Monolingual and bilingual adults can successfully learn foreign language words implicitly

Hana Zjakic[1,2], Alba Tuninetti[1,2], Paola Escudero[1,2]

[1]MARCS Institute for Brain, Behaviour, & Development, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia
[2]Centre of Excellence for the Dynamics of Language, Australian Research Council, Canberra, ACT, Australia

H.Zjakic@westernsydney.edu.au, A.Tuninetti@westernsydney.edu.au,
Paola.Escudero@westernsydney.edu.au

Abstract

Assessing word learning methods has important implications for classroom settings and developing language learning programs. Previous research has shown that adults can learn native-language words implicitly [1] and second-language (L2) words explicitly [2]. We tested whether adults could learn L2 words and attend to phonetic detail implicitly using cross-situational word learning, where participants make word-object associations by tracking word-object co-occurrences across learning trials. Results show that participants learned L2 words implicitly with above-chance accuracy, and that the amount of phonetic detail needed to distinguish words determined their accuracy levels. We discuss how these results compare to explicit L2 word learning.

Index Terms: second language learning, word learning, implicit

1. Introduction

Learning words is a challenging process as learners will often be presented with a number of potential referents for a single linguistic label at once [1]. Word learning can occur in two ways: implicit word learning happens without the learner's awareness and learners' access of implicit knowledge is automatic [3]. Alternatively, explicit word learning is a more conscious process and the acquired knowledge can only be accessed when time constraints are not an issue and the learner's focus in on accuracy [4]. Importantly, implicit and explicit word learning are not mutually exclusive as both play a vital role in language learning. However, learners may prefer one method over the other depending on factors such as age, language exposure and experience [5] or instructional context, whether it be formal (e.g. in a classroom) or naturalistic (e.g. in an individual's home) [6].

Previous studies have shown that the acquisition of knowledge becomes more explicit over the course of an individual's lifespan [7]. According to the Maturational Hypothesis [5], children are maturationally constrained to rely on implicit learning mechanisms, while adults rely primarily on explicit methods of learning. For example, Escudero, Broersma and Simon [8] tested native Dutch listeners on their ability to learn novel Dutch words using an explicit word learning task. Participants completed a learning phase in which each trial consisted of a novel image being presented on a screen along with the audio for the corresponding novel word in the form of *'This is an X'*. The same word was then repeated *('Click on the X')*, along with two images appearing on the screen (i.e. the target object and a distractor object) and participants were instructed to select the correct object. During learning, participants were familiarised with all 12 novel words and their visual referents and were then tested on their ability to match the words and objects during a test phase. Results indicated that participants had higher accuracy for the non-minimal pairs (i.e. sets of words that differ in two or more segments: nonMPs) than minimal pairs (i.e. sets of words that differ in only one segment: MPs).

Using a similar task, Elvin, Escudero, Williams, Shaw and Best [2] tested 20 native Australian-English (AusE) monolingual adults on their ability to learn novel Brazilian Portuguese (BP) words. Participants were more accurate at learning nonMPs than MPs. Accuracy rates were also higher for easy-MPs than difficult-MPs, which were categorized according to how difficult the vowel contrasts were to discriminate. The lower accuracy scores for the difficult-MPs provide support for models of speech perception such as the Perceptual Assimilation Model (PAM) [9] and Second-Language Linguistic Perception Model (L2LP) [10, 11, 12] which state that the perception of L2 sounds is influenced by the acoustic properties of an individual's native language (L1). According to both models, when L2 listeners perceive two non-native sounds as one native category (known as single-category assimilation in PAM and NEW scenario in L2LP), they will face the difficult task of either creating a new L2 category for one of the non-native sounds or splitting an existing L1 category. Mapping two non-native sounds on to two separate L1 categories (known as a two-category assimilation in PAM and SIMILAR scenario in L2LP) poses another learning challenge for L2 learners as they must adjust the boundaries of their existing L1 categories to match those of the L2.

