Understanding Bilingual Word Learning: The Role of Phonotactic Probability and Phonological Neighborhood Density

Vishnu KK Nair[1,2]

Britta Biedermann[2,3]

Lyndsey Nickels[2]

1: Discipline of Speech Pathology and Audiology, Flinders University, Adelaide, Australia

2: ARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University, Sydney, Australia

3: School of Psychology and Speech Pathology, Curtin University, Perth, Australia

Correspondence concerning this article should be addressed to Vishnu KK Nair, Discipline of Speech Pathology and Audiology, Lvl 7 FMC, Flinders University, SA 5001, Australia.

Email: vishnu.nair@flinders.edu.au

Abstract

Purpose: Previous research has shown that the language learning mechanism is affected by bilingualism resulting in a novel word learning advantage for bilingual speakers. However, less is known about the factors that might influence this advantage. This paper reports an investigation of two factors: phonotactic probability and phonological neighborhood density.

Method: Acquisition of fifteen novel words varying in phonotactic probability and phonological neighborhood density was examined in high proficiency, early onset, Mandarin-English bilinguals and English monolinguals.

Results: Both bilinguals and monolinguals demonstrated a significant effect of phonotactic probability and phonological neighborhood density. Novel word learning improved when the phonological neighborhood density was higher, in contrast, higher phonotactic probability resulted in worse learning. Although, the bilingual speakers showed significantly better novel word learning than monolingual speakers, this did not interact with phonotactic probability and phonological neighbourhood density manipulations.

***Conclusion*:** Both bilingual and monolingual word learning abilities are constrained by the same learning mechanisms. However, bilingual advantages may be underpinned by more effective allocation of cognitive resources due to their dual language experience.

Keywords: bilingualism, novel word learning, phonotactic probability, phonological neighborhood density

Understanding Bilingual Word Learning: The Role of Phonotactic Probability and Phonological Neighborhood Density

A large body of research has found converging evidence for a positive relationship between bilingualism and non-linguistic skills (e.g., Abutalebi, Della Rosa, Green, Hernandez, Scifo, Keim, Cappa & Costa, 2011; Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Craik, & Ryan, 2006; Costa, Hernandez, & Sebastián-Gallés, 2008). In contrast, studies examining the impact of bilingualism on linguistic skills have yielded mixed findings. For example, while Sheng, McGregor, and Marian (2006) found a bilingual advantage in children for a lexical production task, Rogers, Lister, Febo, Besing, and Abrams (2006) showed that bilingual adults have more difficulty in recognising words in challenging (noisy) environments compared to monolinguals. It has also been reported that, bilingual adults take longer to retrieve words from the mental lexicon (i.e., slower lexical access) (e.g., Bialystok, 2009; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ivanova & Costa, 2008) and may demonstrate a disadvantage in picture naming tasks (Gollan et al., 2005). Despite these mixed reports, superior bilingual performance has been consistently found for linguistic tasks associated with language learning (e.g., Antoniou, Liang, Ettlinger, & Wong, 2015; Kaushanskaya & Marian, 2008, 2009a, 2009b; Kaushanskaya, 2012; Nair, Biedermann, & Nickels, 2015; Papagno & Vallar, 1995; Van Hell & Mahn, 1997). For instance, Kaushanskaya and Marian (2009b) showed that early bilinguals were significantly better than monolinguals in learning novel words. Nair, Biedermann, and Nickels (2016) found that this bilingual advantage in novel word learning remained present even in a population of late bilinguals with delayed onset of learning their second language (L2 speaking acquired after 12.45 years of age). Although evidence suggests a generally positive influence of bilingualism on word learning, the factors underpinning this bilingual effect remain unclear. Further investigation of the constraints on the effects of bilingualism on the language learning mechanism will enable us to not only further specify how bilinguals differ from monolinguals but also will contribute to the development of theories explaining why bilinguals differ from monolinguals. A variety of participant-related factors have been found to influence the novel word learning advantage in bilinguals, such as age of second language acquisition (Kaushanskaya & Marian, 2008; Nair, Biedermann & Nickels, 2015), length of second language exposure (Hernandez & Li, 2007), cognitive control (Bradley, King, & Hernandez, 2013), and, for developing bilinguals (second language learners), the linguistic structures of their first language (L1; for theoretical accounts see e.g., Flege, 1982, Lado, 1957, and Zobl, 1980). However, more recent investigations with monolingual speakers have shown that phonological and lexical properties of novel words, such as phonotactic probability and phonological neighbourhood density, may also influence word learning outcomes (Storkel, Armbrüster, & Hogan, 2006). However, it is not known how these factors affect word learning in bilinguals. An investigation of whether (and how) these factors affect word learning differently in bilinguals and monolinguals would provide further insights into the changes bilingualism brings to cognitive processing.

Phonotactic probability refers to the likelihood of sounds and sound combinations occurring in a given language (Vitevitch & Luce, 2005). Phonological neighborhood density refers to the total number of words that sound similar to a given word (Luce & Pisoni, 1998). Although these two variables are different, they are related: words with common phoneme sequences (high phonotactic probability) tend to have dense neighborhoods, whereas words with rare sequences generally have sparse neighborhoods (e.g., Vitevitch, Luce, Pisoni, & Auer, 1999). Nevertheless, Storkel et al. (2006) argued that phonotactic probability and phonological neighborhood density have distinct influences on word learning. They examined the independent effects of these two variables on novel word learning in the context of a story in monolingual adults. During the initial stages of word learning (i.e., analysing partially correct responses), they found a disadvantage for words with high phonotactic probability compared to words with lower phonotactic probability. These effects of phonotactic probability found during initial stages of word learning differed from developmental studies where an advantage is usually found with increasing phonotactic probability (e.g., Storkel, 2001; but see also Storkel & Lee, 2011; Hoover, Storkel, & Hogan, 2010 for contrasting evidence). In contrast, the effect of phonological neighborhood density was not significant: Initial learning performance did not vary between words of high and low density.

For the later stages of word learning (i.e., analysing completely correct responses), Storkel et al. (2006) found a contrasting pattern: no significant effect of phonotactic probability but an advantage for words of high neighborhood density. Storkel et al. suggest that the advantage for low phonotactic probability words may be due to these stimuli being less word-like. For instance, these words could be more easily detected, with detection triggering differing processing for novel words (learning) and known words (lexical access).

