

[THIS IS A PREPRINT OF A MANUSCRIPT ACCEPTED IN THE JOURNAL
“BILINGUALISM: LANGUAGE AND COGNITION”]

Running-Head: Bilingual Statistical Learning

Title: Improved statistical learning abilities in adult bilinguals

Author’s names: Luca Onnis, Win Ee Chun, Matthew Lou-Magnuson

Affiliation: Nanyang Technological University, 14 Nanyang Drive, Singapore 637332

Acknowledgments

Support comes from Singapore Ministry of Education Tier 1 grant RG81/14 to L.O., and NTU-URECA fund to L.O. and W.E.C. Shimon Edelman, Beth O’Brien, and Kathleen Ang for commenting on earlier versions of this manuscript.

Address for correspondence: lucao@ntu.edu.sg

Keywords: Artificial grammar learning; Bilingualism; Implicit Learning; Language Dominance; Mixed-effects models; Statistical learning.

Abstract

Using multiple languages may confer distinct advantages in cognitive control, yet it is unclear whether bilingualism is associated with better implicit statistical learning, a core cognitive ability underlying language. We tested bilingual adults on a challenging task requiring simultaneous learning of two miniature grammars characterized by different statistics. We found that participants learned each grammar significantly better than chance and both grammars equally well. Crucially, a validated continuous measure of bilingual dominance predicted accuracy scores for both artificial grammars in a generalized linear model. The study thus demonstrates the first graded advantage in learning novel statistical relations in adult bilinguals.

1. Introduction

The ability to track raw frequencies, probabilities, and other patterns of day-to-day sensory experience, and to abstract patterns of regularities embedded in sensory information is evidence of a powerful human cognitive process known as statistical learning (SL). This set of mechanisms works behind the scenes, as a form of implicit knowledge construction. In experimental contexts with presented sequences of seemingly meaningless syllable, shape, or sound stimuli, adults and children alike require only a few minutes of exposure to implicitly track low level statistical information embedded in those sequences (e.g., frequencies of individual and grouped elements, or transitional probabilities of adjacent and non-adjacent dependencies). The few minutes' exposure is sufficient to allow differentiation between novel sequences that have either the same, or a different structure – similar to the way we notice when a radio station switches between two unfamiliar languages. Several studies have linked individual variance on SL tasks with individual variance in tasks of natural language learning and processing, thus suggesting that the two abilities are related (e.g., Conway, Christiansen, & Onnis, 2012; Misyak & Christiansen, 2012). For example, infants exhibit individual differences in statistical learning skills that may modulate their language development trajectories as young children, including the development of comprehension of syntax (e.g., Kidd and Arciuli, 2016; Kidd, 2012). Studies with older children have also linked poor implicit statistical skills with concurrent language and/or reading difficulties (Evans, Saffran and Robe-Torres 2009; Yim and Windsor, 2010).

Considering that the experience of acquiring two languages requires bilinguals to track multiple distinct sets of statistical regularities and thus may be related to their sensitivity and approach to novel statistical regularities (Weiss, Poepsel & Gerfen, 2015), a pertinent research interest would be the relationship between bilingualism and SL. The limited literature on the topic to date suggests that the relationship between bilingualism and SL is nuanced. Bilingual experience was demonstrated to enhance learning outcomes in speakers for an artificial tone

language more than prior experience with tonal languages (Wang & Saffran, 2014). In a word learning task that required participants to extract and learn novel words comprising of pure tones based on Morse code in a continuous auditory stream, participants with higher bilingual experience were found to be better at learning words through the tracking of transitional probabilities (Bartolotti et al., 2011). In a more recent study, Escudero et al. (2016) found that bilingual adults were more accurate than monolinguals at picking up novel word-referent mappings in an implicit cross-situational word learning task. However, in contrast to the positive findings above, Yim and Rudoy (2013) did not find any differences between monolingual and bilingual children on a nonlinguistic auditory tone task and a visual statistical learning task. Similarly, Potter, Wang and Saffran (2015) found an advantage in performance on an artificial tonal language learning task, but not for a visual statistical learning task between a Mandarin-learning group and a control group. Thus, the mixed findings presented suggest that the relationship between bilingualism and SL is not straightforward. Furthermore, the focus on comparing bilingual groups and monolingual controls in this line of research underscores an underlying assumption that bilingualism is an all-or-none phenomenon. In reality, bilinguals vary in their relative proficiency, experience and dominance in their individual languages. Consequently, this study aims to examine how individual differences in bilingual dominance may be related to individual differences in SL.

