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
Available online 22 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

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 their statistical learning abilities. In particular, we were interested in exploring this phenomenon when the input affords the learner an opportunity to form multiple statistical mappings, such as when multiple objects could be mapped to a single word. Bilinguals may be more likely to assume that there are multiple causal models generating the surface statistics. To the best of our knowledge, this notion has yet to be formally tested in the context of statistical learning. Consequently, we approached this problem by comparing a group of functional monolinguals with two groups of sequential bilinguals (Chinese–English and English–Spanish) who acquired their L2 subsequent to mastering their L1. This provided a rigorous test of whether proficiency with a second language could impact statistical learning even in the absence of early learning experiences with two languages that have been shown to result in a relaxation of the mutual exclusivity constraint (e.g., Byers-Heinlein & Werker, 2009). We first explored whether there are differences between functional monolinguals bilinguals on cross-situational learning in the context of one-to-one mappings (Experiment 1). Next, we tested whether these groups differed when the input afforded two-to-one mappings for a subset of objects and labels (Experiment 2).

2. Experiment 1

In the initial experiment we explore whether functionally monolingual learners, Chinese–English bilinguals, and English–Spanish bilinguals might differ in their abilities to track statistical information across scenes in a CSSL task. There were three conditions that varied in the number of items presented simultaneously (ranging between 2 and 4).

2.1. Participants

Seventeen students (11 female, 6 male; mean age 19.7 years, SD = 1.4) at Penn State University were given course credit for their participation in this experiment. Based on language history questionnaire (LHQ) data, these participants were native speakers of English who self-rated their English proficiency at an average of 9.6 (SD = .79) on a 10-point scale, on which 10 indicated maximum proficiency. Due to a foreign language requirement at Penn State University for undergraduates, all participants indicated exposure to a second language in the course of their education. Participants self-rated their second-language proficiency at an average of 1.2 (SD = 1.4) on the ten-point scale, and all participants rated below a 4. As such, we considered these participants to be functionally monolingual.

Seventeen Chinese–English bilinguals (14 female, 3 male; mean age 22.2 years, SD = 1.4) from Beijing Normal University in Beijing, China also participated for payment. Participants self-rated their Mandarin proficiency at an average of 9.3 (SD = 1.2) on the same 10-point scale used above. Bilinguals began learning English at the age of 11.1 years (SD = 2.1) and self-rated their English proficiency at an average of 6 (SD = 1.3).

Sixteen Penn State students (15 female, 1 male; mean age 21 years, SD = .78) who were English–Spanish bilinguals also participated for payment. These participants self-rated their English proficiency at an average of 9.9 (out of 10; SD = .27), their Spanish proficiency at an average of 6.9 (SD = .89) out of 10, and began learning their L2 at an average age of 10.9 years (SD = 4.4).

2.2. Stimuli

The stimuli for Experiment 1 consisted of fifty-four unique word-object pairs created by randomly pairing novel objects with nonce words. The objects consisted of black and white complex line drawings. Eight objects appeared in the stimuli used by Creel, Aslin, and Tanenhaus (2011) and served as a template for creating the remaining 46 objects (using MS Paint ©). All objects were converted to a .jpeg file format with a size of 150 × 150 pixels.

Nonce words consisted of an equal distribution of monosyllabic, disyllabic, and trisyllabic items chosen from the English Lexicon Project (ELP) non-word database (http://elexicon.wustl.edu) (see Table 1 for a full listing of the nonce words used in Experiment 1 and 2). All nonce words had American English phonological patterns, were between 4 and 10 characters in length, and based on data from the ELP had an average of 2.2 orthographic neighbors and a bigram mean of 2022. The words were created in a female American English voice (*Crystal*) via the AT&T Natural Voices text-to-speech synthesizer (http://www.naturalvoices.att.com), and converted into WAV files sampled at 22,050 Hz. The fifty-four word-object pairs were separated into three non-overlapping sets of eighteen pairs, in which word length was equally distributed.

2.3. Procedure

During familiarization, participants watched objects appear on a computer screen while listening to words presented over speakers. Each participant completed three familiarization phases, each containing 18 unique word-object pairs, distinguished by the number of words and objects presented in a trial; there was a 2 × 2 condition (in which participants saw two objects and heard two words), a 3 × 3 condition, and a 4 × 4 condition. The order in which participants encountered these familiarizations was randomized, and the set of words and objects presented in each familiarization was non-overlapping. Preceding each trial, a fixation cross appeared for 750 ms. During the trial, two to four objects appeared simultaneously while the corresponding nonce words were played serially at 3 s intervals. The onset of the visual presentation of objects was synchronized with the presentation of the first word of the trial. There was no systematic relationship within a trial between the placement of an object in the visual array and the location of its corresponding word in the auditory stream; object locations and word orders were randomly assigned. Progression through the trials was automatic in that the end of one trial cued the presentation of the next. The order of the trials was pseudo-randomized such that no word-object pair appeared in two consecutive trials. Within each condition, every word-object pair was presented 6 times, and across all trials participants saw

Table 1
Nonce words used in Experiments 1 & 2, organized by syllable count.

Monosyllabic	Bisyllabic	Trisyllabic
barsh	briskle	baturate
blep	crinklow	calorix
chost	dounger	caprion
crid	durrow	clamoreck
daint	grinter	coronick
drock	haser	haterfront
dulch	lattle	interlade
feech	masset	jatterside
frane	mubble	latercress
glack	murler	naureate
glink	pangle	overlood
gotch	patchet	perminal
plock	peadle	rentacle
plunt	pedline	tanderer
scown	pritter	thermistar
slute	tallot	todular
sunch	tarren	tonogram
veam	thecker	ventuker

a total of 108 objects and heard 108 words. Accordingly, the total number of trials varied by condition: there were fifty-four trials in the 2 × 2 condition, 36 in the 3 × 3, and 27 in the 4 × 4. Total time of familiarization was constant at 320 s across all conditions. Prior to beginning the experiment, participants were told that they would be learning novel names for novel objects.

Following each familiarization phase, participants completed a 4 alternative forced-choice (4AFC) test consisting of 18 test trials (i.e., one trial for each word-object pair presented during familiarization). On each test trial, participants viewed four objects (in randomized positions) while hearing a single word. Three of these objects were distractors randomly selected from the set of objects presented within a given condition. The remaining object was the correct referent for the presented word. All objects within a test trial were presented simultaneously, with one object located in each corner of the screen and labeled with a number (1–4). Participants were asked to press the number key corresponding to the correct object without any time limit.