In contrast, the advantage for words with high neighborhood density during the later stages of word learning was suggested to be due to these words activating more neighbors from long term memory. These neighbors may facilitate acquisition of novel words by strengthening representations through feedback from shared phonemes.

In sum, Storkel et al. (2006) found that monolingual adults demonstrate a disadvantage for learning words with high phonotactic probability but an advantage for words with high phonological neighborhood density (albeit over different phases of learning). Nevertheless, it is clear that the effects of phonotactic probability and phonological neighborhood density on language processing in general are far from straightforward.

It seems that phonological neighborhood density effects on speech production vary depending not only on the task but also across languages and populations studied (e.g., healthy speakers vs individuals with aphasia; see Sadat et al., 2014 for a detailed discussion). By extension, it is possible that monolingual and bilingual adults may differ in the effects of phonological neighbourhood density and phonotactic probability on their word learning.

Moreover, given that bilinguals show advantages for word learning compared to monolinguals, perhaps this advantage is influenced or modulated by the effects of phonotactic probability and phonological neighbourhood density. It is also possible that the phonotactic probability and phonological neighbourhood effects are different in bilinguals compared to monolinguals given that these variables are known to be sensitive to different methodologies (e.g., word recognition, word learning, non-word repetition) and participant characteristics (e.g., language experience, language performance, age). For example, in monolingual adults, high phonotactic probability is generally associated with facilitatory effects for word recognition (e.g., Vitevitch & Luce, 1999) and non-word repetition (e.g., Vitevitch & Luce, 2005). A high phonological neighborhood density may produce an advantage for word production (e.g., Vitevitch, 2002; Baus, Costa, & Carreiras, 2008) and word learning (e.g., Storkel et al., 2006) but not for word recognition (e.g., Luce & Pisoni, 1998).

Bilinguals have extensive experience in learning unfamiliar words in a second language; it is possible that this experience generalises to better learning of novel words regardless of their similarity to other words in their lexicon (i.e., even those words with lower phonotactic probability and lower phonological neighborhood density). In the context of second language learning, two contrasting theories have also been developed specifically to account for learning familiar or unfamiliar phonemes/phoneme strings. For example, the contrastive analysis hypothesis suggests that L2 learning may be susceptible to the structural properties of L1 (Lado, 1957). L2 phonemic structures that closer resemble L1 structures may be easier to learn than L2 structures that are dissimilar to L1. These prior L1 knowledge effects on L2 learning have been proposed to be highly selective depending on the linguistic structure of the languages (Zobl, 1980). However, the speech learning model (Flege, 1982; 1987) predicts that phonemes that are less familiar and unique in a second language will be better learned than commonly occurring phonemes (irrespective of L1).

In a broadly related study, Kaushanskaya, Yoo and van Hecke (2013) found that for English speakers with varying L2 (Spanish) exposure, increased second language experience was associated with enhanced novel word acquisition, but only for novel words that were phonologically unfamiliar (non English/Spanish sounds) paired with familiar semantic referents (animals). This supports the speech learning model (Flege, 1982; 1987) that argues that bilinguals are adept at learning unfamiliar phonological combinations. Kaushanskaya et al. note that, while these are the items that most closely simulate the L2 language learning experience, broader benefits for other items (e.g., familiar phonological words with unfamiliar semantic referents) may be found with more second language exposure. Kaushanskaya et al.'s (2013) findings seem to suggest, therefore, that bilingual experience might be expected to specifically facilitate learning of less frequent sound combinations (words with low phonotactic probability) and/or words that are less similar to words in the lexicon (words with low phonological neighborhood). If a bilingual advantage is demonstrated only for words with low phonotactic probability/low neighborhood density then this would indicate that the cognitive mechanisms that underpin the bilingual advantage are sensitive to these psycholinguistic variables. However, to the best of our knowledge, there have been no attempts to investigate the impact of bilingualism on the effects of phonological neighborhood density or phonotactic probability in word learning or vice versa.

In sum, the critical question from the previous literature is whether the bilingual advantage for novel word learning is specific to only certain phonotactic patterns (e.g., words with low phonotactic probability) and of certain phonological neighborhood density (low neighborhood density) or whether bilinguals exhibit an overall word learning advantage regardless of the phonotactic and neighborhood patterns of the novel word. This question has theoretical implications for the understanding of bilingual word learning, will inform theories of second language learning, and for the specification of the cognitive mechanisms of word learning more generally, and hence is the focus of the research presented here.

Method

Participants

The participants were 20 monolingual native speakers of English (13 females and 7 males) and 20 Mandarin-English early onset, highly proficient, bilingual speakers (11 females and 9 males). The participants were all university undergraduate students and the groups were matched for age (Bilinguals: mean = 21.55 years, standard deviation = 1.00; monolinguals = 21.47 years, standard deviation = 0.94, t (38) = -.245, p = .807).

All bilingual participants rated their second language proficiency across four language categories ranging from 0 (no proficiency) to 4 (native-like). The self-reported proficiency questionnaire captured both language proficiency as well as the linguistic history of the participants and was similar to other second language proficiency measures such as the language experience and proficiency questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007), and the international second language proficiency rating scale (Ingram & Wylie, 1999) and language proficiency categories (Collier, 2007).

The indicator of language proficiency used in this study was age of active bilingualism (indexed through speaking). The mean start age of active bilingualism was 6.15 years (age range = 5-7 years, standard deviation=0.88years) with an average of almost fifteen years of exposure to English (see Table 1). The bilingual participants were native speakers of Mandarin (L1) and had acquired English (L2) in both a classroom context and by immersion in an English-language environment. The language history revealed that the bilingual participants were born to immigrant parents of Chinese background. The participants mean start age of active bilingualism often coincided with their mean age when their parents migrated to Australia. This indicates that although the bilingual participants acquired English in classrooms, the L2 learning happened in a classroom context where English is spoken as the native language. This is a critical indicator for our bilingual participants increased proficiency in L2 in contrast to other bilinguals who learn English as a second language in an impoverished classroom (non-native) context (a scenario that is commonplace for most non-native speakers of English).