A concurrent theoretical reason to explore bilingual implicit learning is the putative bilingual cognitive advantage. Bilinguals may benefit from enhanced executive control, particularly aspects of attentional control and inhibition of irrelevant information (Bialystok, Craik & Luk, 2012), as well as greater flexibility in shifting between mental sets (e.g., Prior & MacWhinney, 2010). While others have questioned the bilingual cognitive advantage on methodological grounds (De Bruin, Treccani and Della Sala 2014; Paap, Johnson and Sawi 2015), thus keeping the debate open (Bialystok, Kroll, Green, MacWhinney and Craik, 2015; Paap, Johnson and Sawi, 2016), here we note that it has largely focused on executive function

skills, and mostly neglected *learning* abilities. Thus, considering the evidence above that bilingual experience may facilitate learning outcomes, we hypothesized that degree of bilingualism is related to implicit learning ability; more balanced bilinguals are expected to exhibit greater SL ability than those who are comparatively dominant in one language.

1.1. Method

To investigate whether adult bilinguals exhibit heightened statistical learning, we looked at their ability to learn two artificial grammars concurrently, and predicted their learning scores from their degree of bilingualism.

1.1.1. Participants. Fifty-five undergraduate students (33 females, mean age=22.02 years, $sd=1.49$) at a university in Singapore participated for a small monetary token. All but one were born and lived in Singapore all their lives, and were educated through the same national educational system. They reported being bilingual in English and in their heritage language (Mandarin=51, Malay=3, Filipino=1). This sample size was calculated using the *pwr.f2.test* function in R to achieve a medium effect size with a significance value less than .05 and a power of .8, for a planned multiple regression. In addition, our sample size is more than 4 times the one required to obtain a medium effect size on mean proportion accuracy in the artificial language task planned (see below).

Language dominance. In Singapore multilingualism is the norm and monolingualism the exception. At the same time, Singaporeans' individual experiences with multiple languages vary greatly, and can yield quite different linguistic profiles from person to person. We decided to capitalize on such variability, rather than treating bilingualism as a homogeneous variable to be tested against monolingualism. The Bilingual Language Profile (BLP) is a validated questionnaire for assessing language dominance through self-reports (Birdsong, Gertken & Amengual, 2012; for measures of validity and reliability see Gertken, Amengual, & Birdsong, 2014). It produces a continuous dominance score and a general bilingual profile taking into

account variables of linguistic background, use, and ability. Each question response in the BLP is a scalar associated with a certain point value. We first tallied the point totals based on self-reported values in each language separately for four subcomponents¹: Language History (6 questions), Language Use (5 questions), Language Proficiency (5 questions), and Language Attitude (4 questions). To obtain absolute scores of the four subcomponents independently for each language, the questionnaire was administered twice, once in English and once in the equivalent translation in the other language of the participants. We then obtained a weighted sum of the four subcomponents to ensure that each subcomponent received equal weighting, to yield a global score for each language. The maximum total score possible for each language is 218 points. Finally, we obtained a composite language dominance index for each participant, by subtracting their least dominant language total from the most dominant one. This rendered a dominance score that ranged from 0 (perfectly balanced in both languages) to +218 (strongly dominant in one language only). We used this composite measure as index of bilingual dominance.

