Cognition 152 (2016) 9–19

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

The influence of bilingualism on statistical word learning

Timothy J. Poepsel *, Daniel J. Weiss

Department of Psychology and Program in Linguistics, Pennsylvania State University, University Park, PA, United States

ARTICLE INFO

Article history:
Received 12 February 2015
Revised 1 March 2016
Accepted 4 March 2016

Keywords:
Statistical learning
Bilingualism
Mutual exclusivity
Cross-situational learning
Word learning

ABSTRACT

Statistical learning is a fundamental component of language acquisition, yet to date, relatively few studies have examined whether these abilities differ in bilinguals. In the present study, we examine this issue by comparing English monolinguals with Chinese–English and English–Spanish bilinguals in a cross-situational statistical learning (CSSL) task. In Experiment 1, we assessed the ability of both monolinguals and bilinguals on a basic CSSL task that contained only one-to-one mappings. In Experiment 2, learners were asked to form both one-to-one and two-to-one mappings, and were tested at three points during familiarization. Overall, monolinguals and bilinguals did not differ in their learning of one-to-one mappings. However, bilinguals more quickly acquired two-to-one mappings, while also exhibiting greater proficiency than monolinguals. We conclude that the fundamental SL mechanism may not be affected by language experience, in accord with previous studies. However, when the input contains greater variability, bilinguals may be more prone to detecting the presence of multiple structures.

1. Introduction

Statistical learning can be described as the process of detecting structure by monitoring distributional information available in the sensory input. For the past two decades, research on statistical learning has had a dramatic impact on our understanding of language acquisition. Yet despite many advances in this line of inquiry, very few investigations have approached this problem from the perspective of bilingualism. In order to acquire two languages, bilinguals must be able to establish and maintain multiple statistical representations. This experience could influence how bilinguals approach new statistical information (see Weiss, Gerfen, & Poepsel, 2015). Consequently, in the present study we endeavor to explore whether there are consequences of bilingualism for statistical word learning.

To date, only a handful of studies have compared statistical learning in bilinguals relative to monolingual abilities and thus far the results have been mixed. Perhaps the most straightforward investigation was conducted by Yim and Rudoy (2013). They tested monolingual and sequential bilingual children (who acquired their second language after age 3) between 5 and 13 years of age on a nonlinguistic auditory tones task, as well as a visual statistical learning task. There was no advantage for bilinguals on either task as learning was equivalent across both groups. This suggests that the most fundamental sequential statistical learning abilities may not be influenced by multi-language exposure. By contrast, Wang and Saffran (2014) found that adult bilingual learners were advantaged relative to monolinguals when tracking an artificial speech stream that contained compatible syllabic transitional probabilities and tonal cues to word boundaries. The authors note that the tones appear to have increased the difficulty of the segmentation task rather than simplified it, and therefore may have required suppression in order to successfully segment the stream. This conjecture accords with the observation that bilinguals who are not proficient in a tone language outperformed Chinese monolinguals on this task. Further, Bartolotti, Marian, Schroeder, and Shook (2011) presented participants with a statistical learning task using International Morse Code. Participants listened to two Morse Code languages in the context of either a high or low interference condition (a competing pause cue conflicted with the statistics in one condition and reinforced it in the other; see also Weiss, Gerfen, & Mitchel, 2010). Bilingual experience improved performance in the low interference condition, and inhibitory control (as measured by the Simon task) correlated with improved learning when interference was high. The authors suggest that the improvement shown by bilingual learners may stem from a bilingual advantage in phonological working memory (e.g., Majerus, Poncelet, van der Linden, & Weekes, 2008; see also Misyak & Christiansen, 2007). Similarly, Nation and McLaughlin (1986) found a bilingual advantage for implicit learning using an artificial grammar-learning task. They reported that multilingual learners were better at acquiring the grammar when they did not explicitly attend to the rules (there

* Corresponding author at: 007 Moore Bldg., Psychology Department, Penn State University, University Park, PA 16802, United States.
E-mail address: tjp19@psu.edu (T.J. Poepsel).

http://dx.doi.org/10.1016/j.cognition.2016.03.001
0010-0277/

was no advantage when they did). In sum, the differences reported to date for statistical learning between monolinguals and bilinguals have been quite nuanced. Our goal was to extend this literature in two ways: first, by comparing functionally monolingual and bilingual performance in a new domain of inquiry, namely statistical word learning; and second, by providing learners with the opportunity to acquire multiple sets of statistics, a situation that may mirror the real-world challenges confronting bilinguals.

1.1. Statistical word learning

A primary challenge for learning words is mapping them to their correct referents. This task is complex because words can potentially refer to any object, feature, or event in an environment (e.g., Quine, 2013). Accordingly, a prominent suggestion in the literature has been that learners may be constrained in the types of word-object mappings that they will consider. For example, it has been proposed that language learners may have a preference for assigning novel labels to novel objects (Markman & Wachtel, 1988), a preference for labeling whole-objects (Markman, 1991) and may also be limited by social-pragmatic constraints (e.g., Baron-Cohen, 1997; Clark, 1987; Diesendruck & Markson, 2001; Tomasello & Barton, 1994). However, constraining the problem space is not the only tool for word learners to alleviate the word-world mapping problem. Statistical learning has recently been proposed as another mechanism that helps learners overcome the challenge of indeterminacy (e.g., Yu & Smith, 2007). Word meanings may seem ambiguous in the context of one learning environment, yet if learners can aggregate information across multiple environments then statistical information (such as co-occurrence probabilities) may help them disambiguate which words belong with which objects.