These participants reported that they spoke Mandarin at home and social situations especially while communicating with other family members and friends from similar cultural and linguistic backgrounds. English was used to communicate with friends in both formal (e.g., university) and informal (e.g., social) settings. All bilingual participants currently lived and attended university in Australia and met the (stringent) English language requirements for admission. Their proficiency was also reflected in the fact that their English was spoken with a near-native or native accent. The participants did not report any significant language or cognitive impairments.

Before the learning phase, subtests from Comprehensive Test of Phonological Processing (CTOPP, Wagner, Torgesen, & Rashotte, 1999) were used to test participants' non-word repetition and digit span abilities. Participant's demographic characteristics and self-ratings of bilingual language proficiency are reported in Table 1. The participant groups did not differ in their non-word repetition (t (38) = .493, p = .625) or digit span scores (t (38) =.-220, p = .827).

Table 1

Demographic and background data of participants. Means and standard deviations (in parentheses).

Demographic variables	**Monolinguals**	**Bilinguals**	**p value**
Age (years)	21.47 (0.938)	21.55 (0.998)	.807
Non-word repetition^{a}	70.65 (6.70)	69.55 (7.39)	.625
Digit span^{b}	68.10 (9.91)	69.15 (10.17)	.746

L2 acquisition age (speaking)	___	6.15 (0.812)	___
Proficiency ratings[c]			
Speaking	___	3.05 (0.394)	___
Listening	___	3.40 (0.502)	___
Reading	___	3.35 (0.489)	___
Writing	___	3.28 (0.487)	___

Notes
N=20 for both participant groups
[a & b] Non-word repetition (n=18) and digit span (n=21) percentile scores (subtests of Comprehensive Test of Phonological Processing)
p value = significance of two-sample t-test (two tailed)
[c] Proficiency Ratings from 0 = not proficient to 4 = highly proficient

Stimuli

To examine the effects of phonotactic probability and phonological neighbourhood density, we first created 35 bisyllabic non-words with varying English phonotatctic probability and phonological neighborhood density as calculated using the English vocabulary of the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) using an algorithm based on the conventions used by Storkel et al. (2006). The non-words were created by first selecting real words that contained either high or low bigram and trigram frequencies (bi-/tri-gram frequencies refer to the number of occurrences of a particular two or three letter string across all the words in a language), or high or low phonological neighborhood density[1]. We changed these real words into a non-word by deleting or changing the position of a single phoneme. This generated a list of bisyllabic non-words with varying bigram and trigram frequencies, whose phonological neighborhood size, biphone frequency, positional segment frequency and summed biphone frequency were once again extracted from CELEX.

---

[1] In our calculation of bigram and trigram frequencies, and phonological neighborhood density we excluded any item for which the headword was a simple contraction, complex contraction, letter or abbreviation, any item with a spelling containing a non-alphabetic character (e.g., hyphen, space) or a capital letter in a position other than the first. This left 65030 unique pronunciations in the database. In cases where a pronunciation occurred in multiple entries the frequencies were summed to get a single total frequency for that specific pronunciation.

We then constructed matched sets that were high or low in each variable (i.e, higher or lower than the median value of the set). We originally had hoped to have orthogonal contrasts (i.e., four sets that were high in phonotactic probability and either high or low in phonological neighbourhood density, or low in phonotactic probability and either high or low in phonological neighbourhood density). However it was not possible to create a set of non-words that were both low in phonotactic probability and high in phonological neighbourhood density. Therefore the final stimuli consisted of three sets of five words that were: 1) low in phonotactic probability and phonological neighborhood density (LP-LD); 2) high in phonotactic probability and low in phonological neighborhood density (HP-LD); and 3) high in phonotactic probability and high in phonological neighborhood density (HP-HD). Although these three categories enabled us to fulfil our aims, they did not allow for examination of the full range of one variable while manipulating the other. Therefore, we examined phonotactic probability and phonological neighbourhood density effects in three ways: a) effects of phonotactic probability were examined when phonological neighborhood density was low; b) effects of phonological neighborhood density were examined when phonotactic probability was high c) correlated effects of phonotactic probability and phonological neighborhood density were examined by comparing the low phonotactic probability/ low phonological neighborhood density set to the high phonotactic probability/ high phonological neighborhood density set. This final comparison allowed us to look at what might be considered the 'norm' in language learning: when stimuli are high in both variables or low in both variables.

The similarity of the target words to Mandarin was rated by five native speakers of Mandarin based on a 1-4 point rating scale (1 = no resemblance to Mandarin, 4 = close resemblance to Mandarin). The ratings indicated that none of the novel words resembled Mandarin words. While carrying the familiarity ratings for Mandarin we asked the judges to identify if the stimuli contained any sounds that were unusual for Mandarin. The judges were not able to identify any such sounds. This was not completed for English, because the non-words were constructed using English phonemes (see Appendix).

Referents

Each novel word was paired with a novel picture as a referent. The pictures consisted of color images of fifteen novel alien creatures differing in physical appearance and characteristics selected from the Gupta et al. (2004) stimulus set in such a way that all fifteen were visually distinct. Each alien was also assigned with attributes (definition) relating to physical or mental characteristics of the alien (e.g., “/tæbɛk/ likes flowers and owns a beautiful garden”) and unrelated to the physical appearance. The novel words and their definitions are given in the Appendix.

Procedure

All 15 novel words were presented for learning in one session. The word learning session followed background testing and completion of the language proficiency questionnaire. Each learning phase consisted of presentation of the referent picture on a Mac OS X 10.7 laptop monitor together with simultaneous presentation of an audio recording of the novel word and its definition (to provide additional semantic/ associative elaboration). For example, “this is *fɒni:s*, *fɒni:s* can sing beautifully and is known as the heavenly singer". Pictures remained on the screen for 30 seconds, and were followed by the next stimulus.

Following presentation of all stimuli, they were presented again in a different random order four more times. In the first four presentations, participants were told to look at the picture and listen to, and memorise the name of the picture and the sentence about its characteristics. After the final (fifth) presentation of each item, while the picture of the item remained on the screen, the participants were asked to repeat the word aloud three times to maximize the learning.