1.1.2. Material and procedure. The task was modeled and adapted from Conway and Christiansen (2006). We generated two different finite-state grammars, Grammar A and Grammar B (Figure 1), each with its own sets of non-overlapping stimuli. We used 9 grammatical sequences from each grammar in the training phase and 10 grammatical sequences from each grammar in the test phase; all sequences contained between three and seven elements. For a given grammar, each shape was randomly paired with a pseudoword from one of two lexicons – the paired unit is indicated by a letter symbol (see Figure 1). The pairing of shapes and pseudowords for each symbol of each grammar within a lexicon was random for each different participant, thus reducing undesired group-wise sequence-specific biases on learning. Therefore, each letter symbol was mapped onto both a shape lexicon (five different

¹ <https://sites.la.utexas.edu/bilingual/blp-components/>

shapes) and an auditory lexicon (five spoken pseudowords). Specifically, Grammar A was associated with 5 blue shapes, and 5 pseudowords (*rud, pel, dak, vot, jic*) generated with the English speaking voice Victoria available from the Speech System Manager of Mac OS X. Grammar B was associated with a different lexicon of 5 red shapes, and 5 pseudowords (*ginot, labou, liva, taret, kimosse*) generated with the French speaking voice Thomas available from the same software. Conway & Christiansen (2006, Experiment 3) showed that learning suffered when the two grammars used lexicons along the same perceptual dimension; in this case, statistical learning was limited to just one of the two grammars. In a similar task involving statistically-based segmentation of two continuous streams of sounds into word-like units, Weiss, Gerfen & Mitchel (2009) found that participants were unable to extract two sets of regularities when the same speaker voice was used. Thus, the purpose of using shapes of different color and pseudowords of different pronunciations in our stimuli was to provide robust perceptual cues for the change of grammars. The pseudowords and shapes were presented concurrently, thus forming a multimodal stimulus percept. Visual stimuli were presented in a serial format in the center of a computer screen. Pseudoword stimuli were also presented in sequence and timed to the visual stimuli via headphones. Each shape of a particular sequence was presented for 500 ms, with 100 ms occurring between shapes. A 1,700-ms pause separated each sequence from the next.

Participants were told that they would experience sequences of shapes and sounds, and it was important to pay attention to the stimuli because afterward they would be tested on what they had observed. The instructions did not indicate that the sequences followed underlying rules or regularities of any kind, nor that two sets of regularities were present. In the training phase 18 sequences (9 from each grammar) were pseudo-randomly presented in 6 blocks of interleaved sequences, mimicking the language switches characteristic of bilingual speech. Thus, a total of 108 sequences were presented during training. Participants could take a brief break in between blocks.

Before the beginning of the test phase, participants were informed that the stimuli they had observed were generated according to a complex set of rules that determined the order of the elements within each sequence. Participants were told they would next be exposed to other sequences. Some of these sequences would conform to the same set of rules as before, whereas the others would be different. Their task was to judge which of the sequences followed the same rules as before and which did not.

For the test phase, a total of 20 test sentences were used, 10 that were grammatical according to the word order of Grammar A, and 10 that were grammatical according to the word order of Grammar B. For scoring purposes, the test sequences from the grammar that was instantiated with the same stimuli as in the training phase were deemed grammatical, whereas the test sequences from the other grammar were deemed ungrammatical. This effectively implemented a crossover design (Conway & Christiansen, 2006) in which half of the grammatical test sequences of one grammar were used as the ungrammatical test sequences for the other grammar. Crucially, for participants to show that they learned the statistical regularities specific to each grammar, they ought to classify a sequence as grammatical only if it was presented in the same stimuli as were the training sequences generated from the same grammar (i.e., with the same visual stimuli mapped to the same pseudowords spoken with the English or French voice). As an analogy from natural languages, word order in Japanese and English differs, and a Japanese-English bilingual can tell that "I you like" is ungrammatical in English while the same order in Japanese would be grammatical. This ability requires subtly internalizing the statistical preferences of each grammar.

All test items were presented in random interleaved order. Participants pressed "Y" on their keyboard if they thought that a particular test sequence conformed to the rules, and "N" if they did not. Classification judgment was scored as correct if the test sequence was judged as grammatical and its stimuli were the same as those of the training sequences that were generated from the same grammar. Similarly, a classification judgment was also scored as

correct if the test sequence was judged as ungrammatical and its stimuli were different from those of the training sequences that were generated from the same grammar. In all other cases, a classification judgment was scored as incorrect.

