984113).

**T. Christina Zhao (United States)** is a postdoctoral fellow under the supervision of Patricia Kuhl at the Institute for Learning & Brain Sciences, University of Washington. Zhao's work centers on examining the effects of early auditory experiences (e.g., language experience, music experience) on complex sound processing (e.g., music sounds, speech sounds) in the brain. So far, she has focused on the early experience of music on the brain's ability to process complex auditory information (including both music and speech), hypothesizing that music experience tunes higher level processes in the brain such as attentional control and predictive coding, which are also critical for speech perception. Her training has been interdisciplinary; she has a background in psychology, biology, piano performance, and speech and hearing sciences. Outside of work, she enjoys a wide range of activities including attending concerts and traveling throughout the world.

**Patricia K. Kuhl (United States)** is internationally recognized for her research on early language and brain development, and studies showing how young children learn. Her work has played a major role in demonstrating how early exposure to language alters the brain. It has implications for critical periods in development, for bilingual education and reading readiness, for developmental disabilities involving language, and for research on computer understanding of speech. She is a member of the National Academy of Sciences, the Rodin Academy, and the Norwegian Academy of Science and Letters. She was awarded the IPSEN Foundation's Jean-Louis Signoret Neuropsychology Prize in 2011, the William James Lifetime Achievement Award by the Association for Psychological Science in 2013, became Honoris Causa Doctor of Stockholm University in 2014, and received the George A. Miller Prize in Cognitive Neuroscience from the Cognitive Neuroscience Society in 2015. She has appeared on the Discovery television series *The Baby Human*, the NOVA series *The Mind*, and PBS's *The Power of Ideas* and *The Secret Life of the Brain*. In 2000, she co-authored *The Scientist in the Crib: Minds, Brains, and How Children Learn* (HarperCollins).