We carried out two phases of testing, one test immediately following the five presentations for word learning and a second test, one week later. Each participant was assessed on the acquisition of the novel words using a picture naming task. The picture naming task aimed to evaluate learning of the word form and its association with the picture referent, and also allowed for examination of the effects of phonotactic probability and phonological neighborhood density on learning. In this task the target picture was presented and the participant was asked to name it as quickly as possible. The responses were audio recorded for later analysis.

Analyses

Analyses were performed on response accuracy (raw number of correct responses out of five) for each task. Object naming was considered correct if all phonemes were produced correctly. This was scored on the basis of review of the audio recordings by the first author, on two separate occasions to ensure reliability. On three instances when the response was unclear, an independent researcher reviewed the recording. The total correct responses for all stimuli subsets were calculated.

We analysed accuracy for naming using a series of mixed analyses of variance (ANOVAs). These ANOVAs comprised one between subject factor, language group (bilingual/monolingual) and two within subject factors: testing time and condition (high/low in the variable under consideration). The first analysis examined phonotactic probability (high vs low), the second phonological neighbourhood density (high vs low) and the final analysis the correlated comparisons (high in both PP and PND, vs low in both PP and PND).

Results

Figure 1 provides the results of the naming task and Table 2 gives the results of the statistical analyses. For clarity, we first summarise the effects of group and time across all analyses and then report the results of the analyses manipulating phonotactic probability and phonological neighborhood density.

Figure 1. Effects of phonotactic probability (High = stimuli with high phonotactic probability, Low = stimuli with low phonotactic probability)

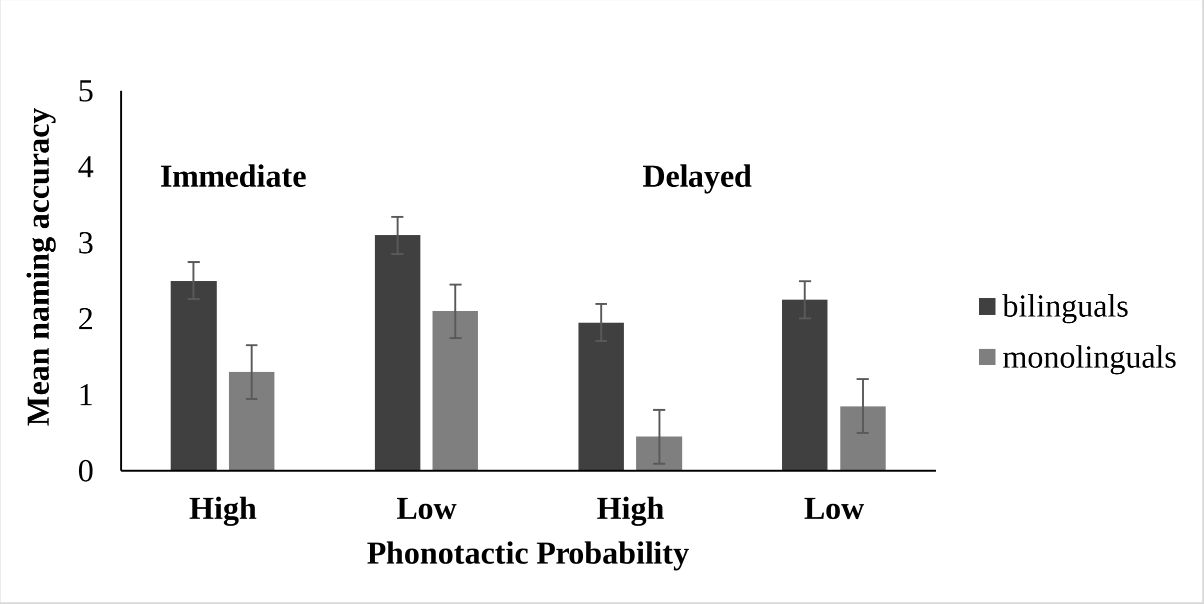

Figure 2: Effects of Phonological neighborhood density (High = stimuli with high phonological neighbourhood density, Low = stimuli with low phonological neighbourhood density)

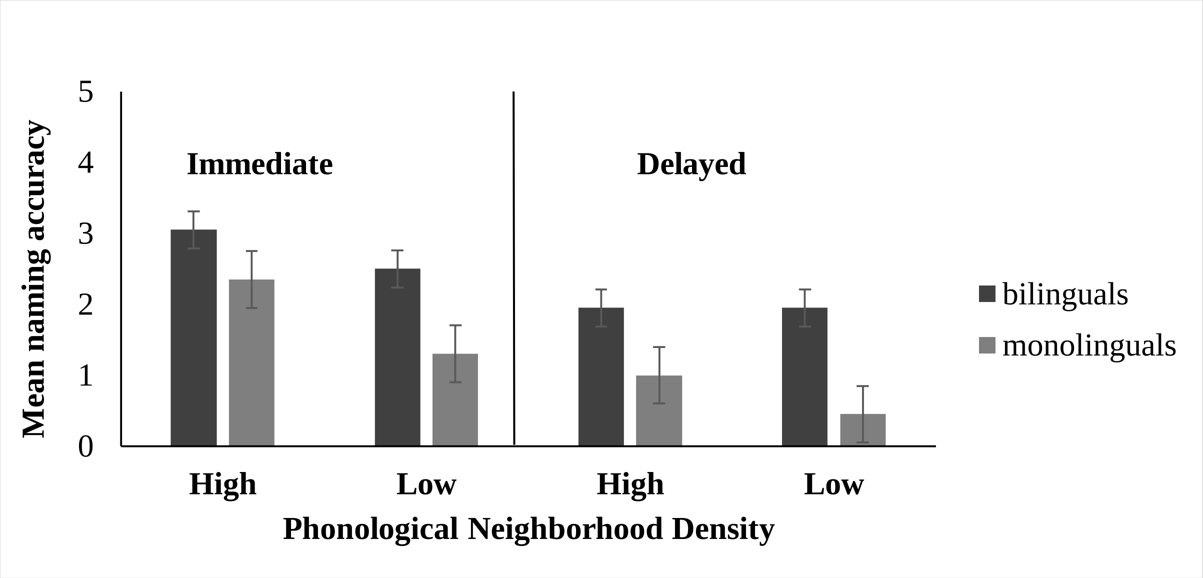

Figure 3: Correlated comparisons (High = words with high phonotactic probability and phonological neighbourhood density, Low = words with low phonotactic probability and phonological neighbourhood density)