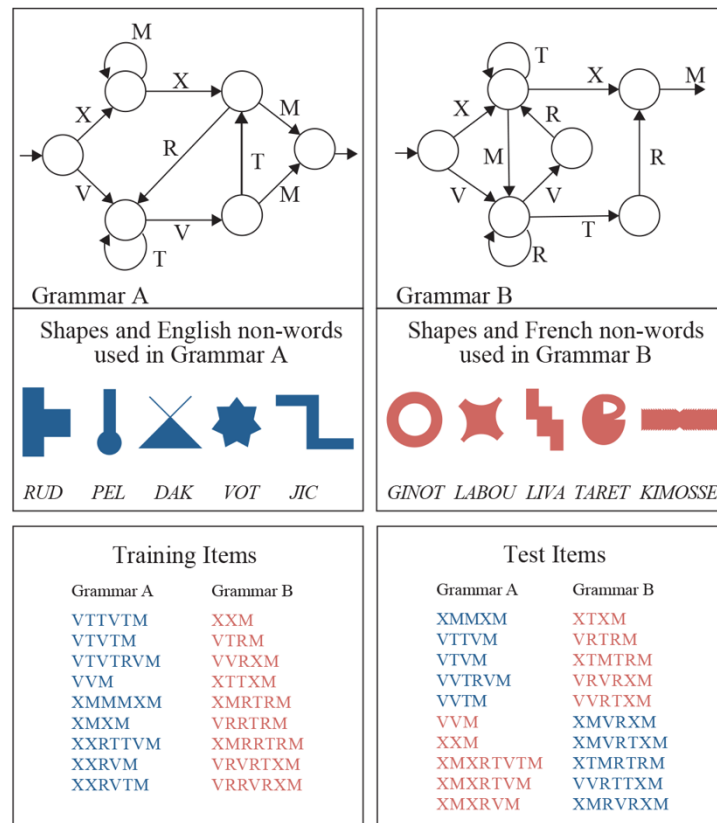


Figure 1. The grammars, training items, and test items used in the experiment. The top panels depict the state transitions generating all grammatical sequences by Grammar A and Grammar B respectively. The training items were mapped onto two lexicons of non-words and coloured shapes (centre panels). To create ungrammatical items at test, half of the test items for each grammar were mapped onto the non-words and shapes of the other grammar.

1.2. Results

Data from two participants were not included in the statistical analyses because incomplete in one of the two critical measures. We first calculated accuracy scores on the dual grammar task for each participant and each artificial grammar as the proportion of correct

endorsements to grammatical test items and correct rejections to ungrammatical test items. As a group, our participants learned both languages significantly above chance (Figure 2): Grammar A (mean=0.57, sd=0.18, $t(52)=2.69$, $p<0.01$, $d=0.37$) and Grammar B (mean=0.58, sd=0.19, $t(52)=3.21$, $p<0.01$, $d=0.45$). The measure of Language dominance obtained from the questionnaire ranged from 1.58 (close to equally bilingual) to 85.65 (closer to dominant in one language), (mean=32.15, sd=20.78), indicating variability in our sample (Figure 3, top). Forty-four participants were more dominant in English (the global score of English was higher than the global score of the Other language), while 9 were more dominant in the other language. Across all participants, global scores of English calculated by summing scores for Language History, Use, Proficiency, and Attitudes were generally higher (min= 101, max=181, mean=158.5), than global scores obtained from the other language (min= 58, max=176, mean=115.7).

To assess whether language dominance (measured as the subtraction of the least dominant language global score from the most dominant one) predicted accuracy scores of the artificial grammar task, we fitted a mixed-effects generalized linear model using the *glmer* function in R (R Core Team, 2015). Accuracy scores were modeled as a binomial distribution and regressed against Grammaticality of test item (Grammatical, Ungrammatical), Grammar (A, B), and Language Dominance as fixed effects. We also regressed Participants and Test Items as random effects. We first fitted a maximal model with all predictors and interactions between the fixed factors. We then performed a stepwise model selection by AIC to select the most parsimonious model. The final model indicated two main effects: Grammaticality and Language Dominance. Grammaticality ($\beta=0.66$, $z=2.73$, $p<0.01$, $d=1.92$) suggests that participants were better at endorsing grammatical test items than rejecting ungrammatical test items. This is in line with previous literature on artificial grammar learning (e.g., Pothos, 2007). In addition, Language dominance independently predicted accuracy scores ($\beta=-0.23$, $z=-2.155$, $p=0.03$, $d=0.79$); see Figure 3, bottom). Thus, participants with a more balanced bilingual

profile performed better than less bilingual individuals.