2.4. Experiment 1: Results

Functionally monolingual English participants in Experiment 1 learned 87% (SD = 17%) of word-object pairs in the 2 × 2 condition; 70% (SD = 27%) of pairs in the 3 × 3 condition, and 53% (SD = 19%) of pairs in the 4 × 4 condition. Chance performance was 25% (since there were four alternatives at test). Single-sample t-tests confirmed that performance in each of the three conditions was significantly above chance (2 × 2: $t(16) = 15.3$, $p < .001$; 3 × 3: $t(16) = 6.8$, $p < .001$; 4 × 4: $t(16) = 5.9$, $p < .001$). Chinese–English bilinguals learned 78% (SD = 19%) of word-object pairs in the 2 × 2 condition; 66% (SD = 24%) of pairs in the 3 × 3 condition, and 38% (SD = 13%) of pairs in the 4 × 4 condition. Single-sample t-tests confirmed that performance in each of the three conditions was significantly above chance (2 × 2: $t(16) = 11.2$, $p < .001$; 3 × 3: $t(16) = 7.1$, $p < .001$; 4 × 4: $t(16) = 4.3$, $p < .01$). English–Spanish bilinguals learned 86.5% (SD = 11%) of word object pairs in the 2 × 2 condition, 73% (SD = 16%) of pairs in the 3 × 3 condition and 56% (SD = 20%) of pairs in the 4 × 4 condition. Single-sample t-tests confirmed that performance in all three conditions for English–Spanish bilinguals was significantly above the level of chance (2 × 2: $t(15) = 23.2$, $p < .001$; 3 × 3: $t(15) = 11.4$, $p < .001$; 4 × 4: $t(15) = 6.2$, $p < .001$) (see Fig. 1).

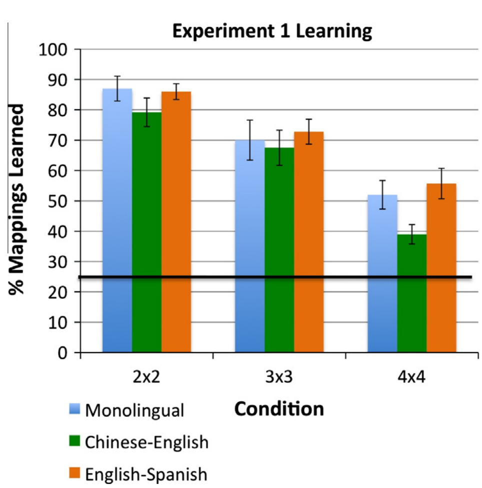

Fig. 1. Results from all conditions in Experiment 1 for English monolinguals, Chinese–English bilinguals, and English–Spanish bilinguals, where the black line represents chance performance (25%). Learning is equivalent among the groups in the 2 × 2 and 3 × 3 conditions, although in the 4 × 4 condition the Chinese–English bilinguals acquired significantly fewer mappings (see text). Error bars represent 1 SE of the mean.

A Shapiro–Wilk test confirmed that participants' performance across Experiment 1 was normally distributed ($W(50) = .97$, $p = .22$). We used a 3 (Group) × 3 (Condition) repeated measures ANOVA to investigate factors that influence learning in the CSSL task. Group (English monolingual, Chinese–English bilingual, English–Spanish bilingual) was a between-subjects factor, while Condition (2 × 2, 3 × 3, or 4 × 4) was a within-subjects factor. There was a main effect of test ($F(2,94) = 71.9$, $p < .001$, $\eta^2 = .6$), such that learning was greatest in the 2 × 2 condition ($M = 84\%$, SE = 2.3%), followed by the 3 × 3 condition ($M = 70\%$, SE = 3.3%) and the 4 × 4 condition ($M = 50\%$, SE = 2.5%). The interaction between test and group was not significant ($F(4,94) = .88$, $p = .48$, $\eta^2 = .04$), nor was the between-subjects factor of Group ($F(2,47) = 2.6$, $p = .08$, $\eta^2 = .09$), indicating that English monolinguals, Chinese–English bilinguals and English–Spanish bilinguals were not significantly different in their learning performance across the three conditions of Experiment 1 (L1 English: $M = 70.3\%$, SE = 3.7%; Chinese–English bilinguals: $M = 61.2\%$, SE = 3.7%; English–Spanish bilinguals: $M = 72.4\%$, SE = 3.8%). A series of one-way ANOVAs compared performance within each learning condition (i.e., 2 × 2, 3 × 3, 4 × 4) between the samples of English monolinguals, Chinese–English bilinguals, and English–Spanish bilinguals. The difference was not significant in the 2 × 2 condition ($F(2,47) = 1.9$, $p = .2$) or 3 × 3 condition ($F(2,47) = .37$, $p = .7$), but was significant in the 4 × 4 condition ($F(2,47) = 4.8$, $p < .05$). Bonferroni corrected post-hoc tests confirmed that there were no differences between the groups in either the 2 × 2 or 3 × 3 conditions (all $ps > .23$), no differences between monolinguals and English–Spanish bilinguals in the 4 × 4 condition ($p > .99$), and significant differences between Chinese–English bilinguals and both monolinguals ($p = .04$) and English–Spanish bilinguals ($p = .012$) in the 4 × 4 condition, as the Chinese–English bilinguals acquired fewer mappings than the other two groups.

A final analysis examined whether the number of syllables in a word as well as the first language of the participant (i.e., English or Chinese) influenced learnability. We used a 3 (Number of Syllables) × 2 (L1 English or L1 Chinese) ANOVA to investigate this issue. Number of Syllables was a within-subjects factor, while L1 was a between-subjects factor. There was a main effect of number of syllables ($F(2,96) = 5.9$, $p < .01$, $\eta^2 = .11$) indicating that learning of monosyllabic words was significantly higher than bisyllabic and trisyllabic words (mono: $M = 70.4\%$, SE = 2.4%; bi: $M = 63\%$, SE = 2.6%; tri: $M = 65.1\%$, SE = 2.6%). The between-subjects factor of L1 was also significant ($F(1,48) = 5.2$, $p < .05$, $\eta^2 = .1$), such that learning for L1 English participants ($M = 72.4\%$, SE = 3%) was more robust than that of L1 Chinese participants ($M = 61.2\%$, SE = 4%) across all conditions. The interaction between Number of Syllables and L1 was also significant ($F(4,96) = 3.9$, $p < .05$, $\eta^2 = .08$). Planned follow-up tests investigating this interaction showed that L1 English participants did not perform differently based on the number of syllables ($F(2,64) = 1$, $p = .36$, $\eta^2 = .03$; monosyllabic: $M = 73.3\%$, SE = 3.1%; bisyllabic: $M = 72\%$, SE = 3.1%; trisyllabic: $M = 70\%$, SE = 3.3%), while L1 Chinese participants performed significantly higher on monosyllabic words relative to bisyllabic or trisyllabic words ($F(2,34) = 6.9$, $p < .01$, $\eta^2 = .3$; monosyllabic: $M = 68\%$, SE = 3.4%; bisyllabic: $M = 55\%$, SE = 4.3%; trisyllabic: $M = 61\%$, SE = 4.1%).