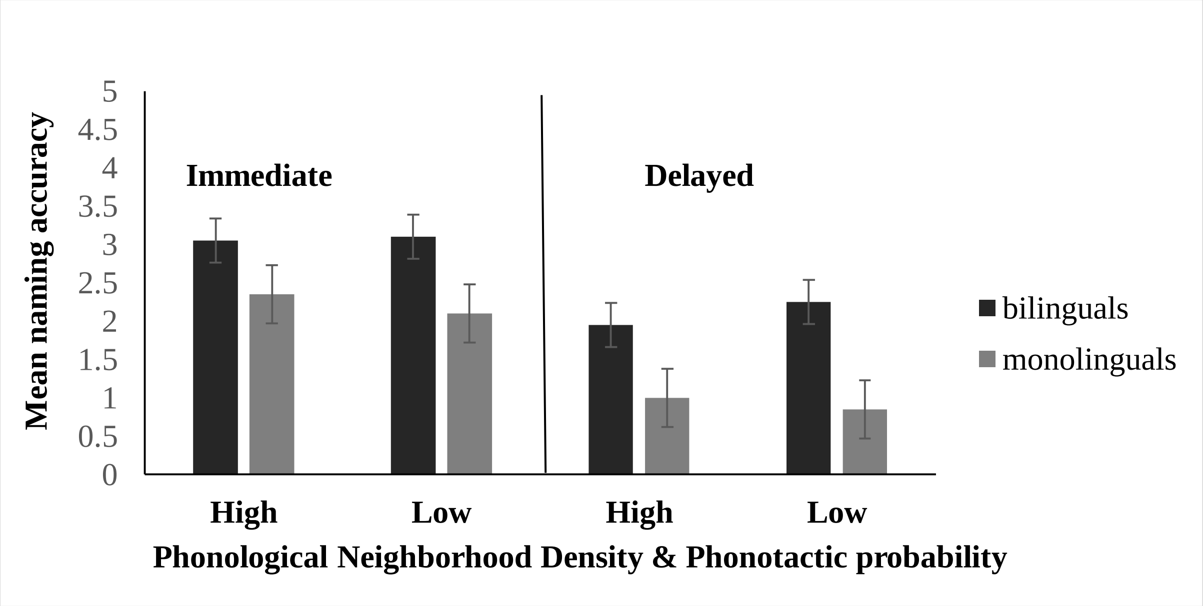

Table 2

Results of phonotactic probability, phonological neighborhood density and correlated comparisons.

Effect	**Degrees of freedom**	**F**	**p**	**n^{2}p**
Analysis 1: Phonotactic Probability				
Group (monolingual vs bilingual)	1,38	15.70	<.001**	.292
Time (Immediate vs Delayed)	1,38	35.10	<.001**	.480
PP (Hi vs Low)	1,38	18.41	<.001**	.326
Time*Group	1,38	1.40	.243	___
Time*PP	1,38	1.96	.170	___
Group*PP	1,38	.376	.543	___
Time*Group*PP	1,38	.040	.843	___
Analysis 2: Phonological Neighbourhood Density				
Group (monolingual vs bilingual)	1,38	12.51	.001**	.248
Time (Immediate vs Delayed)	1,38	43.41	<.001**	.533
PND (Hi vs Low)	1,38	10.72	.002**	.220
Time*Group	1,38	.886	.353	___

Time*PND	1,38	3.22	.080	___
Group*PND	1,38	2.55	.118	___
Time*Group*PND	1,38	.007	.932	___
Analysis 3:Correlated Comparisons				
Group (monolingual vs bilingual)	1,38	11.31	.002**	.229
Time (Immediate vs Delayed)	1,38	85.74	<.001**	.693
Condition (HiPP/PND vs LowPP/PND)	1,38	.005	.942	___
Time*Group	1,38	1.75	.194	___
Time* Condition	1,38	.498	.485	___
Group*Condition	1,38	1.20	.279	___
Time*Group*Condition	1,38	.091	.764	___

Notes
PP=Phonotactic Probability, PND=Phonological Neighborhood Density, Hi=High, HiPP/PND=High Phonotactic Probability/Phonological Neighborhood Density
**p < .01, η^2_p (effect size) = Partial eta square

Effects of Group and Time

All three analyses showed a significant main effect of group indicating better naming for bilinguals compared to monolinguals. There was also, as might be expected, a significant main effect of time: participants performed better immediately compared to at one week delay. The interactions between group and time were not significant.

Analysis 1: Manipulating Phonotactic probability

This analysis examined the effects of phonotactic probability when the sets were matched on phonological neighborhood density (both low in phonological neighborhood density). There was a significant main effect of phonotactic probability with lower probability words better named than higher probability words (see Figure 1, Table 2). There were no significant two-or three-way interactions with group and time. This showed that the bilingual advantage in learning, as indexed by the naming task, did not vary according to phonotactic probability.

Analysis 2: Manipulating Phonological neighborhood density

This analysis examined the effects of phonological neighborhood density when sets were matched for phonotactic probability (both high in phonotactic probability) and found a main effect of phonological neighborhood density with higher accuracy for high density words than low density words, but no interactions with group or time (see Figure 2, Table 2). Once again this indicated that the bilingual advantage did not differ across high and low phonological neighborhood density novel word learning.

Analysis 3: Correlated Comparisons

This analysis examined the effects of phonological properties of stimuli on naming comparing sets that were high in both phonotactic probability and phonological neighborhood density with sets that were low in both variables. This analysis found no effect of this manipulation on naming, nor any significant interactions (Figure 3, Table 2).

Discussion

This study aimed not only to replicate the bilingual advantage that has been found for novel word learning (Kaushanskaya & Marian, 2009b), but also, more importantly, examine whether this advantage was influenced by the phonotactic probability and phonological neighborhood density of novel words. In order to investigate this, we compared the performance of high proficiency early bilinguals and matched monolinguals on a word learning task that manipulated the phonotactic probability and phonological neighborhood density of the novel words. We found that bilinguals outperformed monolinguals in learning as measured by a picturenaming task. There were also clear effects of phonotactic probability and phonological neighborhood density on word learning. However, there was no difference between bilinguals and monolinguals in the extent to which these variables influenced learning.