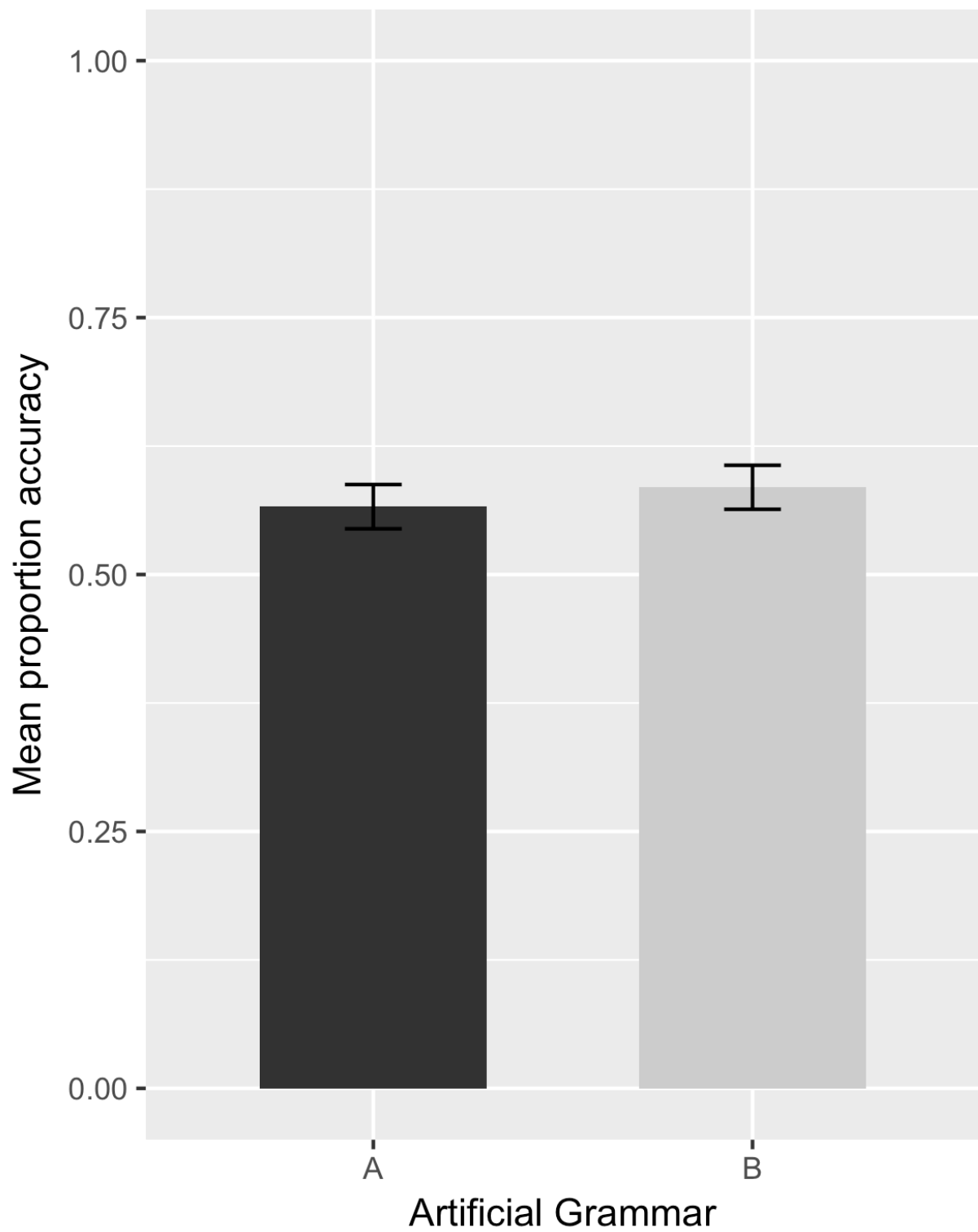


Figure 2. Mean proportion accuracy scores in each artificial grammar, with standard error bars.