Interestingly, previous research has also shown that adults are able to implicitly learn word-object pairings [13] and encode words with fine phonological detail [1] in their L1 via cross-situational word learning (XSWL) – a paradigm which involves listeners inferring word-object associations by tracking word-object co-occurrences across a number of learning trials [14]. Using this paradigm, native AusE monolingual and bilingual adults were tested on their ability to learn novel AusE words [1]. Participants were presented with novel words in CVC structure which formed nonMPs, vowel-MPs or consonant-MPs. Vowel-MPs differed only in their vowel (e.g. DIT-DEET) and consonant-MPs differed only in

their initial consonant (e.g. BON-TON). While word learning performance was worse for vowel-MPs than nonMPs and consonant-MPs, listeners performed above chance for all pair types, demonstrating that adults can successfully learn novel L1 words implicitly regardless of their language background.

Therefore, the present study aimed to test whether AusE-speaking monolingual and bilingual adults were able to implicitly learn novel L2 words via XSWL. Using an implicit version of the task reported in Elvin et al. [2], we manipulated the perceptual difficulty of the MPs based on the vowel categories (i.e. easy-MPs, difficult-MPs). The difficult-MPs refer to the vowel contrasts in which both vowels are acoustically closer together based on their F1 (tongue height) and F2 (tongue backness) values, while the easy-MPs refer to the vowel contrasts in which the vowels are acoustically more distinct. We examined participants' accuracy when identifying the novel BP words after exposure to the learning trials. We hypothesized that participants would show an overall accuracy and that both nonMPs and MPs will be above chance, demonstrating that adults can learn foreign language words and at the same time pay attention to phonetic detail when learning words implicitly, that is, without explicit instruction of which word corresponds to which object referent.

In line with Elvin et al.'s [2] findings for explicit L2 word learning, we also predicted higher accuracy for nonMPs compared to MPs, and easy-MPs compared to difficult-MPs. It was further anticipated that there would be no significant difference in the word learning accuracy of the monolinguals compared to the bilinguals [1].

2. Method

2.1. Participants

Participants included 10 first year psychology undergraduates from Western Sydney University, ranging from 17-38 years of age. Five participants were native AusE monolinguals, while five were AusE-speaking bilinguals with intermediate to native-like proficiency in at least one other language. Participants had no prior knowledge of the Portuguese language so as to yield reliable results. Participants recruited through Western Sydney University's research participation system SONA received course credit, while those recruited by word of mouth received monetary compensation for their participation.

2.2. Stimuli

The stimuli for the experiment included 14 novel BP words consisting of the Portuguese vowels /i, e, ɛ, a, o, ɔ, u/ which were used in [2] and were selected from Escudero, Boersma, Rauber and Bion's [15] corpus. Seven target words were in fVfe form and contained each of the seven vowels. The other seven words (kɔko, kuke, pipe, popo, sase, sɛso, teko) were distractor words recorded by a native female speaker of BP. The word pairs comprised 70 nonMPs, 15 easy-MPs and 4 difficult-MPs.

The six difficult vowel contrasts were /a-ɔ/, /a-ɛ/, /e-i/, /o-u/, /e-ɛ/, and /o-ɔ/, based on the close positioning of the vowels with respect to their F1 and F2 values (see Figure 1 which shows the values for the seven Portuguese oral vowels from [15] and the 12 AusE vowels from [16]). The remaining 15 vowel contrasts were considered perceptually easy: /a-e/, /a-o/, /a-i/, /a-u/, /e-o/, /e-ɔ/, /e-u/, /ɛ-i/, /ɛ-o/, /ɛ-ɔ/, /ɛ-u/, /i-o/, /i-ɔ/, /i-u/, and /ɔ-u/, as the vowels in each contrast are acoustically more distinct from each other (see Figure 1).

Each word was randomly paired with a novel object. All participants were presented with the same word-object pairings across learning and test trials.
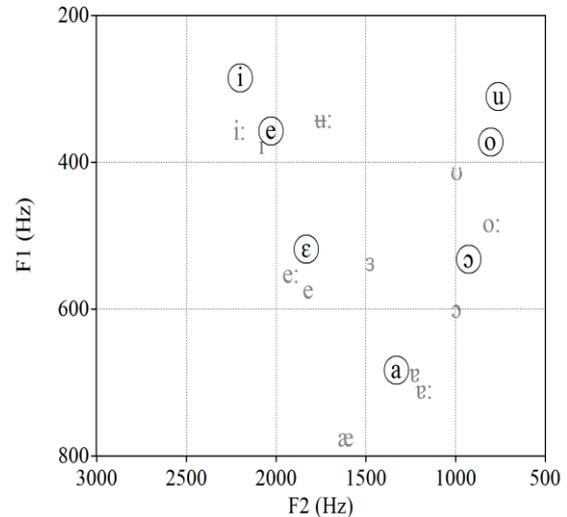