Interestingly, there was no difference in learning between stimuli that were high in both phonotactic probability and phonological neighborhood density and low in both. This result is consistent with recent studies with monolinguals that have reported non-significant effects for correlated comparisons with phonotactic probability and phonological neighbourhood density (e.g., Storkel & Lee, 2011). Our results provide initial evidence for non-significant effects for correlated comparisons in bilinguals too. However, while it was not possible to develop stimuli for a full factorial design, when we examined the effects of these two variables independently, we found that the two factors had opposite effects on learning.

When manipulating phonotactic probability (in stimuli that were low in phonological neighborhood density), both bilinguals and monolinguals demonstrated better accuracy for learning words of lower phonotactic probability. This effect is consistent with that found by Storkel et al. (2006), although they only found an effect of phonotactic probability for the initial stages of learning, arguing that learning was triggered better when the stimuli were more novel. However, this explanation seems less appropriate to a context, such as that used here, where the words were explicitly flagged as novel (and do not need to be detected in a story context). Perhaps, in a direct learning task, such as that used here, learning abilities are at a peak when the novelty associated with the task increases. In the context of our task, this would occur specifically for low phonotactic probability items because of their unfamiliarity (and therefore novelty) compared to high phonotactic probability items. Consequently, this may result in a low probability advantage in low neighborhood density words for both bilinguals and monolinguals.

For the phonological neighborhood density manipulation (within words of high phonotactic probability), words of higher neighborhood density were better learned. Our results therefore replicate Storkel et al.'s (2006) findings on monolingual adult word learning in a story context. Storkel et al. suggested that the advantage for high neighborhood density words could be due to better consolidation of representations through the links with many neighbors. They suggest that the activated neighbors of the novel word would activate their phonemes, and that these phonemes may provide feedback that facilitates the acquisition of the novel word. This interactive process results in strengthening of the representation for the novel word. This could also hold for our task. Hence, while low phonotactic probability may enhance initial learning, higher phonological neighbourhood could enhance consolidation of that learning.

An advantage for words with low phonotactic probability also may speak to the speech learning model (Flege, 1987). This model suggests that the structural similarities between L1 and L2 may have little effect on L2 learning. Flege (1987) argues that second language learners are highly capable of learning new phonemes in L2 without accessing prior knowledge from L1. When they learn a new phoneme in L2, they are also capable of independently modifying the previously learned similar phonemic patterns in L2 without relating it to L1 structures. Although the speech learning model (Flege, 1987) was developed in the context of second language learning, we suggest that the model in fact may logically suggest that any speaker should be more adept at learning words containing less familiar phonemes. Critically, however, all of our phonemes were of high familiarity to all our speakers, but what varied was the familiarity of the combinations of these phonemes. It is possible, therefore, that not only less familiar phonemes but also less familiar combinations of phonemes are better learned.

Overall, bilinguals performed more accurately in learning, as indexed by naming, irrespective of the phonotactic probability and phonological neighborhood density of the stimuli. This finding replicates the previous research demonstrating facilitatory effects of bilingualism for novel word learning (e.g., Kaushanskaya & Marian 2009a, 2009b; Nair et al., 2015). It is also possible that early experience with a second language could generally facilitate the language learning mechanism (e.g., Bartlotti & Marian, 2012; Bartlotti, Marian, Schroeder, & Shook, 2011; Grey, 2013; Van Hell & Mahn, 1997, Wang & Saffran, 2014; Yoshida, Tran, Bentitez, & Kuwabara, 2011). For example, the phonological system of bilingual participants may have influenced their word learning skills: Kaushanskaya and Marian (2009b) have argued that experience with more than one language makes the bilingual’s phonological system comparatively more open. However, this might be thought to imply that the bilingual's phonological system may therefore be more open to accepting any phonological combination (even unusual combinations), in contrast to the specific phonological tuning that occurs for monolinguals in their native language (e.g., Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992).

How might the effects of phonotactic probability and phonological neighborhood density on bilingual word learning relate to the language production mechanism of a bilingual speaker? This is particularly critical given that we did not find an interaction between either phonotactic probability or phonological neighborhood density and the learning of the different participant groups. This suggests that the bilingual advantage in novel word learning is at least partially rooted in factors other than phonological or lexical properties of the novel words. In other words, it is likely that the bilingual advantage transcends the potential effects of phonotactic probability and phonological neighbourhood density.

In the ‘Inhibitory Control’ model for language production (Green, 1998), the language production mechanism of bilinguals is mediated by the lexico-semantic system, the language task schema and the supervisory attentional system. It is the language task schema that helps in selecting the appropriate language and inhibiting the non-relevant language. In the context of word learning, the most crucial component of the inhibitory control model is the supervisory attentional system. Green (1998) suggested that the supervisory attentional system is a goal oriented mechanism which is especially skilled at facilitating tasks that have been not previously performed. Therefore, when an individual performs a novel task associated with language production, such as novel word learning, the supervisory attentional system is employed to ensure its successful completion.

While it is likely that components of the ‘Inhibitory Control’ model are present in both bilinguals and monolinguals, the supervisory attentional system of bilinguals is argued to be more efficient and more active than in monolinguals (see Bialystok, Craik, & Luk, 2012 for a detailed discussion), and our word learning data appear consistent with this account. We propose that the supervisory attentional system recognises learning of any novel word as a novel task and allocates all available attentional resources to execute the successful completion of the task. This leads to more attentional resources being available in the bilingual speaker than the monolingual, resulting in enhanced learning. Therefore we hypothesise that the bilingual advantage in novel word learning could be due to enhancement of this mechanism that underpins word learning. Moreover, our data show that this mechanism is not sensitive to effects of phonotactic probability or phonological neighborhood density.