This idea was modeled using a cross-situational statistical learning (CSSL) paradigm introduced by Yu and Smith (2007). In their initial study, participants were shown multiple scenes in which two to four objects were displayed on a computer screen while their corresponding labels were played in random order (note that the location of an object on the screen was not related to the position of its label in the auditory stream). Due to this randomization, learners could only assign words to their objects by aggregating information across multiple scenes. That is, since words and objects appeared multiple times in different visual and auditory contexts throughout familiarization (i.e., with different non-target objects and thus with different sets of labels), learners could infer that the most frequently and reliably co-occurring words and objects cohered as pairings. This task has yielded successful learning by both adult and child learners (Fazly et al., 2010; Fitneva & Christiansen, 2011; Kachergis, Yu, & Shiffrin, 2009; Smith & Yu, 2008; Vlach & Sandhofer, 2014; Yu & Smith, 2007).

We note that there has been considerable debate as to whether learning in this task is best described by statistical accumulation of multiple label-object pairings across trials (e.g., McMurray, Horst, & Samuelson, 2012; Vouloumanos, 2008; Yurovsky, Fricker, Yu, & Smith, 2014) or by forming hypotheses related to individual referents (e.g., Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013). One possibility is that task difficulty might determine which strategies learners adopt, as many of the aforementioned studies use different experimental paradigms (see Yurovsky & Frank, in review). While this debate is outside the scope of the present study, we note that the modified procedures employed here are most consistent with studies that are thought to rely on statistical accumulation rather than hypothesis-testing (e.g., see Yu & Smith, 2012).

1.2. Bilingual word learning

For bilinguals, the challenges of word learning are compounded by multiple mappings. These can take the form of translation equivalents (e.g., learners must realize that 'dog' and 'chien' both describe a four-legged pet canine) as well as interlingual homographs (i.e., "false friends", such as the word 'tuna' which refers to a fish in English and a pear in Spanish). While monolingual learners are also confronted with similar challenges in the form of synonymy and polysemy, for bilinguals such multiple mappings are compounded as they are encountered both within each language as well as across languages. Since at least half of the world's population is bilingual, an important question for word-learning research is how learners accommodate bilingual input which routinely violates assumptions of mutual exclusivity (Byers-Heinlein & Werker, 2009; Grosjean, 2008, 2010; Marian & Shook, 2012). One possibility is that bilingual learners are not constrained in the same manner as monolinguals when approaching the word-learning situation. In that vein, a number of recent word-learning studies suggest that the extent to which mutual exclusivity develops may depend on the input that a learner receives. For example, in a study with monolingual, bilingual and trilingual infants, Byers-Heinlein and Werker (2009) demonstrated that 17–18 month-old infants with exposure to multiple languages showed less disambiguation in the context of many-to-one word mappings. Furthermore, this effect was greater for trilinguals than bilinguals, suggesting that increased exposure to language variation predicts less reliance on an assumption of mutual exclusivity in mapping. Houston-Price, Caloghiris, and Raviglione (2010) noted a similar finding in a study with monolingual and bilingual infants using a broader age range (17–22 months). These results are consistent with the computational modeling efforts of McMurray et al. (2012). In their model, the development of a mutual exclusivity preference crucially depends on how many translation equivalents are encountered. We note, however, that to the best of our knowledge, the studies suggesting bilinguals may relax the mutual exclusivity constraint have focused on early or simultaneous bilinguals, and thus it is unknown whether later exposure to a second language might similarly impact learning style.

In the broadest sense, the relaxation of the mutual exclusivity constraint by early bilinguals can be understood within the framework of "learning to learn", a concept that dates back to the early behavioral learning literature. Several discrimination learning studies have demonstrated that when learners (in these studies, rats) receive repeated reversal training, they are more likely to reverse their choice when they encounter a new reversal (Dufort et al., 1954; Krechevsky, 1932; Williams, 1968; summarized in Gallistel, Mark, King, & Latham, 2001). More recently, Gallistel et al. (2001) extended these findings by testing how learners adapt to variability in reward rates and found that the frequency of change in the environment was strongly predictive of the adaptation rate. That is, the learners that experienced more frequent change were able to accommodate change faster than those who experienced less frequent change. Thus, at a very fundamental level, it can be argued that developing a prior expectation for change in a learning environment may enhance the ability to detect changes in new environments (see Qian, Jaeger, & Aslin, 2012 for further discussion of this topic).

In the present study, we investigated whether the statistical learning mechanisms that facilitate word learning might similarly be impacted by the nature of the input to learners. Specifically, we sought to determine whether late bilingual learners perform differently than monolinguals in the cross-situational statistical learning paradigm. Since even late bilinguals contend with an added layer of variability in their mappings (corresponding to the labels generated by each language), we hypothesized that this may impact