Figure 1: *Male speakers' average F1 and F2 values for BP vowels (black with circles) [15] and AusE vowels (grey) [16]*

2.3. Procedure

The study was an implicit version of the task in [2]. Participants first completed a language background form and provided consent. During learning, participants viewed novel objects and heard novel words on a screen without receiving explicit instruction to make any associations between them. The learning phase consisted of 84 trials, each word appearing as the target six times. During each trial, two images appeared side by side on a white screen (800 x 600 resolution). The audio words for the objects were presented 500 ms after the objects appeared, named in a random order with 500 ms between each word (either right/left or left/right, counterbalanced between participants). Participants were not given any instruction during the learning phase. The word pairs in each learning trial formed nonMPs, easy-MPs or difficult-MPs.

During testing, participants were presented with the same objects and words as those in the learning phase on a laptop. For each trial, two images appeared side by side on the screen. After 500 ms, participants heard an audio word corresponding to one of the objects and had to select whether they thought the word corresponded to the left or the right image via key press. The test phase consisted of 280 nonMP trials (70 pairs x 4 times), 60 easy-MP trials (15 pairs x 4 times) and 24 difficult-MP trials (6 pairs x 4 times), resulting in a total of 364 trials during which each word appeared as a target 26 times. The test phase took approximately 25 minutes for the participants to complete.

3. Results

Participants' accuracy rates were examined in a one sample *t*-test and it was found that participants performed better than chance for all three pair types (nonMPs: $M = .82$, $SE = .38$, $t[2799] = 44.317$, $p < .001$; easy-MPs: $M = .61$, $SE = .49$, $t[599] = 5.608$, $p < .001$ and difficult-MPs: $M = .57$, $SE = .50$, $t[239] = 2.080$, $p = .04$. This illustrates that all listeners were

able to infer object-word associations for each pair type. The mean word learning accuracy rates for the monolingual and bilingual groups are illustrated in Figure 2.

The accuracy rates for each language group were analysed in a 2 (Language group: monolingual, bilingual) x 3 (Pair type: nonMPs, easy-MPs, difficult-MPs) repeated measures ANOVA. No main effect of language group was found, F (1, 238) = .142, p = .707 as monolinguals (M = .63, SE = .03) and bilinguals (M = .64, SE = .03) had similar performance in word learning accuracy. A main effect of pair type was found, F (2, 476) = 8.748, p < .001, such that participants had significantly greater overall accuracy for nonMPs compared to difficult-MPs (β = .171, 95% CI [.072, .270], p < .001) and nonMPs compared to easy-MPs (β = .129, 95% CI [.026, .232], p = .008). There was no interaction found between pair type and language group, F (2, 476) = 1.346, p = .261: the mean accuracy scores were similar for all three pair types regardless of language background (monolinguals; nonMPs: M = .71, SE = .04; easy-MPs: M = .64, SE = .05; difficult-MPs: M = .54, SE = .05; and bilinguals; nonMPs: M = .77, SE = .04; easy-MPs: M = .58, SE = .05, difficult-MPs: M = .59, SE = .05.
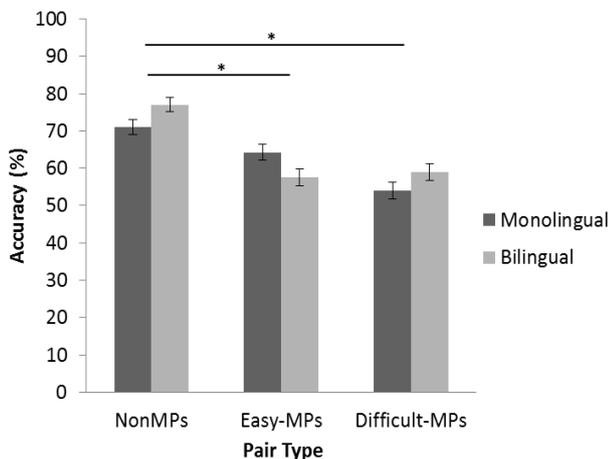

Figure 2: *Mean word learning accuracy rates across the three pair types (nonMPs, easy-MPs, difficult-MPs) for the monolingual and bilingual groups. * represents p < .05*

With respect to the difficult-MPs, no main effect of language group was found, F > 1; monolinguals (M = .54, SE = .05) and bilinguals (M = .59, SE = .05) had similar performance in accurately learning the difficult vowel contrasts. No main effect of difficult-MP type was found, F (5, 190) = 1.339, p = .249, and post-hoc pairwise comparisons showed no significant differences between difficult-MP types. There was no interaction found between language group and difficult-MP type, F > 1; mean accuracy scores were similar for all six difficult-MP types regardless of language background.