2.5. Experiment 1: Discussion

The overall findings from Experiment 1 suggest that monolinguals and bilinguals do not significantly differ with respect to their ability to engage in cross-situational statistical learning. The only difference in performance between the three groups emerged in the 4×4 condition, which was the most difficult condition for all participants. Given that performance in the 2×2 and 3×3 conditions was equivalent across groups, and that the English–Spanish bilinguals performed equivalently to the English monolinguals on all conditions, we doubt that the difference in the 4×4 condition reflects a true difference in statistical learning abilities. Rather, the stimuli conformed to English phonology and this may have advantaged native English-speaking participants, as evidenced by the decrement in performance on multisyllabic words that occurred for Chinese–English bilinguals but not for the native English speakers. In the 4×4 condition, the probability of encountering a multisyllabic word in every scene was higher than in the other conditions. This could explain why this condition was particularly hard for the Chinese–English bilinguals relative to the English monolinguals and English–Spanish bilinguals. Future experiments using words that conform to the native phonology of the Chinese–English bilinguals could further elucidate the source of this performance difference.

3. Experiment 2

Having discovered relatively similar levels of performance by functionally monolingual and bilingual learners on the standard version of CSSL, in Experiment 2, we extended the paradigm to provide learners with the opportunity to form multiple mappings. Several previous studies have investigated whether learners obey mutual exclusivity in CSSL when they have the opportunity to map one item with multiple objects or labels. The results of these studies suggest that there is a bias toward mutual exclusivity.

Yurovsky, Yu, and Smith (2013), for example, exposed learners to an initial set of 1:1 mappings and in a subsequent training session remapped a subset of words to new objects. Learners were able to acquire both the first (primacy) and second (recency) referent of remapped words, but in direct preference tests demonstrated a bias towards the primacy referent. A follow-up condition found that learners could acquire two mappings for a word within the same training session. Similarly, Ichinco, Frank, and Saxe (2009) presented learners with a set of 18 one-to-one mappings in a 3×3 design, followed by a second training phase that included a new set of one-to-one mappings as well as a subset of items transferred from the first familiarization phase. These items appeared with their original mapping but could also be mapped to one of the new items. In this way, learners were afforded the opportunity to form a second mapping for the transferred item, but still encountered information consistent with the primacy mapping. Learners in this paradigm preferred the primacy mapping and did not acquire the new (recency) mapping. A more recent study from our lab replicated this result and also found that the addition of contextual cues to the second familiarization (e.g., a change in speaker voice) facilitated acquisition of the recency mapping and significantly reduced participants' preference for primacy mappings (Poepsel, Gerfen, & Weiss, 2012). Finally, Kachergis et al. (2009) modeled CSSL in environments presenting learners with two-to-one mappings, and concluded that learners could modulate their reliance on mutual exclusivity in response to the proportion of mappings in the input that either followed or violated ME. Increased exposure to mappings that violated mutual exclusivity predicted better acquisition of new mappings.

In Experiment 2 we compared the learning performance of a group of functionally monolingual English speakers to late-learning Chinese–English and English–Spanish bilinguals using a CSSL task that presented learners with a mix of one-to-one and two-to-one mappings. As noted previously, there is empirical support in the developmental literature for the idea that bilinguals may be more likely to assume multiple underlying structures (here mappings) relative to monolinguals as a consequence of frequent exposure to multiple languages (e.g., Byers-Heinlein & Werker, 2009; Houston-Price et al., 2010; Kovács & Mehler, 2009), but this was in the context of tasks that did not require statistical learning and the participants all had early exposure to multiple languages. We chose to study late learning bilinguals to investigate whether experience with multiple languages could influence statistical learning, even for learners who shared similar early L1 experiences with monolinguals (i.e., sequential bilinguals presumably did not relax the mutual exclusivity constraint during development).

We chose to implement a modified version of the task used by Yurovsky et al. (2013), as this paradigm facilitated learning of multiple mappings (with a preference for the mutual exclusivity mapping) and could therefore permit measures of learning related to the amount of exposure. In order to assess whether there might be differences in how quickly the mappings were acquired, we provided participants with a series of three distinct familiarization phases, each followed by a test. Following Yurovsky et al. (2013), we provided learners with multiple words mapped to a single object. This conferred a practical advantage relative to mapping multiple objects to a single word in that the new objects appeared on the screen throughout the trial, whereas novel words would have been highly transient (and in our previous work, remapping words was more effective; Poepsel et al., 2012). With respect to bilingual acquisition, interlingual homographs are known to be challenging, particularly when presented in conjunction with cognates (see Brenders, van Hell, & Dijkstra, 2011). Since interlingual homographs occur less frequently than translation equivalents, this manipulation arguably provided a subtler test of differences in mapping abilities between monolinguals and late-learning bilinguals.

3.1. Participants

Sixteen English monolinguals (11 female, 5 male) from Penn State University who did not participate in Experiment 1 participated for course credit. Based on language history questionnaire data, these participants had a mean age of 18.6 years (SD = 0.65) and self-rated their proficiency in English at 10 (SD = 0) on the ten-point scale used above. Six of these participants had been classroom learners of Spanish, who began receiving instruction at a mean age of 13.2 years (SD = 1.5) and self-rated their L2 proficiency at a mean of 2.5 (SD = 1.2) on the same ten-point scale.

Sixteen Chinese–English bilinguals (13 female, 3 male) from Beijing Normal University in Beijing, China who did not participate in Experiment 1 participated for payment. Based on language history questionnaire data, these participants had a mean age of 22.9 years (SD = 2.4), began learning English at an average age of 10 (SD = 4), and self-rated their proficiency in English at 6.6 (SD = 1.4) on a ten-point scale.

Sixteen English–Spanish bilinguals (14 female, 2 male) from Penn State University who did not participate in Experiment 1 participated for payment. These participants had a mean age of 20.2 years (SD = 1.1) and self-rated their overall proficiency in English at 10 out of 10 (SD = 0). These participants indicated that they began learning their L2 at a mean age of 12.6 years (SD = 1.8) and rated their L2 proficiency at an average of 6.4 out of 10 (SD = 1.8).

3.2. Stimuli

Stimuli consisted of 18 nonce words and 24 novel objects, grouped into twelve 1:1 (one object to one word) mappings and six 2:1 (two objects to one word) mappings. All stimuli were chosen from the 54 word-object pairs used in Experiment 1. Due to the results of Experiment 1, words chosen for this experiment were all monosyllabic, between four and five characters in length, and contained between three and five phonemes. They followed one of four syllable patterns (CVC, CCVC, CVCC, CCVCC). Based on data from the ELP, nonce words had an average of 4.3 orthographic neighbors and a bigram mean of 4983.

3.3. Procedure

Participants completed three familiarization phases, each of which was followed by a test phase (see Fig. 2). Within each familiarization phase, participants were exposed to the same set of twelve 1:1 and six 2:1 object-word mappings across twenty-four training trials. Before each trial, a fixation cross appeared for 500 ms. Trials presented three objects and three words as in the 3×3 condition of Experiment 1. Participants were exposed to four instances of each 1:1 mapping within a familiarization. Each pairing for the 2:1 mappings (e.g., object A – word 1; object B – word 1) was presented twice. The ordering of the 24 trials within each familiarization was pseudo-randomized such that no word-object pair appeared in consecutive trials.

