Brain & Language 127 (2013) 452–462

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Two faces, two languages: An fMRI study of bilingual picture naming

Yunqing Li [a], Jing Yang [b,a], K. Suzanne Scherf [a], Ping Li [a,*]

[a] Department of Psychology and Center for Brain, Behavior, and Cognition, Pennsylvania State University, University Park, PA 16802, USA

[b] Guangwai Brain and Language Lab and National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou 510420, China

ARTICLE INFO

Article history:
Accepted 13 September 2013
Available online 13 October 2013

Keywords:
Bilingual language processing
Picture naming with face cues
Cultural priming
Congruence effect
fMRI

ABSTRACT

This fMRI study explores how nonlinguistic cues modulate lexical activation in the bilingual brain. We examined the influence of face race on bilingual language production in a picture-naming paradigm. Chinese–English bilinguals were presented with pictures of objects and images of faces (Asian or Caucasian). Participants named the picture in their first or second language (Chinese or English) in separate blocks. Face race and naming language were either congruent (e.g., naming in Chinese when seeing an Asian face) or incongruent (e.g., naming in English when seeing an Asian face). Our results revealed that face cues facilitate naming when the socio-cultural identity of the face is congruent with the naming language. The congruence effects are reflected as effective integration of lexical and facial cues in key brain regions including IFG, MFG, ACC, and caudate. Implications of the findings in light of theories of language processing and cultural priming are discussed.

1. Introduction

Nonverbal cues, such as facial expressions and oral movements, are important for linguistic communications, especially in multilingual contexts where speakers and listeners are not from the same linguistic community. A large number of studies have indicated that listeners integrate both visual and auditory information in speech perception, as demonstrated by the well-known McGurk effect (Dick, Solodkin, & Small, 2010; McGurk & MacDonald, 1976; Skipper, Van Wassenhove, Nusbaum, & Small, 2007). Also, facial cues may facilitate language comprehension in bilingual or multilingual environments. For example, Sueyoshi and Hardison (2005) showed that advanced second language learners of English exhibit enhanced performance in an audiovisual English sentence comprehension task when the auditory information is presented together with faces. Facial cues can provide critical information regarding the linguistic and socio-cultural identity of the speaker or the listener (the interlocutor). Young children also use facial cues from an early age in language acquisition, as shown in recent studies of infants who discriminate between two languages on the basis of speaking faces in silent videos (e.g., Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012; Weikum et al., 2007).

Throughout our human history, face cues like race have likely been a strong predictor of the language or languages that a speaker or listener knows (e.g., Latino face -> Spanish). However, with increasing globalization, face race may no longer be as reliable a predictor of the language of the speaker or listener, especially in a multi-cultural and multilingual society. In fact, face cues such as race could also potentially interfere with second language processing. In a recent study Zhang, Morris, Cheng, and Yap (2013) tested Chinese–English bilinguals in the United States. The experimenters asked bilingual participants to engage in a computer-mediated conversation while viewing a photograph of the imagined interlocutor, who was either Caucasian or Chinese. Interestingly, bilingual participants spoke their second language (English) less fluently after viewing an image of an imagined Chinese interlocutor than a Caucasian interlocutor. The authors interpreted this finding to reflect implicit cultural priming. That is, face cues and other culturally-laden icons or symbols can automatically prime activation of cultural and linguistic schemas, especially the socio-cultural and linguistic identity of the interlocutor. In so doing, the bilingual speaker may elevate the accessibility of the native language (in this case Chinese), which disrupts the fluency of production in the second language (English).

Understanding how facial cues facilitate or hinder speech processing in the bilingual brain is important in light of current evidence regarding bilingual language activation. There is a body of literature indicating that bilinguals simultaneously activate both languages while speaking or listening to only one language (e.g. Dijkstra & van Heuven, 1998; Marian & Spivey, 2003; see Grosjean & Li, 2013, Chapter 4 for review). Brain imaging studies have identified the neural substrates that support their ability to control the use of multiple languages during speech production, which include the inferior frontal gyrus, the inferior parietal cortex, the anterior cingulate cortex, and the basal ganglia (Abutelabi, 2008; Abutalebi & Green, 2007; Abutalebi et al., 2013; Crinion et al., 2006; Hernandez, 2009; Price, Green, & Von Studnitz, 1999). One way in which

* Corresponding author.
E-mail address: pul8@psu.edu (P. Li).

0093-934X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandl.2013.09.005

bilingual speakers use these cognitive control networks to modulate access and use of their multiple languages may be explicit and/or implicit reliance on cues to help them avoid significant interference from the non-target language; after all, bilingual speakers only rarely make mistakes and produce words from the language that they did not intend to use. Such cues could be language specific (e.g., specific words in the language being processed) and/or language non-specific from the communicative context. We suggest that faces may serve as one of these powerful nonlinguistic cues to help bilinguals in this regard. However, given the limited evidence so far, we are only beginning to understand how face cues support or impair bilingual language processes, and there are no data on the underlying neurocognitive mechanisms.

