Title
Deep in the Trenches: First language performance predicts primacy in statistical learning of two structures

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Chough, Jino
Zinszer, Benjamin D.

Publication Date
2022

Peer reviewed

Deep in the Trenches: First language performance predicts primacy in statistical learning of two structures

Jino Chough (dwnchgh7@gmail.com)
Department of Psychology, 500 College Ave
Swarthmore, PA 19081

Benjamin Zinszer (bzinszer@swarthmore.edu)
Department of Psychology, 500 College Ave
Swarthmore, PA 19081

Abstract

While statistical learning is a well-established language learning mechanism, its usefulness in multiple language contexts is more unknown. A phenomenon known as entrenchment has been proposed, in which learning one language prevents the acquisition of a second language in the same speech stream. The observed L1 advantage or primacy effect has been previously mitigated with various cues to the presence of a second structure (L2). The present study manipulates the number of transitions between L1 and L2 to influence entrenchment. One condition was designed to replicate previous findings of entrenchment and the other was designed to overcome entrenchment. We find that adding more transitions between languages did not increase L2 learning, and second language learning is more dependent on the first learned language than on manipulations of the transitions between languages.

Keywords: speech segmentation; statistical learning; primacy; bilingualism

Introduction

Statistical learning (SL) is a mechanism that contributes to word segmentation in language learning, a discovery dating to the initial study showing infants are able to parse artificial language speech streams (Saffran, Aslin, & Newport, 1996). Adults have demonstrated similar proficiency for artificial language learning, which indicates that SL is not purely an ability for individuals in early stages of development (Saffran, Newport, & Aslin, 1996). While the artificial languages used in most SL studies are dissimilar to any real language due to differences in pitch, tone, and uniformity, SL is also effective for natural language segmentation in infants (Pelucchi, Hay, & Saffran, 2009). Not only that, but SL performance on artificial languages correlates with natural syntactic and semantic language skills, as well as vocabulary and reading ability (Newman et al., 2006; Evans, Saffran, & Robe-Torres, 2009; Arciuli & Simpson, 2012). Despite the differences between artificial and natural languages, the correlation between SL and aspects of natural language learning is well established.

With bilingualism, however, the relationship with SL is more complex. Bilingual individuals have varying ages and levels of acquisition with their second language. In short, real world language acquisition is varied by tone, pitch, or speaker, and the addition of another language makes learning by statistical regularities even more difficult (Qian, Jaeger, & Aslin, 2012). In the context of SL, however, the most prominent issue is how individuals are able to form different statistical representations for each artificial language instead of simply aggregating the two as one language (Weiss, Gerfen, & Mitchell, 2009). Using completely different sets of syllables leaves open the possibility that individuals perceive the entire sequence as a single language (Antovich & Graf Estes, 2017). Therefore, multiple artificial language studies often incorporate at least a partially overlapping syllable inventory to avoid that confound.

Gebhart, Newport, and Aslin (2009) tested how well participants were able to acquire two artificial languages sharing a partial syllable inventory and incompatible distributional statistics in a single speech stream. The authors found a strong primacy effect, in which participants were able to acquire L1 but were unable to acquire L2. This “entrenchment” in L1 was only offset with the use of an explicit cue, or when L2 exceeded L1 in duration by 3 times the length. This phenomenon of entrenchment parallels the native language neural commitment (NLNC) hypothesis proposed by Kuhl (2004) for natural language acquisition. Kuhl argues that, based on previous literature showing that language experience affects later language learning capabilities (Dehaene-Lambertz et. al, 2000; Callan et al., 2004), language learning produces neural networks that code the patterns of the experienced language, interfering with learning foreign languages that do not conform to learned patterns.