After each familiarization phase, participants completed a 2AFC test in which the learning of both 1:1 and 2:1 mappings was assessed. Test trials in Experiment 2 were largely similar in their presentation to those of Experiment 1. On each test trial, participants heard a single word and saw two objects, one a distractor and one the correct referent of the presented word. The order of trials was randomized in all three tests.

For each test following the first and second familiarizations, participants completed eighteen test trials. All twelve 1:1 mappings were tested once. Participants also received one test trial for each of the six 2:1 mappings; each possible referent (i.e., primacy and recency) was tested once across the first two tests, and the order in which both referents were tested was pseudorandomized and counterbalanced across participants. That is, if the primacy mapping was tested after the first familiarization (i.e., the first label for an object encountered by the learner), the recency mapping (i.e., the second label for an object) was tested after the second familiarization, and vice versa. This procedure was instantiated in order to ensure that participants were not explicitly cued to the presence of multiple mappings for some objects and not reinforced for one mapping over another. Moreover, in each of the first two tests, half of the 2:1 mapping trials probed primacy mappings, while the other half probed recency mappings. In the test following the third familiarization, participants completed 36 test trials; each 1:1 mapping was tested twice, and both the primacy and recency referents of words with 2:1 mappings were tested once, as the aforementioned concern was no longer relevant. Participants were given as much time as needed to make each response.

3.4. Experiment 2: Results

In the test following the 1st familiarization, English monolinguals learned 70% (SD = 13%) of 1:1 object-word mappings and were accurate on 56% (SD = 21%) of trials probing 2:1 mappings, Chinese–English bilinguals learned 71% (SD = 13%) 1:1 mappings and were accurate on 61% (SD = 20%) of trials probing either side of the 2:1 mappings, and English–Spanish bilinguals learned 72% (SD = 15%) of 1:1 mappings and were accurate on 66% (SD = 15%) of trials probing either side of the 2:1 mappings. In the test following the 2nd familiarization, English monolinguals learned 81% (SD = 14%) of 1:1 mappings and were accurate on 70% (SD = 20%) of trials probing 2:1 mappings, Chinese–English bilinguals learned 84% (SD = 14%) of 1:1 mappings and were accurate on 82% (SD = 15%) of trials probing either side of the 2:1 mappings, and English–Spanish bilinguals learned 92% (SD = 12%) of 1:1 mappings and were accurate on 79% (SD = 20%) of trials probing 2:1 mappings. In the test after the third familiarization, English monolinguals learned 92% (SD = 16%) of 1:1 mappings and were accurate on 77% (SD = 22%) of trials probing 2:1 mappings, Chinese–English bilinguals learned 93% (SD = 10%) of 1:1 mappings and were accurate on 88% (SD = 12%) of trials probing 2:1 mappings, and English–Spanish bilinguals learned 97% (SD = 7%) of 1:1 mappings and were accurate on 94% (SD = 7%) of trials probing 2:1 mappings.

The learning averages reported above were compared against chance (which was set at 50% as a result of the 2AFC design) in a series of single sample *t*-tests. For monolinguals, learning was significantly above chance for 1:1 mappings in all three tests and for 2:1 mappings in both the second and third tests (all $ps < .01$). Monolinguals did not exceed chance performance on 2:1 mappings following in the first test ($t(16) = 1.1$, $p = 0.27$). For both Chinese–English (C–E) and English–Spanish (E–S) bilinguals, learning was

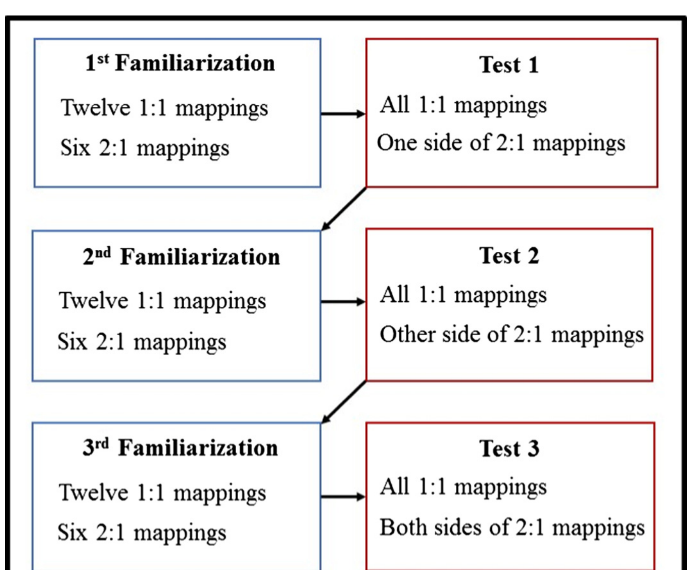

Fig. 2. Experiment 2 was composed of three familiarization phases, each of which was followed by a test. In tests 1 and 2, participants completed a trial for either the primacy or recency mapping of each 2:1 mapping. In test 3, both the primacy and recency mappings were tested.

significantly above chance for 1:1 and 2:1 mappings in all tests (C–E bilinguals all $ps < .05$; E–S bilinguals all $ps < .01$).

As our primary goal was to compare the statistical learning abilities of our monolingual participants to those of our bilinguals, we carried out several analyses to determine whether the Chinese–English and English–Spanish bilinguals were statistically equivalent in their performance in Experiment 2. Thus we compared their performance using a 2 (Group) × 3 (Test) × 2 (Mapping Type) repeated-measures ANOVA. Test and Mapping Type were within-subjects factors, while Group (C–E and E–S bilinguals) was a between-subjects factor. There was a main effect of Test ($F(2,60) = 59.3$, $p < .001$, $\eta^2 = .66$) indicating that learning performance increased from Test 1 to Test 3. There was also a main effect of Mapping Type ($F(1,30) = 21.2$, $p < .001$, $\eta^2 = .41$), demonstrating that performance on 1:1 mappings was higher than that for trials probing either side of the 2:1 mappings. Critically, the between-subjects factor of Group was not significant ($F(1,30) = 1.38$, $p = .25$, $\eta^2 = .04$), suggesting that both groups of bilinguals did not perform differently across both mapping types in Experiment 2. Furthermore, no interaction terms reached significance (all $Fs < 1.31$, all $ps > .28$, all $\eta^2s < .04$). To further verify that the bilingual groups were equivalent in performance, we ran a follow-up 2 (Group) × 3 (Test) ANOVA for each mapping type. As above, Test was a within-subjects factor and Group was a between-subjects factor. For both 1:1 mappings and 2:1 mappings, there was a significant main effect of Test (1:1 mappings: $F(2,60) = 43.4$, $p < .001$, $\eta^2 = .59$; 2:1 mappings: $F(2,60) = 25.4$, $p < .001$, $\eta^2 = .46$) such that performance increased from Test 1 to Test 3. Again, the between-subjects factor of Group did not reach significance for either mapping type (1:1 mappings: $F(1,30) = 1.96$, $p = .17$, $\eta^2 = .06$; 2:1 mappings: $F(1,30) = .54$, $p = .47$, $\eta^2 = .01$), nor did the interaction of Test and Group for either mapping type (1:1 mappings: $F(2,60) = .72$, $p = .49$, $\eta^2 = .02$; 2:1 mappings: $F(2,60) = .79$, $p = .46$, $\eta^2 = .03$) suggesting that English–Spanish and Chinese–English bilinguals performed equivalently on all mappings in Experiment 2.