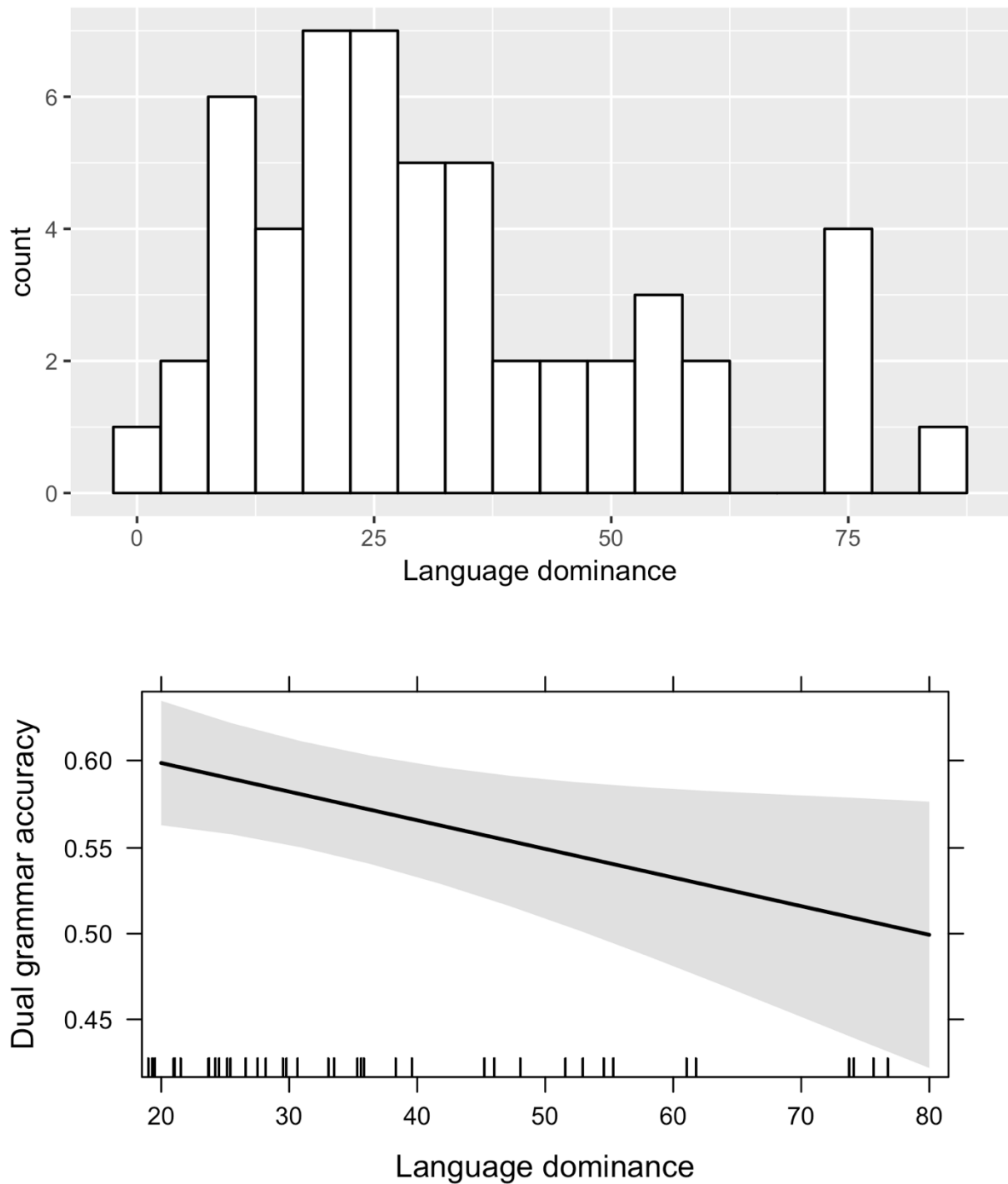


Figure 3. Top: the distribution of language dominance scores in the sample of participants. Higher scores indicate a more monolingual profile. Bottom: Plotted effects (and confidence intervals) of Language dominance on the combined mean accuracy scores from both grammars, obtained from a mixed-effects logistic regression. Higher bilingualism predicts higher accuracy scores on the dual grammar task.