Other studies have manipulated the consistency of presentation of L1 and L2 to test whether multiple switches between languages would lead to acquisition of both languages instead of relying on L2 overexposure (Zinszer & Weiss, 2013). The authors compared the degree of primacy, the difference between performance on L1 and L2, for different configurations of the L1 and L2 exposure, finding the greatest L2 learning when participants were exposed to five and a half minutes of L1, then alternating 2 minute, 45 second blocks of L2, L1, and L2. In a similar study, Weiss, Gerfen, and Mitchell (2009) presented L1 and L2 in

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

alternating blocks of two minutes for 24 minutes total. They found that learning for both languages exceeded chance. The "unstable" exposure to L1 and L2 avoided the entrenchment effect in Weiss et al.'s study, which offers further support for the idea that the NLNC is a model for artificial language entrenchment. However, the unstable design of Zinszer & Weiss also overcame the same initial L1 entrenchment observed by Gebhart et al. despite a net increase of L1 exposure. These findings raise the question of how switching between languages may support L2 learning, by preventing entrenchment altogether or by providing a mechanism for overcoming entrenchment faster.

Karuza et al. (2016) raises the possibility that the brain's tendency for efficiency could result in entrenchment-like performance. The authors found that participants who learned less of L2 from a speech stream had decreased activity in the fronto-subcortical and posterior parietal regions, which are associated with attention and executive function. This suggests that failure to acquire L2 could be due to a lack of attention in the latter half of the speech stream, and the authors suggested that "inefficient" learning systems are more sensitive to structural changes of language and more likely to acquire L2. In line with these findings, an unstable speech stream, with more frequent transitions between languages as seen in Zinszer & Weiss (2013) or Weiss, Gerfen, and Mitchell (2009) may be more successful by capturing attention with sudden or earlier switches between languages. Furthermore, the study suggests that failure to learn a second language occurs due to a prolonged exposure to L1 changing attentional allocation rather than directly impeding learning.

Bulgarelli and Weiss (2016) aimed to directly limit entrenchment of L1 by providing frequent tests throughout L1 exposure and switching participants to L2 exposure once they scored well in L1, thus ensuring that they learned the language and transitioned immediately after said learning. Under this paradigm, participants were able to learn both languages, suggesting that previous literature replicating entrenchment was a result of overlearning L1, but the frequent starts and stops for assessment may have had an unintended result of sustaining participant attention as well. While in a uniformly unstable exposure to language, the frequent and early onset transitions between languages will not be as accurate in preventing overlearning, the study provides support that instability is more effective for multiple language learning due to reduced overexposure. Bulgarelli & Weiss's conclusions about overexposure, like the other entrenchment studies, offer support for the principles of the NLNC that state L1 learning directly prevents L2 acquisition.

In the context of these previous studies on entrenchment, we sought to test the link between over-learning and switch effects seen in artificial speech segmentation studies and the NLNC theory for natural language acquisition. We initially briefly exposed participants to L1 to estimate their level of L1 learning and avoid overlearning. The next day, we exposed participants to L1 and L2 in either a stable or unstable transition speech stream, as seen with Zinszer and Weiss (2013), and Karuza et al. (2016), in the hopes of respectively replicating the entrenchment effect and preventing L2 acquisition, from Gebhart, Newport, and Aslin (2009), or avoiding an overcommitment in one language and allowing L2 learning, as Weiss et al. (2009), Zinszer and Weiss (2013), and Bulgarelli and Weiss (2016) showed.

Based on previous literature concerning entrenchment, we predicted that faster initial learning of L1 on the first day would be negatively correlated with later L2 learning on Day 2. For our exploration of the NLNC, we predicted that L2 learning would be better for participants presented with the unstable speech stream, and within that unstable condition, we predicted that L1 on Day 1 and L2 on Day 2 performance would be positively correlated. Unlike many previous studies, which have screened exclusively for monolingual participants, we also recruited participants with varying levels of L2 proficiency allowing a secondary exploration of how experience with a natural L2 might modulate sensitivity to switch cues or entrenchment effects.

Method

Participants 64 undergraduate students (27/37/0 M/F/other) from Swarthmore College were recruited through physical advertisements posted around campus, word of mouth, or an introduction to psychology course. Participants were compensated with either ten dollars or course credit, and they were randomly assigned to one of four groups based on one of two experimental conditions and one of two first languages; each group had 16 participants. The mean age of the group was 19.8 years (20.1y for males, 18.8 y for females). The participants were not excluded based on their language history, resulting in a wide range of second language proficiency levels.