Having established this equivalence, we combined these groups into a single bilingual group for comparison against the sample of English monolinguals. Subsequently, we used a 2 (Group) × 3 (Test) × 2 (Mapping Type) repeated-measures ANOVA to determine the influence of bilingual experience on learning in Experiment 2. A Shapiro–Wilk test confirmed that participants' performance across Experiment 2 was normally distributed ($W(48) = .97$, $p = .23$). Group (monolingual, bilingual) was a between-subjects factor, while Test (performance on post-tests 1, 2 and 3) and Mapping Type (1:1 or 2:1) were within-subjects factors. There was a main effect of Test ($F(2,92) = 62.9$, $p < .001$, $\eta^2 = .58$), indicating that learning increased over the course of the experiment (Test 1: $M = 65.2\%$, SE = 2%; Test 2: $M = 79.9\%$, SE = 1.9%; Test 3: $M = 88.6\%$, SE = 1.7%). We also found a main effect of Mapping Type ($F(1,46) = 52.9$, $p < .001$, $\eta^2 = .54$), indicating that participants' performance was significantly higher on trials probing 1:1 mappings than 2:1 mappings (1:1 mappings: $M = 82.8\%$, SE = 1.4%; 2:1 mappings: $M = 73\%$, SE = 1.7%). The interaction of Mapping Type and Group was also significant ($F(1,46) = 6.2$, $p < .05$, $\eta^2 = .12$) as bilinguals exhibited a smaller difference in performance between 1:1 and 2:1 mappings relative to monolinguals. The between-subjects factor of Group was significant ($F(1,46) = 7.0$, $p < .05$, $\eta^2 = .13$) providing evidence that bilinguals performed better overall in Experiment 2 relative to monolinguals (bilinguals: $M = 81.7\%$, SE = 1.7%; monolinguals: 74.1%, SE = 2.3%).

Planned follow-up tests further explored the significant Group factor, directly comparing monolinguals' and bilinguals' performance on trials probing 1:1 and 2:1 mappings in separate ANOVAs. Here we found no significant difference in performance on 1:1 mappings between the groups ($F(1,46) = 2.2$, $p = .14$, $\eta^2 = .04$), but a significant difference in performance on trials probing 2:1 mappings ($F(1,46) = 9.8$, $p < .01$, $\eta^2 = .18$), as bilinguals outperformed monolinguals on this test trial type (monolinguals: $M = 67.5\%$, SE = 2.9%; bilinguals: $M = 78.5\%$, SE = 2%). The factors for Test × Group, Test × Mapping Type, and Test × Mapping Type × Group did not reach significance (all $Fs < .85$, all $ps > .43$).

The analyses carried out above for 2:1 mappings were based on participants' accuracy on trials probing one of the two sides of these mappings; this does not necessarily provide information on how frequently participants were accurate on both sides of a 2:1 mapping. Across the first two tests, participants completed trials probing both the primacy and recency mappings for each 2:1 mapping. In the third test, both the primacy and recency mappings were probed (see Methods for further detail). In assessing performance across both sides of the 2:1 mapping, we again compared English monolinguals to a combined group of Chinese–English and English–Spanish bilinguals. Across the first two tests, monolinguals were accurate on both sides of a 2:1 mapping in 36% (SD = 23%) of cases, while bilinguals were accurate on both sides in 50% (SD = 21%) of cases. This difference was significant ($t(46) = 2.03$, $p < .05$), suggesting that bilinguals learned more 2:1 mappings than monolinguals across the first two familiarizations (see Fig. 3). In the third test, monolinguals were accurate on both sides of a 2:1 mapping 59% (SD = 33%) of the time, while bilinguals achieved an accuracy of 83% (SD = 19%). This difference was also significant ($t(46) = 3.18$, $p < .01$), again suggesting that bilinguals acquired more 2:1 mappings than monolinguals by the end of training in Experiment 2.

For words with 2:1 mappings, we were also interested in whether participants exhibited a bias toward either primacy or recency mappings and whether bilingual experience impacted mapping preferences. We used a 3 (Test) × 2 (Mapping Type) × 2 (Group) ANOVA to investigate the factors that influenced acquisition of primacy and recency mappings in Experiment 2. Test (first, second, third) and Mapping Type (primacy, recency) were within-subjects factors, while Group (monolingual, bilingual) was a between-subjects factor. There was a main effect of Test ($F(2,92) = 22.4$, $p < .001$, $\eta^2 = .33$) demonstrating that overall performance increased across the three tests (Test 1, $M = 60.7\%$, SE = 2.8%; Test 2: $M = 75.5\%$, SE = 2.8%; Test 3: $M = 83.9\%$, SE = 2.3%). There was also a main effect of Mapping Type ($F(1,46) = 12.6$, $p < .001$, $\eta^2 = .22$), as performance on primacy mappings was higher relative to recency mappings (primacy: $M = 77.7\%$, SE = 2.1%; recency: $M = 68.9\%$, SE = 2.1%). The between-subjects factor also reached significance ($F(1,48) = 10.8$, $p < .01$, $\eta^2 = .19$) suggesting that bilinguals acquired more mappings (both 1:1 and 2:1) than monolinguals across

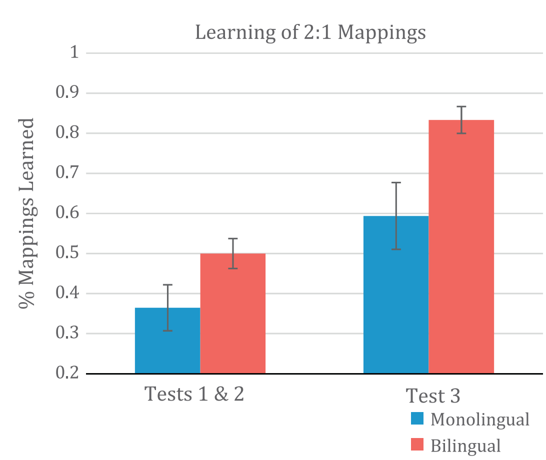

Fig. 3. Bilinguals acquired significantly more 2:1 mappings than monolinguals across the first two tests and in the third test. Error bars represent 1 SE of the mean.

