Singing Facilitates Word Learning and Memory

Dominique Horn, Peyton Jennings, Tatiana Mcgraw, Peng Zhou, and Weiyi Ma

1. Introduction

Language learning is essential for human survival and development, as language allows us to use systems of complex symbols to communicate information, to express ourselves, to influence the behavior of others, and to build and maintain social relationships. Like language, music is a major auditory channel of communication, present universally in all known human cultures. A significant and growing body of research has examined overlaps in the cognitive processing of language and music (see Patel, 2008). The boundary between these two domains is porous especially in the vocal communicative channel including singing and speech production. This study compares word learning and memory in song and adult-directed speech (ADS) in children, aiming to determine whether the intentional use of musical properties facilitates word learning.

Throughout history, songs have been used as a means of transmitting messages through generations (the oral tradition), suggesting that the combination of music and speech may also enhance learning and recall of the material the music is paired with (Milliron, 2017). A seminal study by Wallace (1994) showed that sung text can both facilitate and interfere with language recall, which is necessary for word learning. A number of studies have also shown that sung words enhance foreign language learning in both children (Good, Russo, & Sullivan, 2015) and adults (Ludke, Ferreira, & Overy, 2014; Rukholm, Helms-Park, Odgaard, & Smyth, 2018), suggesting that song enhances the memory representation for these words. Further, Schön et al. (2008) showed that novel, sung syllables were learnt more effectively in a statistical learning paradigm compared to novel spoken syllables, and that this effect was

* Dominique Horn, School of Human Environmental Sciences, University of Arkansas, Fayetteville, AR, 72701. Peyton Jennings, Department of Psychological Science, University of Arkansas, Fayetteville, AR, 72701. Tatiana Mcgraw, School of Human Environmental Sciences, University of Arkansas, Fayetteville, AR, 72701. Peng Zhou, Department of Foreign Languages and Literatures, Tsinghua University, Beijing 100084, China. Weiyi Ma, School of Human Environmental Sciences, University of Arkansas, Fayetteville, AR, 72701, weiyima@uark.edu. W.M. is supported by The University of Arkansas Startup Fund and the Provost's Collaborative Research Grant. P.Z is supported by Tsinghua University Initiative Scientific Research Program [2016THZWLJ14].

enhanced when syllables were consistently matched with particular tones. Furthermore, a growing body of research suggests that music training may enhance language processing ability (Fujii & Wan, 2014; Kraus & Chandrasekaran, 2010; Maess, Koelsch, Gunter, & Friederici, 2001; Patel, 2011; Tierney & Kraus, 2014, Yang, Ma, Gong, Hu, & Yao, 2014),

In addition, another line of research suggests that adults capitalize on infant sensitivity to musical features by using a special form of speech known as infant-directed speech (IDS). IDS is characterized by features such as exaggerated prosody, slower transitions, and increased repetition: all elements characteristic of music (Fernald, 1989). The use of IDS can facilitate language learning in children and adults (e.g., Estes & Hurley, 2013; Golinkoff & Alioto, 1995; Foursha-Stevenson, Schembri, Nicoladis, & Erikson, 2017; Ma, Fiveash, Margulis, Behrend, Thompson, In press; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011; Thiessen, Hill, & Saffran, 2005; Singh, Nestor, Parikh, & Yull, 2009).

2. Present study

The current study aimed 1) to investigate whether word learning in children differed depending on whether words were presented in ADS or song, and to 2) document the development of the effect of singing on child word learning.

2.1. Participants

Three- to five-year-old children participated in this study at the Jean Tyson Child Development Study Center and the Head Start Schools in the Northwestern Arkansas Region.

2.2. Apparatus and stimuli

A child-friendly eye-tracking method – Intermodal Preferential Looking Paradigm (IPLP) will be used (Golinkoff, Ma, Song, & Hirsh-Pasek, 2013; Ma, Zhou, Singh, & Gao, 2017). In the IPLP, infants’ language knowledge is measured by their differential visual fixation to two images presented side-by-side when only one of the images matches an accompanying linguistic stimulus.

2.3. Procedure

An IPLP set-up was used. Each child sat on a research assistant’s lap facing a 40in TV monitor. An experiment consists of a training phase where children learned several new, made-up words by building associations between visual and auditory stimuli, and a test phase, where their understanding of these new words is assessed. Visual stimuli were displayed in the center (during training) or to the left and right of the TV screen (during test) at eye level. Auditory stimuli were presented through internal speakers within the TV monitor. During a training phase, children were shown animated videos of two novel objects,
each paired with one training sentence (e.g., Look! This is a “X”) in which the new word “X” was either sung or spoken in adult-directed speech. During a test phase, children were shown the static version of the two novel objects side-by-side on the TV screen and be directed to look at one of them on each test trial. A hidden camera recorded children’s visual fixation to the display. An experiment lasted approximately 5 minutes.

2.4. Coding and data analysis

Using SuperCoder, participants’ eye movements were coded frame-by-frame to the thirtieth of a second. Mean increases in looking time at the target image across phases were calculated for each type of classifiers.

Table 1. An example of experimental procedure

<table>
<thead>
<tr>
<th></th>
<th>Left side</th>
<th>Center</th>
<th>Right side</th>
<th>Speech stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Salience</strong></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td>No audio</td>
</tr>
<tr>
<td><strong>Training</strong></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td>Look here! It’s a modi! See the modi. That’s the modi…</td>
</tr>
<tr>
<td>Animations of objects (4 trials, 24 s each).</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td>Look here! It’s a pabu! See the pabu. That’s the pabu…</td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td>Modi! Where’s the modi? Look at the modi! There’s the modi.</td>
</tr>
<tr>
<td>(4 trials: 7 s each)</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td>Pabu! Where’s the pabu? Look at the pabu! There’s the pabu.</td>
</tr>
</tbody>
</table>

3. Results and discussion

Three major findings emerged. First, the proportion of looking time to the target was compared against chance (0.5). Successful word learning is conventionally indicated by fixation to target that exceeds chance (0.50). Results showed that both the younger children’s (2- to 3-year-olds) and the older children’s (4- to 5-year-olds) successfully learned the words – regardless of
whether they were sung or spoken. Second, children’s word learning performance was compared between age groups. Results showed that older children’s word learning performance was significantly better than young children’s for both the sung and spoken words. Third, word learning performance was compared between the sung and spoken words within each age group. In the 4- and 5-year-olds, word learning performance did not differ between the sung and spoken words. However, in the 2- and 3-year-olds, word learning performance was better for the sung words than for the spoken words.

These findings suggested that singing facilitated young children’s word learning, and that the effect of singing on child word learning attenuates with age. How did song facilitate immediate word learning? First, the prosodic cues in song, such as such as pausing, pre-boundary lengthening, and intonation, are important for speech segmentation. Second, the sentence and word duration in song may facilitate word segmentation and word association. Third, song and IDS prosody may also facilitate novel word learning because they elicit more attention in participants than ADS. A defining feature of song and IDS is their deviation from a speaker’s mean acoustic profile, functioning to attract attention (Ma & Thompson, 2015; Ma, Fiveash, & Thompson, In press).

References


Ma, Weiyi., Fiveash, Anna., & Thompson, William. (In press). Spontaneous emergence of language-like and music-like vocalizations from an artificial protolanguage. *Semiotica*. (Published online first)


Yang, Hua., Ma, Weiyi., Gong, Diankun., Hu, Jiehui., & Yao, Dezhong. (2014). A longitudinal study on children’s music training experience and academic development. *Scientific Reports, 4*, 5854