Language History Participants self-reported their learned languages and rated their own reading, listening, writing, and speaking skills with each language. They also self-reported frequency of exposure to each language via media, school, family, or friends, as well as the age they were first exposed to the language. For the purposes of this study, language proficiency was calculated using the mean of listening and speaking skills, allowing us to assign numerical scores to each natural language that the participants knew. All participants reported maximum or near-maximum proficiency in at least one language (i.e., their native language), and second language proficiency was calculated as the second highest score after that native language.

Speech Stream Languages The speech stream consisted of two languages previously used by Gebhart et al. (2009, Experiment 1b). Each language was composed of 12 syllables equally divided into four words. For both languages, there were two vowel frames and six consonants to define the words. In one language, the first syllable of a triplet led to one of two syllables, which then led to one of two other syllables. In this manner, within one language, each syllable was reliably placed in the first, second, or third syllable of the word triplet. However, the two languages shared 6 syllables with each other, leading to a 50% overlap, but the shared syllables marked different positions of the words within the language and could not be used as consistent cues for segmentation. The languages were able to be segmented only by tracking the transitional probabilities (TPs) between syllables, with high TPs between syllables or vowels composing a word and low TPs between syllables and vowels creating a word boundary. The within-word TP was 1 for vowels and 0.5 for syllables, and the between-words TP was 0.5 for vowels and 0.25 for syllables.

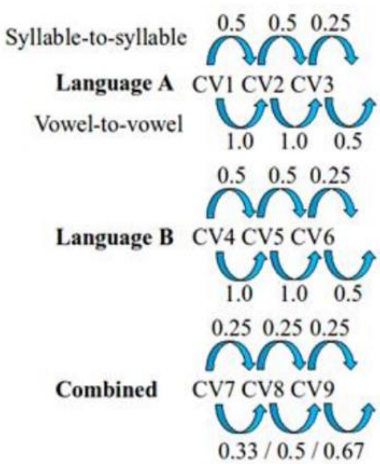

Figure 1: Transitional probabilities defining each language's structure. The transitional probabilities result in an unlearnable structure when combined.

Procedure Participants participated in a two day familiarization phase. On the first day, participants were instructed to listen to a recording of a speech stream and were informed that they would be quizzed on what they had heard. The participants listened to 2 minutes and 45 seconds of their first language, termed L1. After this initial exposure of L1, participants completed a test phase with a 16 two-alternative forced choice task. Participants selected between the statistically defined words from the stream and "part-words" that consisted of the same syllables. Participants were instructed to choose which between the word and part-word had occurred in the stream. Each option was presented with a visual cue on the left or right side of the monitor to indicate where participants should click to select their choice. The test items were presented one at a time, counterbalanced between which came first and between the words and part-words set against each other. The participants were given 10 seconds after the second choice was presented to decide before the test automatically advanced to the next question. The questions were randomly sorted, and spaced between each other by 1 second.

On the second day, participants listened to a speech stream according to one of two conditions, noted by the number of transitions between languages. The second language that they were exposed to was termed L2. Assignment of the two languages to the L1 and L2 positions was counterbalanced across all participants. In the "stable" condition, participants listened to an additional 5 minutes and 30 seconds of the same L1 from Day 1, with one transition (1T) to 5 minutes and 30 seconds of the L2. In the "unstable" condition, participants listened to L1 for a block of 2 minutes and 45 seconds, then blocks of L2, L1, and L2 for the same amount of time each, for a total of three transitions (3T) between languages. The total length of the speech streams were the same between conditions (11 minutes), as well as the exposure to L1 and L2 (Day 1: 2:45 L1; Day 2: 5:30 L1 and 5:30 L2). After the speech stream, participants completed another 2-AFC task of 32 questions. Half of the test items contrasted words and part-words in L1, and the other half of the test items contrasted words and part-words of L2. The order of the 32 questions was randomly sorted for each participant.

Analysis We estimated linear regression models for the magnitude of the primacy effect on Day 2 (L1-L2), as well as for L1 and L2 performance on Day 2 separately, L2 proficiency based on the switch condition, and language as predictor variables. Models included counterbalancing (which language was assigned as L1), condition (1T or 3T), Day 1 score on L1, and second language proficiency (scaled 0-1) as predictors. For the L1 and L2 models, we also included the other language's Day 2 accuracy score to assess whether performance in the two languages were positively or negatively correlated with each other.