The results show that on average, the bilingual group had difficulty learning the BP difficult vowel contrasts in the following order (from least difficult to most difficult): /e-i/, /a-ɛ/, /a-ɔ/, /e-ɛ/, /o-u/, /o-ɔ/, while the following order was found for the monolingual group: /e-i/, /a-ɛ/, /a-ɔ/, /e-ɛ/, /o-ɔ/, /o-u/.

Table 1: *Average accuracy scores and standard errors for the bilingual and monolingual groups across the six difficult-MP types*

	Difficult-MP Type	Mean (%)	Std. Error
Bilingual	a-ɔ	60	.113
	a-ɛ	65	.109
	e-i	70	.107
	o-u	55	.112
	e-ɛ	60	.113
	o-ɔ	45	.114
Monolingual	a-ɔ	55	.113
	a-ɛ	65	.109
	e-i	65	.107
	o-u	35	.112
	e-ɛ	55	.113
	o-ɔ	50	.114

4. Discussion

The present study shows that adults can successfully learn L2 words through implicit instruction. The finding that listeners were able to learn words regardless of their language background supports the findings of [1] who found that native AusE monolingual and bilingual adults were able to learn novel AusE words using XSWL. While adults rely primarily on explicit learning mechanisms, adults are able to develop explicit knowledge even in implicit training conditions [5] which could explain why the participants successfully learned novel L2 words across all three minimal pair types.

The finding that participants' accuracy scores were lower for the easy-MPs and difficult-MPs across both word learning conditions compared to nonMPs provides support for models of speech perception such as PAM [9] and L2LP [10, 11, 12] which state that L2 learners will initially perceive the sounds in the L2 based on the acoustic properties of their L1 categories. While the BP vowel contrasts in the study are present in English (see Figure 1), they would have been difficult for learners to perceive as the acoustic properties of the vowels vary in each language. The higher accuracy scores for the nonMPs (e.g. fofe-teko) might have been due to the fact that consonants are considered to be more salient in word recognition and have a larger lexical role in English [17]. Therefore, the words that differed in both their consonants and vowel would have been easier to discriminate than words that only differed in their vowel (e.g. fife-fafe). Furthermore, participants' overall lower accuracy on the easy-MPs and difficult-MPs may have been due to the fact that vowels are more important than consonants when identifying words in continuous speech [17] while the present study only presented listeners with isolated novel words.

While it is known that both explicit and implicit methods are viable for word learning [1, 2], there has been little research comparing implicit and explicit word learning in adults. In order to investigate whether adults' word learning performance is better when learning words implicitly or explicitly, we looked at the mean accuracy scores for the 10 participants in the present study and 10 participants from Elvin et al.'s [2] study.

Table 2: *Average accuracy scores for implicit and explicit word learning across nonMPs, easy-MPs and difficult-MPs (2 d.p.)*

	NonMPs	Easy-MPs	Difficult-MPs
Implicit	82.11	61.17	56.67
Explicit	96.21	79.67	60.83

Based on the average word learning accuracy scores for participants in the implicit and explicit conditions, it can be argued that adults perhaps learn L2 words better when receiving explicit instruction rather than implicit instruction. Participants in both conditions were better able to learn words that formed nonMPs than words that formed MPs, and easy-MPs compared to difficult-MPs. While we investigated implicit word learning of individuals with mixed language backgrounds, testing a larger number of AusE monolinguals and then comparing our findings to those of [2] could determine whether one method of word learning is stronger than the other.