In this study, we investigated the influence of face cues in facilitating or interfering with language production in bilingual participants, and the underlying neural architecture supporting this potential influence. According to the theory of "audience design" (Clark, 1996), speakers and listeners seek a shared knowledge base, a common ground in conversation, in order for successful communication to occur. To reach the common ground, speakers often adjust the form and style of speech so that the utterances are tailored to the particular knowledge of the addressee. For example, when talking to listeners in their non-native language, we quickly adapt our speech to a level that reflects the linguistic knowledge and competence of the listener (e.g., simplifying the use of certain words or structures we think may be used by the listener, Van Engen et al., 2010). Race of the interlocutor may be a useful cue to facilitate the mutual understanding for common ground in communication. Another way in which common grounds may be achieved is via "interactive alignment" in speech production, during which interlocutors mimic each others' language patterns, including lexical items, grammatical structures, and phonological patterns during conversation (Pickering & Garrod, 2004). The function of such alignment of vocabulary, structure, and phonology is to help interlocutors converge on the same ideas more easily. Given that the "audience design" and "interactive alignment" hypotheses highlight how inferred knowledge about the interlocutor might influence our language production, it is not surprising that such theories are resonant with hypotheses about "cultural priming" that are based on social psychology studies of inter-group dynamics (Fu, Chiu, Morris, & Young, 2007). These hypotheses predict that in-group biases (favoritism due to 'being the same') facilitate cultural communication, whereas out-group interactions introduce cross-group social anxiety that potentially hinders communication (Morris & Mok, 2011; Zhang et al., 2013). Importantly, face cues can be powerful indicators of in-group versus out-group memberships and, therefore, may serve to facilitate and/or hinder communication via cultural priming.

We designed a functional magnetic resonance imaging (fMRI) study to evaluate whether faces can function as cues to modulate bilingual lexical activation during language processing. We were specifically interested in the degree to which face race (Asian versus Caucasian) may facilitate activation of the intended language or inhibit activation of the unintended language. To do so, we tested Chinese–English bilingual participants performing picture-naming tasks in both their first (L1: Chinese) and second (L2: English) languages separately. Importantly, on each trial, prior to naming a picture of a nonliving object, participants saw either an Asian or a Caucasian face.

We predicted that the race of a face could facilitate naming in both the L1 and the L2 provided there was consistency in the expected mapping between the race of the face and the language to be spoken (i.e., congruence of the face–language pairings). For example, if Chinese–English bilinguals form strong expectations about hearing Chinese upon seeing an Asian (specifically Chinese) face and about hearing English upon seeing a Caucasian face (in the target language environment, the United States), we would expect to see faster naming times in their L1 (Chinese) when they are primed with a Chinese face. Similarly, we would expect to see facilitation of naming in the L2 (English) when the Chinese–English bilinguals are primed with a Caucasian face. Moreover, we predict reduced facilitation in naming in the absence of these face cues, and potential interference in naming when participants are primed with a face–language pairing that violates these strong expectations (e.g., being primed with a Caucasian face and having to name in Chinese). We can empirically differentiate facilitation from interference effects by using a "no face" condition as the baseline for comparison. Facilitation will be reflected in faster naming latencies compared to the no face condition, whereas interference will be reflected in slower naming latencies compared to the no face condition. In addition, we expected that these differential patterns of facilitation and/or interference would be reflected in the critical brain structures that modulate the control of multiple languages, such as the anterior cingulate gyrus, inferior frontal gyrus, and the basal ganglia (see Abutalebi, 2008; Abutalebi & Green, 2007; Abutalebi et al., 2013). Finally, in our design we also included a monolingual comparison group to help evaluate and interpret the bilingual results. We predicted that face cues would not be as relevant to the monolingual native speakers in our naming task, and as a result, we expected to see much less facilitation of naming responses when Caucasian faces preceded the naming in English, as well as less activation in the language control systems due to the monolingual group's lack of experience with multiple languages.

2. Methods

2.1. Participants

Fifteen Chinese–English bilinguals (eight females; mean age of 24.44 ± 3.43 years) and eleven monolingual English speakers (five females; mean age of 21.36 ± 3.08 years) from the Pennsylvania State University participated in the experiment and received payment for their participation. Both groups of participants were right-handed as judged by the Snyder and Harris's (1993) handedness questionnaire. Informed written consents were obtained from participants before the experiment. The study was approved by the Institutional Review Board of the Pennsylvania State University, and it followed the research and ethics protocols used by the Penn State Social, Life, and Engineering Sciences Imaging Center.

A language history questionnaire (Li, Sepanski, & Zhao, 2006) was used to measure self-reported language learning history and proficiency in the bilingual and monolingual participants. These bilingual participants started to learn English as their second language at an average age of 10.64 (±2.59) years. Their average exposure per day to Chinese (L1) was estimated at 52% and to English (L2) at 48% for their daily activities. They were also asked to rate their English language abilities on a scale of 1 ("not fluent at all") to 7 ("very fluent"), and their self-reported scores for English reading, writing, speaking, and listening abilities were 5.14 (±0.86), 4.64 (±0.63), 4.5 (±0.94), and 4.86 (±0.86), respectively.

Both the bilingual and the monolingual participants received a standardized test of English receptive vocabulary, the Peabody Picture Vocabulary Test (PPVT-4) (Dunn & Dunn, 1997; Dunn & Dunn, 2007). The native English monolingual speakers differed significantly from the Chinese–English bilinguals in the size of the English vocabulary as measured by PPVT-4 (English monolinguals: mean score = 109 ± 14.99, Chinese–English bilinguals: mean score = 77 ± 7.1; $t_{24} = -7.33$, $p < .001$).

2.2. Design and materials

All participants performed a picture-naming task in a blocked functional MRI design experiment. The bilingual participants