Infants’ Preferential Attention to Sung and Spoken Stimuli

Eugenia Costa-Giomi¹ and Beatriz Ilari²

Abstract
Caregivers and early childhood teachers all over the world use singing and speech to elicit and maintain infants’ attention. Research comparing infants’ preferential attention to music and speech is inconclusive regarding their responses to these two types of auditory stimuli, with one study showing a music bias and another one indicating no differential attention. The purpose of this investigation was to study 11-month-old infants’ preferential attention to spoken and sung renditions of an unfamiliar folk song in a foreign language (n = 24). The results of an infant-controlled preference procedure showed no significant differences in attention to the two types of stimuli. The findings challenge infants’ well-documented bias for speech over nonspeech sounds and provide evidence that music, even when performed by an untrained singer, can be as effective as speech in eliciting infants’ attention.

Keywords
language, music, infant, singing, speech, attention

Singing to infants is a universal child-rearing practice (Trehub, 2013). Parents and early childhood teachers all over the world comfort and entertain babies by singing to them (Custodero, 2006; Ilari, Moura, & Bourscheidt, 2011; Young, 2008) and, by doing so, support language development, strengthen social bonding, and promote musical development (e.g., Costa-Giomi, 2013; Trehub, 2001). Considering that

¹University of Texas at Austin, Austin, TX, USA
²University of Southern California, Los Angeles, CA, USA

Corresponding Author:
Eugenia Costa-Giomi, University of Texas at Austin, 1 University Station E3100, Austin, TX 78712-0435, USA.
Email: costagiomi@mail.utexas.edu
singing to infants is such a natural and effective child-rearing practice, it may come as no surprise that the distinct way in which adults talk to infants actually resembles singing. The characteristics of infant-directed (ID) speech, such as the use of a melodic contour, exaggerated stress patterns, and elongated vowels (Fernald, 1985), are found also in infant music, such as lullabies (Trehub, Unyk, & Trainor, 1993). Infants’ attraction to ID speech is reflected in longer attention to this type of stimulus as compared to adult-directed (AD) speech.

Although it is well known that infants are more attentive to ID than to AD speech (e.g., Fernald, 1985), it is unclear whether they are more attentive to music than to speech. Considering infants’ attraction to the musical nuances of ID speech (e.g., Fernald, 1985), one would expect singing to elicit preferential attention over speech. However, the results of previous studies are contradictory. Two studies on attention to speech and music during the 1st year of life have claimed a music bias (Nakata & Trehub, 2004) or no preferential attention to either type of stimulus (Corbeil, Trehub, & Peretz, 2013). The characteristics of the stimuli were very different in these investigations and may explain their conflicting results. Nakata and Trehub’s (2004) conclusion that music captures and maintains infant attention was based on infants’ longer and more focused attention to videos of their own mothers’ ID singing than to those of their mothers’ ID speaking. On the other hand, Corbeil et al.’s (2013) findings of no preferential attention to music and speech were based on infant’s responses to audio recordings of spoken syllables or hummed melodies by unfamiliar voices. The methodological differences of these investigations obviously make the integration of their findings difficult.

Overall, it is unclear how the nature of the stimulus affects infant attention, considering the differences between the music and speech stimuli used in previous research. For example, the mothers participating in Nakata and Trehub’s (2004) study chose what to sing and say to their 6-month-old infants and were free to use any facial expressions and movements they deemed appropriate during the videotaping session. As a result, the sung and spoken conditions differed not only in the linguistic and music nature of the stimuli but also in the visual and verbal content of the videos. Which of these differences triggered infants’ preferential attention for the singing videos is unknown. On the other hand, in Corbeil et al.’s (2013) study with 4- to 6- and 9- to 11-month-olds, the hummed stimulus was rich in music structure because it consisted of complete melodies, whereas the spoken stimulus was poor in linguistic structure because it depicted isolated syllables. As a result, the structure of the two types of stimuli were confounded with the music and speech variable under study.