There are, of course, limitations related to the current study. First, we did not explicitly manipulate or control phonotactic probability and phonological neighborhood density in Mandarin as well as English. Consequently, we cannot be sure how far for the bilingual’s L1

(Mandarin) affected these variables. For example, it is possible (if unlikely) that some stimuli may have been of higher phonotactic probability and/or phonological neighborhood density for the bilinguals. Future research should examine this potential confound from L1. Second, it would have been preferable to completely orthogonally manipulate phonological neighborhood density and phonotactic probability, however, at least within our stimuli this was not possible. Consequently we could only manipulate phonotactic probability within words with low phonological neighborhood density, and phonological neighborhood density within words with high phonotactic probability. The lack of a full orthogonal manipulation restricted our ability to examine, for example, whether there was an interaction between the two effects.

Conclusions

The present study replicates and extends the findings regarding the effects of phonotactic probability and phonological neighborhood density on learning. Like Storkel et al. (2006) we demonstrate that despite the high correlation between these two variables their effects not only can be dissociated but are in different directions – an inhibitory effect of phonotactic probability and a facilitatory effect of phonological neighborhood on word learning. Moreover we also replicate previous findings of a bilingual advantage in novel word learning and provide two important contributions to the literature on the linguistic effects of bilingualism. First, we demonstrate that the facilitatory effects of bilingualism on novel word learning are stable even when the phonotactic probability and neighborhood density of the novel words varies; and second we propose that the loci of these advantages may be an efficient Supervisory Attentional System. These results have theoretical implications for understanding the effect of cognitive mechanisms on bilingual novel word learning as well as potential future clinical implications.

Acknowledgements

During the preparation of this paper, Vishnu KK Nair was supported by an International Macquarie University Research Excellence Scholarship (iMQRES - 2011105), ARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University. Lyndsey Nickels was supported by an Australian Research Council Future Fellowship (FT 120100102), and Britta Biedermann by an Australian Research Council Postdoctoral Fellowship and Discovery Project (DP 11010079). We would like to thank Katherine Demuth for her suggestions in the planning stages of this research and on an earlier version of the manuscript, Steven Saunders and Lois MacCullagh for their assistance in stimulus preparation and recording.

References

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S., & Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, *22*, 2076-2086. doi: 10.1093/cercor/bhr287

Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, *18*(04), 683-695.doi: http://dx.doi.org/10.1017/S1366728914000777

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2): Linguistic Data Consortium. *University of Pennsylvania, Philadelphia, PA, USA*.

Bartolotti, J., & Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cognitive Science, 36*(6), 1129-1147. doi:10.1111/j.1551-6709.2012.01243.x

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*(324), 1-9. doi: 10.3389/fpsyg.2011.00324

Baus, C., Costa, A., & Carreiras, M. (2008). Neighbourhood density and frequency effects in speech production: A case for interactivity. *Language and Cognitive Processes*, *23*(6), 866-888. doi:10.1080/01690960801962372

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*(2), 290. doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, *16*(4), 240-250. doi: 10.1016/j.tics.2012.03.001

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1341. doi.org/10.1037/0278-7393.32.6.1341

Bradley, K. A., King, K. E., & Hernandez, A. E. (2013). Language experience differentiates prefrontal and subcortical activation of the cognitive control network in novel word learning. *NeuroImage*, *67*, 101-110. doi: 10.1016/j.neuroimage.2012.11.018.

Collier, C. (2007). Cognitive learning strategies. Leveled learning strategies. Retrieved from: http://crosscultured.com/documents/Cognitive,%20Learning%20Strategies/Leveled%20Language%20Strategies.pdf.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59-86. doi:10.1016/j.cognition.2006.12.013

Fabiano-Smith, L., & Goldstein, B. A. (2010). Phonological acquisition in bilingual Spanish–English speaking children. *Journal of Speech, Language, and Hearing Research*, *53*(1), 160-178. doi:10.1044/1092-4388(2009/07-0064)

Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65.

Flege, J. E., & Hammond, R. M. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, *5*(01), 1-17. doi: org/10.1017/S0272263100004563

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220-1234. doi: 10.3758/BF03193224

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(02), 67-81. doi:10.1017/S1366728998000133

Gupta, P., Lipinski, J., Abbs, B., Lin, P. H., Aktunc, E., Ludden, D., Martin, N & Newman, R. (2004). Space aliens and nonwords: Stimuli for investigating the learning of novel word-meaning pairs. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 599-603. doi: 10.3758/BF03206540

Grey, S. E. (2013). *A neurocognitive investigation of bilingual advantages at additional language learning*. Georgetown University.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: its neural and computational mechanisms. *Psychological Bulletin*, *133*(4), 638. doi: 10.1037/0033-2909.133.4.638. 638

Hoover, J. R., Storkel, H. L., & Hogan, T. P. (2010). A cross-sectional comparison of the effects of phonotactic probability and neighborhood density on word learning by preschool children. *Journal of Memory and Language*, *63*(1), 100-116. doi.org/10.1016/j.jml.2010.02.003

Ingram, D. and Wylie, E. (1999) International Second Language Proficiency Ratings.Brisbane:Centre for Applied Linguistics and Languages, Griffith University.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*(2), 277-288. doi:10.1016/j.actpsy.2007.06.003

Kaushanskaya, M., & Marian, V. (2008). Age-of-acquisition effects in the development of a bilingual advantage for word learning. In *Proceedings of the 32nd Annual Boston University Conference on Language Development* (pp. 213-224).

Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*, *15*(03), 470-489. doi: http://dx.doi.org/10.1017/S1366728911000472

Kaushanskaya, M., & Marian, V. (2009a). Bilingualism reduces native-language interference during novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 829. doi: 10.1037/a0015275

Kaushanskaya, M., & Marian, V. (2009b). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705-710. doi*:* 10.3758/PBR.16.4.705

Kaushanskaya, M., Yoo, J., & Van Hecke, S. (2013). Word learning in adults with second-language experience: Effects of phonological and referent familiarity. *Journal of Speech, Language, and Hearing Research*, *56*(2), 667-678. doi:10.1044/1092-4388(2012/11-0084)

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606-608. doi: 10.1126/science.1736364

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967. doi:10.1044/1092-4388(2007/067)

Nair, V. K., Biedermann, B., & Nickels, L. (2016). Consequences of late bilingualism for novel word learning: Evidence from Tamil–English bilingual speakers. *International Journal of Bilingualism, 20*(4), 473-487. doi:10.1177/1367006914567005