Experiment 2 (bilingual: $M = 78.5\%$, $SE = 2.9\%$; monolingual: $M = 67.5\%$, $SE = 2\%$). The factors of Test × Mapping Type, Test × Group, Mapping Type × Group, and Test × Mapping Type × Group did not reach significance (all $Fs < 1.7$, all $ps > .18$).

Planned follow-up tests further investigated the interaction of Test × Mapping Type within the monolingual and bilingual groups. For monolinguals, the interaction of Test × Mapping Type was not significant ($F(2,30) = 1.9$, $p = .17$, $\eta^2 = .1$), while for bilinguals, this interaction was significant ($F(2,62) = 6.6$, $p < .01$, $\eta^2 = .18$), indicating that over the course of learning, bilinguals' performance on recency and primacy showed a greater trend of convergence than monolinguals (see Fig. 4). Another set of planned follow-up tests compared the groups' performance on each mapping type (primacy and recency). For both primacy and recency mappings, the between-subjects factor of Group was significant (primacy: $F(1,46) = 6.5$, $p < .05$, $\eta^2 = .12$; recency: $F(1,46) = 7.6$, $p < .01$, $\eta^2 = .14$), such that bilinguals performed better on each mapping type than monolinguals across Experiment 2.

A final analysis compared the performance of the two groups of English L1 participants (English monolinguals and English–Spanish bilinguals) in Experiment 2. Here we used a 3 (Test) × 2 (Mapping Type) × 2 (Group) ANOVA, setup similarly to those presented above. There was a main effect of Test ($F(2,60) = 48.6$, $p < .001$, $\eta^2 = .62$) indicating an increase in performance across tests, and a significant interaction between Mapping Type and Group ($F(1,30) = 6.0$, $p < .05$, $\eta^2 = .17$). The between-subjects factor of Group was also significant ($F(1,30) = 5.2$, $p < .05$, $\eta^2 = .15$). Two follow-up tests investigated this significant between-subjects factor, comparing both groups in their performance on each mapping type. We found no difference in performance between these groups on 1:1 mappings ($F(1,30) = 1.3$, $p = .26$, $\eta^2 = .04$), but a significant difference in performance on 2:1 mappings ($F(1,30) = 7.6$, $p < .05$, $\eta^2 = .2$), demonstrating that English–Spanish bilinguals outperformed their monolingual counterparts on this mapping type.

4. General discussion

Across two experiments, we compared a group of functional monolinguals to two groups of late-learning bilinguals (Chinese–English and English–Spanish) to determine whether there were differences in the acquisition of 1:1 and 2:1 mappings in cross-situational statistical word learning tasks. In Experiment 1, we replicated the CSSL study of Yu and Smith (2007) with our own set of stimuli. Participants learned unique sets of 1:1 mappings in three conditions that varied with respect to how many word-object pairings appeared at once. Overall, the results of Experiment 1 suggest that the three groups are very similar in their ability to acquire 1:1 mappings in a standard CSSL task. Across all conditions, there was only one significant difference in performance across the three groups. Chinese–English bilinguals had lower performance relative to the other two groups in the most difficult 4 × 4 condition. As noted above, this might be attributable to the fact that the stimuli were presented in the phonology of their L2 (in particular, multi-syllabic stimuli posed difficulties for these bilinguals).

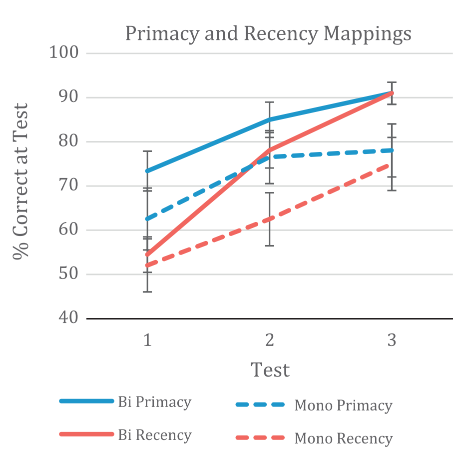

Fig. 4. Performance on primacy and recency mappings across tests for monolinguals and bilinguals. Error bars represent 1 SE of the mean.

In Experiment 2, three new samples of participants drawn from the same populations were familiarized with a mixed set of 1:1 and 2:1 mappings in a 3 × 3 cross-situational word learning design with three consecutive familiarization phases, each of which was followed by a test. As in the 3 × 3 condition of Experiment 1, all groups achieved equivalent levels of performance in learning the 1:1 mappings, providing further evidence that these groups were matched in their core cross-situational statistical learning abilities. However, Chinese–English and English–Spanish bilinguals acquired significantly more 2:1 mappings than the monolinguals across all three tests. Further, bilinguals' performance on both primacy (the first label paired with an object and recency mappings (the second label paired with a given object) was significantly higher than that of the monolinguals across all familiarizations, and bilinguals showed a stronger trend of convergence in their performance on primacy and recency mappings compared to monolinguals. Taken together, these results suggest that bilinguals may have acquired true 2:1 mappings earlier than monolinguals, who showed a more consistent bias toward primacy mappings. In sum, despite broad similarities in performance on 1:1 mappings, we found that Chinese–English and English–Spanish bilinguals seemed to acquire 2:1 mappings with less exposure and greater overall proficiency than monolinguals. Thus, our results suggest that when tracking novel statistical inputs over time, late bilingual learners appear to be more open to the possibility of multiple mappings in the input.

Our findings extend the current knowledge regarding bilingualism and statistical learning in several respects. As noted in the Introduction, to date there has been mixed evidence regarding whether monolinguals and bilinguals differ with respect to their statistical learning abilities (Bartolotti et al., 2011; Wang & Saffran, 2014; Yim & Rudoy, 2013). Whereas the previous studies have focused on the task of speech segmentation, here we extend the investigation to statistical word learning, which may involve different cognitive processes (see below). One interesting parallel that emerged between our studies and previous work is that monolinguals and bilinguals seem to perform similarly on the most straightforward versions of these tasks. For example, when studies required learners to track only the transitional probabilities between adjacent elements without any sources of interference, the results of bilinguals mirror monolinguals (Yim & Rudoy, 2013 with children; see also Bogulski, 2013 with adults). In Experiment 1, we found that on the 1:1 version of the CSSL task, performance was largely similar across two populations of late-learning bilinguals and our functional monolinguals (also for 1:1 mappings in Experiment 2), including Chinese–English bilinguals for whom English phonology may have posed an additional challenge (as evidenced by our analysis of performance based on the number of syllables in a word). While more data is required to draw a firm conclusion, evidence thus far supports the idea that the most basic forms of statistical learning, involving either tracking sequential probabilities or accruing associative information over time, may be relatively unaffected by experience with second language learning.