While no main effect of BP vowel contrast was found in the present study, [2] found a main effect of BP vowel contrast, suggesting that there are in fact specific vowel contrasts that are easier or harder to learn when compared to the L1 vowel repertoire. Therefore, further testing will allow us to answer whether there are specific BP vowel contrasts that lead to better or worse performance in implicit word learning between monolinguals and bilinguals when the sample size is increased. This will elucidate how specific vowel categories influence word learning. [2] found that on average, AusE listeners had difficulty learning the BP vowel contrasts in the following order (from least difficult to most difficult): /a-ɔ/, /a-ɛ/, /e-ɛ/, /i-e/, /o-ɔ/ and /o-u/ (similar to the current findings). The authors suggested that the vowel contrasts /a-ɔ/, /a-ɛ/ were easier for the listeners to recognise due to the fact that both vowels in the contrasts are acoustically similar to different native categories meaning there is less perceptual overlap with other sounds.

Ongoing research at our lab will also assess infants' ability to learn L2 words via implicit and explicit word learning paradigms and then compare the results to the present findings in order to determine whether there is a developmental trend in learning strategies, as proposed by [5]. Regarding possible limitations of the study, as there were a large number of learning trials ($n = 84$), some participants may have guessed the purpose of the learning phase, even without receiving explicit instruction to match the words to the objects. Asking participants if they understood the purpose of the learning trials after completion of the task may allow us to examine if there are any differences in performance between participants who had determined the intention of the study and those who had not. Overall, assessing methods of word learning has important implications for classroom settings and the development of language learning programs, and we believe that our results add to the literature on how adults can effectively learn an L2.

6. References

[1] Escudero, P., Mulak, K. E., & Vlach, H. A., "Cross-situational learning of minimal word pairs", Cognitive Science, 1-11, 2015.

[2] Elvin, J., Escudero, P., Williams, D., Shaw, J., & Best, C.T., "The role of speech perception and individual differences when learning words in a non-native language: English vs Spanish learners of Portuguese", under review.

[3] Hulstijn, J. H., "Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction", Studies in Second Language Acquisition, 27(2): 129-140, 2005.

[4] Ellis, R., "Measuring implicit and explicit knowledge of a second language: a psychometric study", Studies in Second Language Acquisition, 27: 141-172, 2005.

[5] Lichtman, K., "Age and learning environment: Are children implicit second language learners?", Journal of Child Language, 43(3): 707-730, 2016.

[6] Muñoz, C. [Ed.], Age and the rate of foreign language learning (Vol. 19), Multilingual Matters, 2006.

[7] Bialystok, E., "Representation and ways of knowing: three issues in second language acquisition", in N. C. Ellis [Ed.], Implicit and Explicit Learning of Languages, 549-569, San Diego, CA: Academic Press, 1994.

[8] Escudero, P., Broersma, M., & Simon, E., "Learning words in a third language: Effects of vowel inventory and language proficiency", Language and Cognitive Processes, 28(6): 746-761, 2013.

[9] Best, C. T., "A direct realist view of cross-language speech perception", in W. Strange [Ed.], Speech Perception and Linguistic Experience, 171–204, Timonium, MD: York Press, 1995.

[10] Escudero, P., "Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorisation", Ph.D. thesis, LOT Dissertation Series 113, Utrecht University, Utrecht, 2005.

[11] Escudero, P., "The linguistic perception of similar L2 sounds", in P. Boersma and S. Hamann [Ed.], Phonology in Perception, 152-190, Berlin: Mouton de Gruyter, 2009.

[12] van Leussen, J.-W., & Escudero, P., "Learning to perceive and recognize a second language: the L2LP model revised", Frontiers in Psychology, 6: 1–12, 2015.

[13] Fitneva, S. A., & Christiansen, M. H., "Looking in the wrong direction correlates with more accurate word learning", Cognitive Science, 35(2): 367-380, 2011.

[14] Yu, C., & Smith, L. B., "Rapid word learning under uncertainty via cross-situational statistics", Psychological Science, 18(5): 414-420, 2007.

[15] Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A., "A cross-dialect acoustic description of vowels: Brazilian and European Portuguese", The Journal of the Acoustical Society of America, 126(3): 1379-1393, 2009.

[16] Elvin, J., Williams, D. & Escudero, P., "Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English", Journal of the Acoustical Society of America Express Letters, 140(1): 576-581, 2016.

[17] Nespor, M., Peña, M., & Mehler, J., "On the different roles of vowels and consonants in speech processing and language acquisition", Lingue e Linguaggio, 2(2): 203-230, 2003.