Because the number of predictors was relatively large and most were not statistically significant, we used Akaike Information Criterion (AIC) for an automated stepwise regression to identify the best fit model and exclude extraneous predictors.

We also separated participants based on whether they performed better than chance (binomial test for 16 items, accuracy >0.75 for $p<0.05$) in L1 on the first day and computed means for their Day 2 performances.

Results

Participants scored above chance on L1 on both days and in both conditions. Neither group averaged above chance for L2. See Table 1 (next page) for the mean performances and Table 2 for mean participant performance versus chance-level (50%). Performances on Day 2 were not significantly different between the two groups (L1: $t(60)=0.800$, $p = 0.43$; L2: $t(54)=0.98$, $p = 0.33$). See Figure 2 (next page) for the graphic rendition.

The regression model to estimate primacy effect was statistically significant ($F=14.2$, $p<0.001$, adjusted $R^2=0.18$).

Day 1 performance was the only significant predictor retained in the model and showed that for every 1% increase in Day 1 performance on L1, there is a 0.48% increase in the magnitude of the Primacy effect on Day 2 (L1, Day 2 minus L2, Day 2; see Table 3).

We separated participants according to L1 performance on Day 1. Among the 32 participants did not score significantly greater than chance (scored 0.75 or below on 16 items, binomial test p<0.05), 15 were in the 1T condition on Day 2 and 17 in the 3T. Of the 1T group, the average Day 2 score was 0.635 (SD = 0.182) on L1 and 0.582 (SD = 0.237) on L2. For the 3T group, the average Day 2 score was 0.582 (SD = 0.140) on L1 and 0.520 (SD = 0.159) on L2. See Figure 3 for the results.

Among 33 participants scored significantly above chance (>0.75) on Day 1, 18 were in the 1T condition on Day 2 and 15 were in the 3T. Of the 1T group, the average Day 2 score was 0.776 (SD =0.166) on L1 and 0.558 (SD = 0.213) on L2. For the 3T group, the average Day 2 score was 0.786 (SD = 0.152) on L1 and 0.522 (SD = 0.161) on L2. See Figure 3 for the results.

Table 1: Mean Performances by Group

	1T	**3T**	**Combined**
Day 1, L1	0.741 (SD =0.224)	0.747 (SD =0.194)	0.744 (SD = 0.208)
Day 2, L1	0.717 (SD =0.184)	0.681 (SD =0.177)	0.694 (SD = 0.184)
Day 2, L2	0.568 (SD =0.220)	0.521 (SD =0.158)	0.540 (SD = 0.191)

Table 2: T-Test Score Against Chance

	1T	**3T**	**Combined**
Day 1, L1	6.175 (32) (p<0.001)	7.225 (31) (p<0.001)	9.458 (64) (p<0.001)
Day 2, L1	6.575 (30) (p<0.001)	5.677 (30) (p<0.001)	8.700 (61) (p<0.001)
Day 2, L2	1.734 (30) (p = 0.0932)	0.732 (30) (p=0.470)	1.836 (61) (p = 0.0712)

Table 3: Primacy Analysis

	Estimate	Std. Error	T-Value	P-Value
Intercept	-0.20	0.099	-2.07	0.042
L1, Day 1	0.48	0.13	3.77	0.00038

Multiple R^2	Adjusted R^2	F-Statistic	P-Value
0.19	0.18	14.2	0.00038

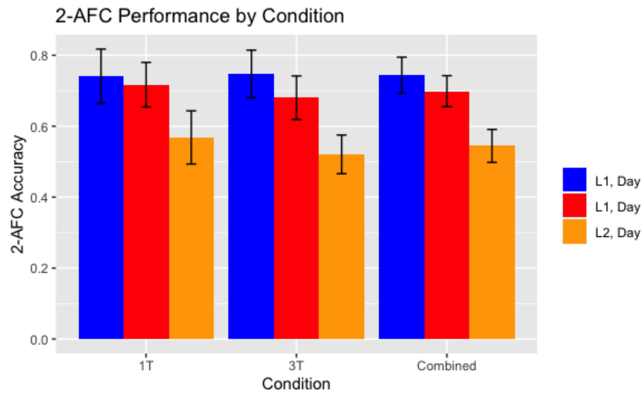

Figure 2: Accuracy of participant responses in the familiarity task. Chance level is at 0.50, and the error bars denote 95% confidence intervals of the mean.