By contrast, the previous statistical learning studies that do report differences between monolingual and bilingual learners have all contained multiple cues to segmentation. Arguably, in theses studies learners must suppress one set of cues in order to correctly segment the stream. In the case of Bartolotti et al. (2011) transitional probability statistics were tracked along with a competing pause cue, whereas in the case of Wang and Saffran (2014) a congruent suprasegmental tone cue appeared to hinder performance, particularly for monolinguals (relative to prior studies in which learners only tracked similar transitional probabilities between adjacent syllables without additional cues) and the authors suggest that learners may have had to inhibit one cue to follow the other. We note, it is unknown whether the bilingual groups tested in the latter study would have also maintained an advantage had the suprasegmental cues been removed. Our study did not involve a second cue type, but did offer learners the opportunity to form a second mapping for one of the elements. In this situation, bilinguals also appear advantaged relative to functional monolinguals, being better able to overcome a mutual exclusivity bias that has been evidenced with adults in previous cross-situational learning studies using a similar experimental paradigm (e.g., see Yurovsky et al., 2013). Future work in this area will need to determine whether all of these advantages arise as a function of bilinguals possessing a general advantage in implicit learning abilities (e.g., Klein, 1995; Nation & McLaughlin, 1986), or whether they emerge due to other cognitive advantages associated with bilingualism such as improved phonological working memory (e.g., Adesope, Lavin, Thompson, & Ungerleider, 2010; Bartolotti et al., 2011; Majerus et al., 2008; Service, Simola, Metsänheimo, & Maury, 2002) or an advantage in inhibitory control (e.g., Bialystok, 1999; Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernández, & Sebastián-Gallés, 2008; Tao, Marzecová, Taft, Asanowicz, & Wodniecka, 2011; Wang & Saffran, 2014), or some combination thereof. It is possible that in our Experiment 2, the first learned mapping might require suppression (i.e., inhibitory control) in order to acquire the second. Future work will include measures of inhibitory control and working memory that may begin to address this issue. Irrespective, another contribution of the present work is that it is the first demonstration, to the best of our knowledge, that late-learning sequential bilinguals are more open to remapping during accumulative statistical learning relative to functional monolinguals.

As noted in the Introduction, the term statistical learning encompasses many forms of learning, and thus it is quite possible that the different types of statistical learning involve a diverse set of cognitive processes. For example, Hsu and Bishop (2010) suggest that the statistical learning involved in word learning may differ from that of grammar learning and also from nonverbal sequence learning (see Fig. 1 of Hsu & Bishop, 2010). Correspondingly, deficits such as Specific Language Impairment are thought to impact the different types of statistical learning to varying degrees (see Hsu & Bishop, 2010; Ullman & Pierpont, 2005). One attempt to formalize this kind of distinction has been proposed by Thiessen et al. (2013) who differentiate between statistical learning tasks involving extraction and those involving integration. Extraction involves holding two elements in working memory and binding them into a chunk (Perruchet & Vinter, 1998) whereas integration involves combining information across chunks to deduce a central tendency (see Erickson & Thiessen, 2015). These distinctions may provide a useful framework for interpreting the results of statistical learning studies comparing monolinguals and bilinguals. Previous studies exploring statistical learning and bilinguals have all involved sequence-learning, which is best characterized as extraction (Erickson & Thiessen, 2015) whereas the present study involve processes related to integration (since information is stored across trials to deduce the correct associations). Thus, it is possible that the bilingual advantage in forming 2:1 mappings may be evidenced in processes involving integration, but not those involving extraction. For example, data collected by Bogulski (2013) suggest that bilinguals may be prone to the same primacy effect as monolinguals when asked to track sequential regularities across two artificial speech streams presented consecutively without any cues to the change in structure (see Gebhart, Aslin, & Newport, 2009; see also Weiss, Gerfen, & Mitchel, 2009). Arguably, this dual-stream task also involves remapping, since there is a fifty percent overlap in elements used in both languages (and thus learners would have to recognize that the initial set of transitional probabilities across elements no longer apply to the overlapping elements found in the second stream). Thus, it is possible that statistical learning abilities involving extraction are equivalent across populations (unless a second cue is added to the stream, see above) while tasks that require accruing statistics over time may be approached differently by bilinguals, who are more open to the possibility of multiple distributions. In our view, the extraction-integration framework could provide some traction in understanding when differences are likely to emerge between monolinguals and bilinguals in statistical learning tasks and our future work will explore this possibility more directly.

With respect to mutual exclusivity, our findings contribute to an existing literature that has largely studied simultaneous or early sequential bilinguals. For example, several studies have reported that bilingual infants are less likely to adhere to the mutual exclusivity constraint relative to monolinguals (e.g., Byers-Heinlein & Werker, 2009; Davidson & Tell, 2005; Houston-Price et al., 2010). By contrast, the bilingual participants in this study were late L2 learners, and consequently their experience acquiring their L1 was likely equivalent to their monolingual peers with respect to adhering to word-learning constraints. Thus, our findings provide suggestive evidence that this constraint may become relaxed (at least in the context of a statistical learning task) even for sequential bilinguals whose acquisition of the L2 comes later in life. This accords with the notion put forth by Markman and Wachtel (1988) that mutual exclusivity likely persists into adulthood (e.g., Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Halberda, 2006) but weakens with age as learners come to experience more overlap in terms of mappings. For bilinguals, who experience translation equivalents and interlingual homographs as well as synonymy and polysemy within each language, the mutual exclusivity constraint may consequently become significantly weaker relative to monolinguals. This also is consistent with a recent model of word learning that demonstrated decrements in mutual exclusivity when multiple labels were present for a single object (though interestingly not when multiple meanings were tested, see McMurray et al., 2012). Thus, our findings add to a growing literature demonstrating that the cognitive impacts of late sequential bilingualism mirror some of the changes associated with simultaneous or early bilingualism (e.g., Vega-Mendoza, West, Sorace, & Bak, 2015).

We also note the importance of testing multiple populations of bilingual learners to compare with functional monolinguals. In order to be confident that the observed differences arose as a consequence of proficiency with a second language, we needed to decrease the likelihood that our findings could arise as a function of idiosyncrasies associated with a particular language. For example, speakers of Chinese must resolve homophonous relationships frequently as a result of great overlap in Chinese characters and syllables (e.g., Chang, 1993; Kuo et al., 2004; Perfetti & Tan, 1998), and thus it is conceivable that any Chinese speaker (including monolinguals) could expect greater success in forming 2:1 mappings relative to English monolinguals who do not encounter homophony on such a scale. Therefore, the finding that English–Spanish bilinguals perform equivalently to the Chinese–English bilinguals (and also outperform functional monolinguals) in forming 2:1 mappings in Experiment 2 offers very suggestive evidence that the effect is language-independent and arises as a consequence of proficiency with two languages.