The purpose of this investigation was to study infants’ preferential attention to music and speech. To allow for a more direct comparison of attention to speech and music, we used stimuli of the same length and verbal content; one was spoken and the other one was sung. Knowing that infants learn the referential meaning of words and are sensitive to the structure of both speech and music from early on (e.g., Jusczyk & Krumhansl, 1993; Kuhl, 2004), we searched for stimuli rich in music and linguistic syntactical structure but devoid of referential meaning. We used a folk tune in a language foreign to the infants because it depicted the complex organization of discrete
music and linguistic units characteristic of music and language (see Patel, 2008, for a review of syntax in music and language) but semantic references unfamiliar to the participants.

Method

Sample

We recruited 34 eleven-month-old infants from the registry of infants born in the city where the study was completed, following required institutional review board procedures. The data from infants who had been exposed to French \((n = 6)\) or who were premature \((n = 1)\) were excluded from the analyses. Three additional infants did not complete the listening test due to fussiness \((n = 1)\) or maternal disruption \((n = 2)\). The final sample consisted of 24 infants, 12 boys and 12 girls, ranging in age from 325 to 356 days \((M = 339, 11 \text{ months} 6 \text{ days})\). Participating children had no history of visual or hearing problems or known disability.

Materials

The musical stimuli consisted of two verses and the refrain of “Gai Lon La, Gai le Rosier,” a French folk tune in 6/8 meter (see Figure 1). The verses of the song were six measures long, and the refrain, two measures long. We chose to use a song in a language foreign to the infants to reduce possible differential effects of word understanding on infants’ attention to the music and spoken stimuli. We also opted to use a traditional song as opposed to an invented melody/text in an attempt to use stimuli that ecologically were valid.

Because children show preference for specific vocal timbres from an early age (DeCasper & Fifer, 1980; Standley and Madsen, 1990) and prefer vocal music performed with reduced vibrato (LeBlanc & Sherrill, 1987), we studied infants’ responses to two different vocal timbres: the voice of a mature female singer and the voice of a younger, untrained woman. The two women were recorded reciting and singing the song expressively “as if they had a baby in their arms.” The trained singer sang the

![Figure 1. “Gai Lon La, Gai le Rosier”](image-url)
The untrained singer sang it in B♭ with no noticeable vibrato. Singers’ renditions of the stimuli were very accurate in pitch, rhythm, and pronunciation of the text. The recordings were edited to equalize the volume, tempo, and length using the software Audacity. The edited recordings lasted 30 seconds each.

The visual stimuli used during the experiment included a flashing circle to direct infants’ attention to the monitor and the picture of a sunflower to maintain children’s gaze on the monitor while the music was playing.

**Procedures**

We completed the study using a preference procedure based on the comparison of cumulative attention to two different stimuli over a period of time (see Johnson & Zamuner, 2010, for a review of this testing procedure). During the experiment, infants sat on their caregiver’s lap facing a monitor equipped with side speakers. A video camera was positioned below the monitor to capture infants’ behaviors and transmit them to an adjacent room, where the researcher observed, analyzed, and recorded all responses. The researcher, who could not hear or see the stimuli being presented in the testing room, registered the duration of infants’ gaze to the monitor during each trial using Habit X software designed specifically for the testing of infants (Cohen, Atkinson, & Chaput, 2004).

Infants were exposed to up to 20 trials each lasting a maximum of 30 seconds (i.e., the duration of the recording). A trial consisted of the presentation of the image of a sunflower on the monitor while one of the recordings played through the speakers. Trials lasted up to 30 seconds and as long as the infant attended to the monitor. If the infant looked away from the monitor for more than 1 second during a trial, the trial stopped. The next trial started after directing the infant’s attention back to the monitor with the flashing image of a colorful circle. Once the infant looked at the monitor, the image of the sunflower appeared and the subsequent audio stimulus began to play.

Each infant listened to the same woman reciting and singing the tune and was exposed to a total of 20 alternating trials, 10 spoken and 10 sung. For half the infants, the spoken stimulus was played first. Twelve infants listened to the recordings of the trained singer, and the other 12 listened to those of the untrained singer.