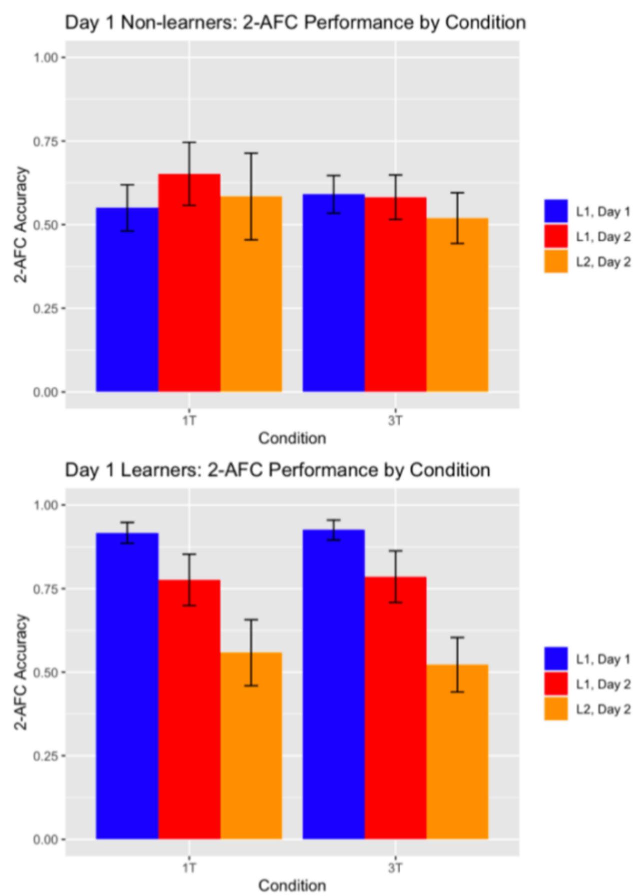

Figure 3: Accuracy of participant response in the familiarity task of the second day separated by their performance on the first day. Chance level is at 0.50, and the error bars denote 95% confidence intervals of the mean.

We repeated the stepwise regression procedure for each language (L1 and L2) on Day 2 to understand their specific contributions in the observed primacy effects. The L2 performance stepwise model only revealed L1, Day 2 performance as a significant predictor of L2 performance, with a 1% increase in L1, Day 2 performance associated with 0.54% increase in L2 performance. However, even the best model was not a good fit (F=1.96, p=0.078), with an adjusted R^2 of 0.098 (results in Table 4).

Table 4: L2 Performance Stepwise Analysis

	Estimate	Std. Error	T-Value	P-Value
Intercept	0.81	0.33	2.47	0.017
L1, Day 1	-0.64	0.35	-1.84	0.071
L1, Day 2	0.54	0.21	2.61	0.012
3T Switch	0.82	0.21	0.39	0.70
Second Language Proficiency	-0.76	0.46	-1.71	0.093

Multiple R^2	Adjusted R^2	F-Statistic	P-Value
0.20	0.099	1.96	0.078

The L1 performance stepwise model showed that performance on L1, Day 1 and L2, Day 2 each predicted variance on L1, Day 2. For each 1% increase in performance on L1, Day 1, there is a 0.46% increase in L1, Day 2, and for each 1% increase in performance on L2, Day 2, there is a 0.25% increase in L1, Day 2. The adjusted R^2 was 0.32 and (F=15.5, p<0.001; results in Table 5).

Table 5: L1, Day 2 Performance Stepwise Analysis

	Estimate	Std. Error	T-Value	P-Value
Intercept	0.22	0.090	2.41	0.019
L1, Day 1	0.46	0.091	5.04	<0.0001
L2, Day 2	0.25	0.099	2.51	0.015

Multiple R^2	Adjusted R^2	F-Statistic	P-Value
0.35	0.32	15.5	<0.0001

Although natural second language proficiency was not a significant predictor in any of the models, we asked whether any bivariate correlation existed between bilingual status and performance on the task because many previous studies have excluded bilingual participants. In comparing natural second language proficiency and L2 performance from the speech stream, the Pearson correlation coefficients were -0.224 and +0.082 for the 1T and 3T groups, respectively. However, neither of these correlations proved to be statistically significant (1T: p=0.23; 3T: p=0.66; see Figure 4).