2. General Discussion

The present study investigated whether individual differences in bilingualism correlate with the ability to learn two distinct sets of statistical regularities concurrently in a dual artificial grammar learning task. Previous research suggests that this is a particularly challenging task (Conway & Christiansen, 2006; Franco, Cleeremans & Destrebecqz, 2011; Weiss, Gerfen & Mitchel, 2009). Furthermore, Yim and Rudoy (2013) and Poepsel and Weiss (2016) found that both monolingual and bilingual children and young adults performed statistical learning tasks with novel miniature grammars equally well, suggesting no bilingual advantage in this cognitive domain. Here we built on and incorporated experimental aspects of all these studies, producing two different sets of regularities associated with different voices, language accents, and visual stimuli. Yet the task required subtle differentiation and learning because of our crossover design. In other words, learning must be language-specific in this paradigm, otherwise performance would be at chance levels as all underlying sequences were grammatical. Participants had to recognize which word-shape order sequence matched a specific grammar, not unlike a bilingual Japanese-English speaker recognizing that "Cats mice eat" is ungrammatical in English but correct in Japanese. Participants learnt both languages at above-chance levels. Crucially, in line with the hypothesis, participants that exhibited a more balanced bilingual profile (relatively equal dominance in both languages) performed better on the grammar test, being more adept at correctly selecting sequences that were grammatical and rejecting sequences that were ungrammatical, irrespective of language.

Research on the extralinguistic benefits of bilingualism has primarily centered on cognitive benefits such as attentional control and task switching (Bialystok, Craik & Luk, 2012). Considering the importance of SL in language, few studies have investigated whether bilingualism modulates SL ability. This study partially fills this gap by proposing that bilinguals may exhibit heightened statistical learning abilities, thus extending the scope of the bilingual advantage. Our findings also contribute to qualify the nature of the advantage as a graded one.

We adopted an individual differences approach and treated bilinguals as a heterogeneous group, rather than as a homogeneous group contrasted to monolinguals. The latter approach is frequent, but may obfuscate gradual dimensions of bilingualism that in turn reflect different cognitive sensitivities. The self-reported Bilingual Language Profile aptly revealed a combined large variability across four holistic dimensions of language history, use, proficiency, and attitudes for each of our participants' languages. By capitalizing on the bilingual variability we found in the questionnaire rather than ignoring it, we unearthed important individual differences that point to the first documented modulating role of bilingualism in adult statistical learning.

One limitation of the current study is that it was not possible to assess whether the bilingual advantage found here involved specific modalities of statistical learning, because in our paradigm the grammars were instantiated by pairing visual non-linguistic with linguistic stimuli. Current empirical evidence suggests that SL may not be a unified domain-general ability, as its modality- and informational specificity across studies suggest that it is instead comprised of multiple distinct subcomponents (Arciuli, 2017; Siegelman et al., 2017). Thus, SL tasks are not equal in their capacity to capture the specific subcomponents of SL; different tasks may capture different facets of SL. This new theoretical understanding of SL may elucidate the mixed findings seen in studies such as Wang and Saffran (2014), where the bilingual advantage in SL was found only in performance on the artificial tonal language learning task, but not a visual statistical learning task. Thus, further studies should investigate whether the bilingual advantage we found in statistical learning is modality-specific.