**Results and Discussion**

We established cumulative attention to speech and music by calculating infants’ total looking time to the spoken and sung trials. Infants’ cumulative attention to speech and music stimuli then was compared through an analysis of variance (ANOVA) with repeated measures (speech/music) with voice (trained/untrained) as a between-subjects factor. The results of the analyses indicated no significant differences in cumulative attention to the spoken and sung renditions of the tune, $F(1, 22) = .114, p = .74$, or the voices of the two singers, $F(1, 22) = 1.1, p = .31$. Although the lack of preferential attention was more evident for the recordings of the trained than the untrained singer (trained, total time for music = 70.6 s, speech = 70.7 s; untrained, total time for
music = 73.1, speech = 62.6), the interaction between voice and stimulus was not significant, $F(1, 22) = 1.09 \ p = .31$. This nonsignificant interaction suggests that infants’ attention to the spoken and sung stimuli was not affected by the idiosyncratic renditions of the two singers.

Our results are in agreement with the findings of Corbeil et al. (2013) and support the conclusion that infants do not show preferential attention to speech or music. Both studies presented the stimuli exclusively in audio mode and used music and language devoid of referential meaning. Whereas we used a foreign language in both speech and music, Corbeil et al. used isolated syllables for the former and no words for the latter. In both studies, the stimuli were expressive and engaging but recorded in the absence of infants and thus could not be characterized as infant directed (for a discussion, see Trainor, 1996).

By contrast, Nakata and Trehub (2004) found evidence of preferential attention to music over speech when using ID stimuli. Their study was completed with videos of the mothers singing and talking to their infants spontaneously and without controlling for the content of the music and spoken discourse or the way in which mothers delivered such content. The difference in results between Nakata and Trehub’s investigation and ours may be attributed to the presence or absence of these controls, but they may also be attributed to infants’ known differential responses to ID and non-ID stimuli (Fernald, 1985; Trainor, 1996). ID singing is mellower in timbre, slower, and more expressive than AD singing (Trainor, 1996) and perhaps particularly attractive to infants. It is possible that infants show preferential attention to music over speech only when exposed to the highly engaging ID stimuli.

An alternative explanation for our finding that 11-month-olds did not attend longer to music than to speech may be the language we selected for the stimuli. Infants may show preferential attention to music when presented with sung and spoken stimuli in their native language, as was the case in Nakata and Trehub’s (2004) study, and not in a foreign language, as was the case in our investigation. However, numerous experiments have shown that infants listen longer to speech than to a variety of nonspeech sounds, including filtered speech, white noise, water sounds, sine waves, monkey calls, and other nonreferential human sounds (Butterfield & Siperstein, 1970; Colombo & Bundy, 1981; Samples & Franklin, 1978; Schultz & Vouloumanos, 2010; Spence & DeCasper, 1987; Vouloumanos, Hauser, Werker, & Martin, 2010; Vouloumanos & Werker, 2004, 2007), even when the spoken stimulus is in a foreign language (Schultz & Vouloumanos, 2010). In light of these findings, it seems unlikely that the lack of music bias in our study was due to the language we used.

Our findings showing a lack of preferential attention to the sung and spoken stimuli are remarkable if one considers infants’ well-established attraction to speech over most other nonspeech sounds. That we did not find a speech bias is important because it strengthens the notion that singing is indeed as powerful as speech in getting and keeping infants’ attention. The results of the study challenge the extent of the well-documented speech bias (e.g., Schultz & Vouloumanos, 2010) and provide evidence that music, even when performed by an untrained singer, can be as effective as speech in eliciting infants’ attention.
The fact that infants do not seem to show preferential attention to song or speech suggests that infancy is a perfect time to expose them to both. Not surprisingly, in cultures all around the globe, parents and caregivers sing and talk to their infants as part of their daily routines (Trehub, 2013; Young, 2008). The results of our study provide support for this ancient child-rearing practice that allows most infants to learn both forms of communication and to become adults who speak their native language and enjoy the music of their culture (Patel, 2008).

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by the Butler Professional Development Fund at the University of Texas–Austin.

References


**Author Biographies**

**Eugenia Costa-Giomi** is professor of music and human learning at the University of Texas–Austin. Her research focuses on musical development, music perception and cognition, and the benefits of music instruction.

**Beatriz Ilari** is assistant professor of music education at the University of Southern California. Her research interests include musical development and growth, music learning, and culture.

Submitted July 28, 2012; accepted April 5, 2013.