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *The Quarterly Journal of Experimental Psychology*, *48*(1), 98-107. doi.org/10.1080/14640749508401378

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*(03), 465-485. doi.org/10.1017/S014271640606036X

Sadat, J., Martin, C. D., Costa, A., & Alario, F. (2014). Reconciling phonological neighbourhood effects in speech production through single trial analysis. *Cognitive Psychology*, *68*, 33-58. doi:10.1016/j.cogpsych.2013.10.001

Sheng, L., McGregor, K. K., & Marian, V. (2006). Lexical–semantic organization in bilingual children: Evidence from a repeated word association task. *Journal of Speech, Language, and Hearing Research*, *49*(3), 572-587. doi:10.1044/1092-4388(2006/041

Storkel, H. L. (2001). Learning New Words. Phonotactic Probability in Language Development. *Journal of Speech, Language, and Hearing Research*, *44*(6), 1321-1337. doi:10.1044/1092-4388(2001/103)

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1175-1192. doi: 10.1044/1092-4388(2006/085)

Storkel, H. L., & Lee, S. Y. (2011). The independent effects of phonotactic probability and neighbourhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, *26*(2), 191-211. doi.org/10.1080/01690961003787609

Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, *47*(3), 507-546. doi: 10.1111/0023-8333.00018

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1), 306-311.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374-408.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*(2), 193-204. doi: 10.1006/brln.1999.2116

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 735. doi:10.1037//0278-7393.28.4.735

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing: CTOPP*. Austin, TX: Pro-ed.

Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, *5:953*. doi:10.3389/fpsyg.2014.00953

Yoshida, H., Tran, D. N., Benitez, V., & Kuwabara, M. (2011). Inhibition and adjective learning in bilingual and monolingual children. *Frontiers in Psychology*, *2*, 210. doi:10.3389/fpsyg.2011.00210

Zobl, H. (1980). The formal and developmental selectivity of LI influence on L2 acquisition. *Language learning*, *30*(1), 43-57. doi: 10.1111/j.1467-1770.1980.tb00150.x

Appendix

Stimuli characteristics, median, mean, standard deviation, definitions and visual referent code number

Novel words	Word category	PP	PND	PN	PSF	BF	MSR	Definitions	Code Number
fɒni:s	HP-HD	-	High	4	0.34	0.02	1	*fɒni:s* can sing beautifully and is known as the heavenly singer.	Set 3 A12-C.25
pɪkɪn	HP-HD	-	High	7	0.45	0.02	1	*pɪkɪn* lives on Mars and owns a big crystal house.	Set 2 A12-C.25
rɛdɪn	HP-HD	-	High	4	0.36	0.03	1	*rɛdɪn* can turn stones into diamonds.	Set 2 A14-C.25
mi:lit	HP-HD	-	High	4	0.36	0.03	1	*mi:lit* creates water and rain in the sky.	Set 1 G04-C.25
dɪtaɪz	HP-HD	-	High	7	0.32	0.04	1	*dɪtaɪz* can create thunder and lightning from his eyes.	Set 1 G02-C.25
mɪgæk	LP-LD	Low	-	0	0.22	0.01	1	*mɪgæk* owns a powerful elephant which has seven heads.	Set 1 F02-C.25
lɛvrəʊ	LP-LD	Low	-	0	0.17	0	1	*lɛvrəʊ* enjoys the beauty of the shining stars.	Set 1 A03-C.25
mi:ɒp	LP-LD	Low	-	0	0.13	0	2	*mi:ɒp* is interested in paintings and fine arts.	Set 2 C16-C.25
trɒgɛm	LP-LD	Low	-	0	0.2	0.02	1	*trɒgɛm* travels to earth in a carriage pulled by five horses.	Set 1 A01-C.25
tæbɛk	LP-LD	Low	-	0	0.18	0.01	1	*tæbɛk* likes flowers and owns a beautiful garden.	Set 1 K01-C.25
tɪsɪv	HP-LD	High	Low	1	0.4	0.03	1	*tɪsɪv* is the eldest alien and the head of the alien family.	Set 2 C10-C.25
dɪmtɛz	HP-LD	High	Low	0	0.4	0.04	1	*dɪmtɛz* enjoys chocolate, milk and sweets very much.	Set 2 A11-C.25
tʃɒnid	HP-LD	High	Low	0	0.33	0.02	1	*tʃɒnid* is very knowledgeable and is regarded as an experienced teacher.	Set 3 D02-C.25
sɛna:k	HP-LD	High	Low	0	0.31	0.02	1	*sɛna:k* is very good at healing diseases.	Set 3 B12-C.25

sɪsrɛt	HP-LD	High	Low	0	0.46	0.03	1	*sɪsrɛt* is fond of travelling and driving around space.	Set 1 D02-C.25
	HP-HD (Mean)			5.2	0.37	0.03			
	LP-LD (Mean)			0	0.18	0.01			
	HP-LD (Mean)			0.2	0.38	0.03			

Notes
HP-HD = High phonotactic probability-High phonological neighborhood density
LP-LD = Low phonotactic probability-Low phonological neighborhood density
HP-LD = High phonotactic probability-Low phonological neighborhood density
PP = Items included in the Phonotactic probability manipulation
PND = Items included in the Phonological Neighborhood Density manipulation
PN = Phonological Neighborhood (minimum =0, maximum = 28, median =1, M = 2.86, SD = 3.81)
PSF = Positional Segment Frequency
Summed Positional Segment Frequency (minimum = 0.002, maximum = 0.805, median = 0.303, M = 0.302, SD = 0.014)
BF=Biphone Frequency
Summed Biphone Frequency (minimum =0, maximum =0.137, median = 0.023, M = 0.025, SD = 0.017)
MSR = Mandarin Similarity Rating for novel word (1 = no resemblance to Mandarin, 4 = close resemblance to Mandarin)
Mean in parenthesis indicates mean values for PN, PSF and BF for all three novel word categories
Code number indicates the specific visual referent (alien) number corresponding to each item (all color images) (Gupta et al. 2004)
Link to the visual referents database http://link.springer.com/article/10.3758/BF03206540