Further research may also build upon this finding by examining the directionality of this relationship, perhaps through structural equation models as Spencer et al. (2015) previously demonstrated. One interpretation compatible with our findings is that in order to succeed in bilingual learning, bilinguals must necessarily learn and keep separate distinct sets of statistical regularities for each language, and in doing so their statistical learning skills might be sharpened. However, the opposite direction of causality can also be entertained in principle: it is

possible that individuals equipped with better SL abilities are able to better learn and maintain two languages in their life. Additionally, future studies that aim to investigate how individual differences in bilingualism modulate SL should also account for general cognitive abilities that may influence task performance, such as IQ, attention, and working memory. Age may also be a potential factor to be controlled for, as current empirical evidence suggests that there are age-related effects on SL ability (Arciuli, 2017).

Together with previous studies documenting a relation between statistical learning and language, a broader picture is emerging: not only is statistical learning associated with the acquisition and processing of language, but also specific experiences with language - here bilingualism - may be implicated in individuals' statistical learning abilities.

[Word count: 3,205]

Author Contributions

L.O. developed the study concept and design, performed the data analysis and interpretation, and drafted the manuscript. W.E.C. performed recruitment and data collection, calculated Language dominance scores, and assisted L.O. in data analysis and manuscript revision. M. L.M. wrote the PsychoPy scripts for the experiment and assisted in data analysis and manuscript revision. All authors approved the final version of the manuscript for submission.

References

- Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society B*, 372: 20160058.
- Bartolotti, J., Marian, V., Schroeder, S., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Cognition*, 2(324), 1-9.
- Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, 16(4), 240–250.
- Bialystok, E., Kroll, J. F., Green, D. W., MacWhinney, B., & Craik, F. I. (2015). Publication bias and the validity of evidence what's the connection? *Psychological science*, 26(6), 944-946.
- Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual language profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*.
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities pitting abstract against stimulus-specific representations. *Psychological science*, 17(10), 905–912.
- Christiansen, M. H., Conway, C. M., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: evidence from event-related brain potentials. *Language and cognitive processes*, 27(2), 231-256.
- De Bruin, A., Treccani, B., & Della Sala, S. (2014). Cognitive advantage in bilingualism an example of publication bias? *Psychological science*, 0956797614557866.
- Escudero, P., Mulak, K. E., Fu, C. S., & Singh, L. (2016). More limitations to monolingualism: bilinguals outperform monolinguals in implicit word learning. *Frontiers in psychology*, 7.
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52(2), 321–335.
- Franco, A., Cleeremans, A., & Destrebecqz, A. (2011). Statistical learning of two artificial languages presented successively: how conscious? *Frontiers in psychology*, 2, 229.

- Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. *Measuring L2 proficiency: Perspectives from SLA*, 208-225.
- Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, 116(3), 321–340.
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental psychology*, 48(1), 171.
- Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child development*, 87(1), 184–193.
- Misyak, J. B. & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, 62(1), 302–331.
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69, 265–278.
- Paap, K. R., Johnson, H. A., & Sawi, O. (2016). Should the search for bilingual advantages in executive functioning continue. *Cortex*, 74(4), 305–314.
- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, 152, 9-19.
- Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological bulletin*, 133(2), 227.
- Potter, C. E, Wang, T. & Saffran, J. R. (2015). Second language experience facilitates statistical learning of novel linguistic materials. *Cognitive Science*, 41(4), 1–15.
- Prior, A. & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13(2), 253–262.
- R Core Team. (2015). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project.org/> (ISBN 3-900051-07-0).
- Siegelman, N., Bogaerts, L., Christiansen, M. H. & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal*

Society B, 372: 20160059.

Spencer, M., Kaschak, M. P., Jones, J. L. & Lonigan, C. J. (2015). Statistical Learning is Related to Early Literacy-Related Skills. *Reading and Writing*, 28(4), 467–490.

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.

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., Poepsel, T., & Gerfen, C. (2015) Tracking multiple inputs: The challenge of bilingual statistical learning. Rebuschat, P. (Ed.) *Implicit and Explicit Learning of Languages*. p167-190. John Benjamins Press.

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.

Yim, D., & Windsor, J. (2010). The roles of nonlinguistic statistical learning and memory in language skill. *Korean Journal of Communication Disorders*, 15(15), 381–396.