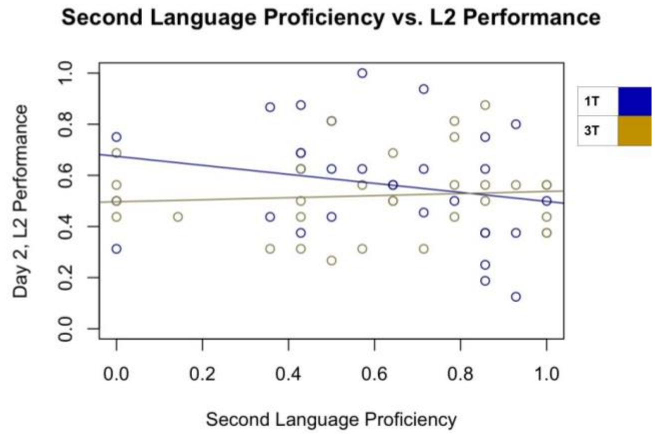

Figure 4: Natural second language proficiency compared with L2, Day 2 performance by condition.

Discussion

Our research was motivated by previous literature surrounding entrenchment in statistical learning of a language in the context of bilingualism. In presenting participants with an unstable speech stream of L1 and L2, we expected to see increased learning in L2, while maintaining learning in L1. The results did show that L1 learning was consistent across both conditions, but L2 learning also proved not to be significantly different between conditions, ultimately contradicting our hypothesis.

In line with an entrenchment or over-learning based explanation for the primacy effect, we also predicted that stronger learning in L1 on the initial exposure would correlate with decreased learning in L2. Based on our results, however, L2 was not acquired in the learned and unlearned groups by condition. The regression model for L2 was not adequate to draw conclusions, but Day 1 learning was a marginally significant and negative predictor of L2 performance on Day 2, which—if supported in future studies—would be consistent with our predictions.

In replicating the primacy measure from Zinszer & Weiss (2013), we found that only performance on L1, Day 1 affected the difference in performance between L1 and L2 on the second day. Primacy was unaffected by the switch condition. With L2 performance also not varying by condition, the results proved contradictory to our hypothesis.

Although this study and previous studies have used primacy as a measure of L2 learning after L1 exposure, primacy is also a function of varying L1 outcomes. Within the unlearned group, only participants in the 1T condition improved their L1 performance from Day 1 to Day 2, suggesting that the stable (1T) vs. unstable (3T) presentations did have an effect on their ability to learn the first language if coming in with little knowledge from Day 1. This pattern somewhat resembles the results of Zinszer & Weiss's (2013) Experiment 1, in which participants also had no prior knowledge of L1 (because there was only one day of testing) and primacy was reduced by largely through a decrease in L1.

In the learned group, participants showed less knowledge of L1 on Day 2 than they did on Day 1, but were still above chance. This group-level effect could be a simple case of regression to the mean (participants selected for doing especially well on Day 1 are more likely to decrease towards the overall group mean than become even more extreme relative to the mean). More importantly, in the stepwise regression model, any increase in an individuals' performance on L1, Day 1 or in L2, Day 2 corresponded to an increase in performance on L1, Day 2.

However, when looking at the influence of L2 performance on L1, Day 2, there are a few possibilities. One potential explanation is that there are individual SL ability differences, in that participants who perform well on the tasks will continue to perform well regardless of the condition. Another possibility relates to the efficiency hypothesis proposed by Karuza et al. (2016), which argues that participants engage in an efficiency versus change-detection tradeoff. According to this theory, participants who more strongly learned L1 reduced their sampling from the structure of L1 as the stream continued. However, this efficiency hampered the acquisition of L2 when it appeared. On the other hand, participants who were more active about learning the language were more prepared for the change to L2 that occurred in the stream. This provides an explanation for our findings as well, as those who pay attention to L1 on Day 2 would have better performance on L2 as well.