Returning to the "learning to learn" framework mentioned in the Introduction, our study provides evidence that experience with multiple languages may fundamentally influence the assumptions made in new learning environments with respect to how many causal models underlie the observed statistics. This notion finds support in developmental studies of early bilingualism (e.g., Kovács & Mehler, 2009), and, as we have shown here, even when proficiency with a second language occurs after the first language has already been mastered. Based on our findings along with previous studies (Bogulski, 2013; Yim & Rudoy, 2013), we argue that the core distributional learning abilities evidenced in transitional probability tasks or cross-situational statistical learning tasks may be unaffected by proficiency with more than one language; just as being exposed to multiple rules is unlikely to impact the most basic principles of learning (e.g., Gallistel et al., 2001). Rather, the differences in performance are more likely to become evident in novel non-stationary environments when learners must determine the number of mappings or structures that underlie the surface input.

Acknowledgments

We wish to thank Dr. Taomei Guo, Haoyun Zhang and Siyao Li in the National Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University and research assistants in the Comparative Communication Lab at the Pennsylvania State University for help with data collection. We would like to thank Federica Bulgarelli for help with recruiting English–Spanish bilinguals. This research was supported by NIH RO1 Grant HD067250 awarded to DJW and NSF PIRE Grant OISE-0968369 to the Center for Language Science at Penn State University.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.03.001.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*, 207–245. http://dx.doi.org/10.3102/0034654310368803.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, 2.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*, 636–644. http://dx.doi.org/10.1111/1467-8624.00046.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychological Aging, 19*, 290–303. http://dx.doi.org/10.1037/0882-7974.19.2.290.

Bogulski, C. (2013). *Are bilinguals better learners? A neurocognitive investigation of the bilingual advantage* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database (UMI No. 1511655851).

Brenders, P., van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology, 109*(4), 383–396.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: infants' language experience influences the development of a word-learning heuristic. *Developmental Science, 12*(5), 815–823.

Chang, C. H. (1993). Corpus-based adaptation mechanisms for Chinese homophone disambiguation. In *Proceedings of the workshop on very large corpora*. USA: Ohio State University.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition, 106*, 59–86. http://dx.doi.org/10.1016/j.cognition.2006.12.013.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2011). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition, 106*(2), 633–664.

Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology, 92*(1), 25–45.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: a pragmatic account. *Developmental Psychology, 37*(5), 630.

Dufort, R. H., Guttman, N., & Kimble, G. A. (1954). One-trial discrimination reversal in the white rat. *Journal of Comparative and Physiological Psychology, 47*(3), 248.

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review, 37*, 66–108.

Fazly, A., Alishahi, A., & Stevenson, S. (2010). A probabilistic computational model of cross-situational word learning. *Cognitive Science, 6*, 1017–1063.

Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate word learning. *Cognitive Science, 35*, 367–380.

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes, 27*(4), 354.

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science, 33*(6), 1087–1116.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28*(1), 99–108.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology, 53*(4), 310–344.

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy, 15*(2), 125–150.

Hsu, H. J., & Bishop, D. V. M. (2010). Grammatical difficulties in children with specific language impairment: Is learning deficient. *Human Development, 53*, 264–277.

Ichinco, D., Frank, M. C., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. In *Proceedings of the 31st annual meeting of the cognitive science society (Vol. 31)*.

Kachergis, G., Yu, C., & Shiffrin, R. (2009). Frequency and contextual diversity effects in cross-situational word learning. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (Vol. 31, pp. 2220–2225). Austin, TX: Cognitive Science Society.

Klein, E. (1995). Second versus third language acquisition: is there a difference?. *Language Learning, 45*, 419–465. http://dx.doi.org/10.1111/j.1467-1770.1995.tb00448.x.

Kovács, Á. M., & Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science, 325*(5940), 611–612.

Krechevsky, I. (1932). *"Hypotheses" versus "chance" in the pre-solution period in sensory discrimination-learning*. University of California Publications in Psychology.

Kuo, W. J., Yeh, T. C., Lee, J. R., Chen, L. F., Lee, P. L., Chen, S. S., & Hsieh, J. C. (2004). Orthographic and phonological processing of Chinese characters: An fMRI study. *Neuroimage, 21*(4), 1721–1731.

Majerus, S., Poncelet, M., van der Linden, M., & Weekes, B. S. (2008). Lexical learning in bilingual adults: the relative importance of short-term memory for serial order and phonological knowledge. *Cognition, 107*, 395–419. http://dx.doi.org/10.1016/j.cognition.2007.10.003.

Marian, V., & Shook, A. (2012). The cognitive benefits of being bilingual. *Cerebrum: The Dana forum on brain science* (Vol. 2012). Dana Foundation.

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. *Perspectives on language and thought: Interrelations in development (pp. 72–106)*.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121–157.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review, 119*(4), 831–877.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences, 108*(22), 9014–9019.

Misyak, J. B. & Christiansen, M. H. (2007). Extending statistical learning farther and further: Long-distance dependencies, and individual differences in statistical learning and language. In: *Proceedings of the 29th annual cognitive science society* (pp. 1307–1312).

Nation, R., & McLaughlin, B. (1986). Novices and experts: An information processing approach to the "good language learner" problem. *Applied Psycholinguistics, 7*(01), 41–55.

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 101.

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language, 39*(246–263).

Poepsel, T., Gerfen, C., & Weiss, D. J. (2012). Context, mutual exclusivity, and the challenge of multiple mappings in word learning. In *Proceedings of the 36th annual Boston conference on language development* (pp. 474–486).

Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in Psychology*, 3.

Quine, W. V. (2013). *Word and object*. MIT press.

Service, E., Simola, M., Metsänheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology, 14*, 383–408. http://dx.doi.org/10.1080/09541440143000140.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558–1568.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in Psychology, 2*, 123. http://dx.doi.org/10.3389/fpsyg.2011.00123.

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*, 792–814.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology, 30*(5), 639.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology, 66*(1), 126–156.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41*(3), 399–433.

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition, 137*, 40–46.

Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical learning. *Cognitive Science, 38*, 757–774. http://dx.doi.org/10.1111/cogs.12092.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition, 107*(2), 729–742.

Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5.

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development, 5*(1), 30–49.

Weiss, D. J., Gerfen, C., & Mitchel, A. (2010). Colliding cues in word segmentation: The role of cue strength and general cognitive processes. *Language and Cognitive Processes, 25*(3), 402–422.

Weiss, D. J., Poepsel, T., & Gerfen, C. (2015). Tracking multiple inputs. *Implicit and Explicit Learning of Languages, 48*, 167.

Williams, J. T. (1968). Reversal-learning in the spectacled caiman. *The American Journal of Psychology, 81*(2), 258–261.

Yim, D., & Rudoy, J. (2013). Implicit statistical learning and language skills in bilingual children. *Journal of Speech, Language, and Hearing Research, 56*(1), 310–322.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420.

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review, 119*(1), 21.

Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review, 21*(1), 1–22.

Yurovsky, D., Yu, C., & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive Science, 37*(5), 891–921.