Ultimately, the results did not show that an unstable presentation of L1 and L2 could lead to stronger L2 learning. We did replicate the primacy effect, as L1 learning was significantly greater than L2 learning in both the 1T and 3T conditions, particularly for participants who showed evidence of L1 learning the day before. These data did not support our hypothesis that entrenchment would be overcome and primacy eliminated under the unstable speech stream. The basis of our hypothesis came from Zinszer & Weiss (2013), as in one of their experiments (Experiment 2) they exposed participants to a speech stream of the same exact presentation and length. The only difference between our 3T condition procedure and their Experiment 2 is we spaced ours over two days, with initial exposure to L1 on the first day instead of in an extended length at the beginning of the larger speech stream.

We spaced our speech stream out in such a manner for the same reason as Karuza et al. (2017): we wanted specifically to focus on how having learned L1 impacts L2 learning instead of how L1 and L2 can be learned at the same time. According to Gebhart et al. (2009), a pause between two speech streams is enough to cue participants towards a potential change in the stream, which (if anything) we would have expected to improve L2 learning and decrease primacy overall, and so we must explore the effect of the 24 hour intermission in our own design. Under our model of unstable transitions, we provide evidence against the native language neural commitment theory, as we find that L1 learning increases with L2 learning in a positive relationship. This indicates that in artificial multiple language learning, participants are hindered from learning L2 as a result of another phenomenon besides overcommitment to L1.

In the future, we will manipulate the speech streams further, eliminating the entrenchment of L1 by either replacing L1 on Day 1 with a third language separate from the Day 2 stream, or by eliminating Day 1 altogether and shortening the presentation of L1 on Day 2. This way we can test for the presence of cue to multiple structures from the 24 hour pause. We will also investigate the use of the EEG as an attention check, to see if individuals who display more attention throughout the speech streams are more likely to detect changes from L1 to L2.

Acknowledgements

We thank the Department of Psychology at Swarthmore College for their support, John Duarte and Chris Zhao for help with data collection, and Dr. Andrea L. Gebhart, Dr. Richard N. Aslin, and Dr. Elissa L. Newport for allowing us (and many other researchers) to use the artificial languages from their study. This work was supported by the Eugene M. Lang Summer Fellowship to J.C. and the Swarthmore College Faculty Research Fund to B.Z.

References

Antovich, D. M., & Graf Estes, K. (2017). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*, 125–148. https://doi.org/10.1111/desc.12548

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36(2),* 286–304. https://doi.org/10.1111/j.15516709.2011.01200.x

Bulgarelli, F., & Weiss, D. J. (2016). Anchors aweigh: The impact of overlearning on entrenchment effects in statistical learning. *Journal of Experimental Psychology: Learning, Memory & Cognition, 42(10),* 1621–1631. https://doi.org/10.1037/xlm0000263

Callan, D. E., Jones, J. A., Callan, A. M. & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage, 22,* 1182–1194.

Dehaene-Lambertz, G., Dupoux, E. & Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *J. Cogn. Neurosci., 12,* 635–647.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *American Speech-Language-Hearing*

Association, 52(2), 321–335. https://doi.org/10.1044/1092-4388(2009/07-0189)

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009), Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cognitive Science, 33*, 1087-1116. https://doi.org/10.1111/j.1551-6709.2009.01041.x

Karuza, E. A., Li, P., Weiss, D. J., Bulgarelli, F., Zinszer, B., & Aslin, R. N. (2016). Sampling over non-uniform distributions: A neural efficiency account of the primacy effect in statistical learning. *J. Cogn. Neurosci, 28(10),* 1484-1500. https://doi.org/10.1162/jocn_a_00990

Kuhl, P. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843. https://doi.org/10.1038/nrn1533

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology, 42(4),* 643–655. https://doi.org/10.1037/0012-1649.42.4.643

Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in Psychology, 3,* 1–9. https://doi.org/10.3389/fpsyg.2012.00228

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development, 80(3),* 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928. http://dx.doi.org/10.1126/science.274.5294.1926

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. https://doi.org/10.1080/15475440802340101

Zinszer, B. D., & Weiss, D. J. (2013). When to hold and when to fold: Detecting structural changes in statistical learning. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35thAnnual Conference of the Cognitive Science Society* (pp. 3858–3863). Austin, TX: Cognitive Science Society.