The Role of Native Language in Statistical Learning Success

Krystal Mendez, An Nguyen, Violet Kozloff, Zhenghan Qi

Language Acquisition and Brain Laboratory, Department of Linguistics and Cognitive Science, University of Delaware, Newark, DE, USA

Introduction

Statistical learning is the ability to extract repeated patterns of regularities and transitional probabilities. We examined whether native language experiences (English vs. Hebrew), perceived familiarity with one's native language, or native language proficiency affects how successful an individual learns an artificial language that is composed of Hebrew syllables.

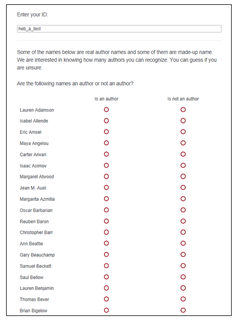
The aim of this study was to confirm these hypotheses:

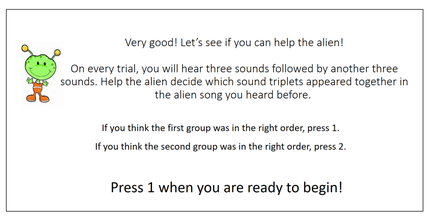
1. English native speakers will perform relatively worse than Hebrew native speakers in learning the artificial language, but equally well in the non-linguistic statistical learning task.
2. There will be a positive correlation between the perceived English likeness of the artificial language and the learning success of an artificial language.
3. People with better verbal knowledge and reading experiences will show greater success in both statistical learning tasks.

Method

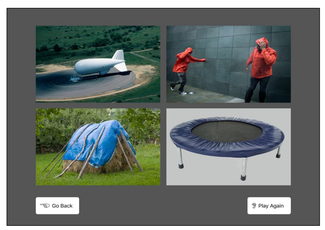
32 adults participated in the study (mean age 21.9 years old, 26 females and 6 males). All were between the ages of 18 and 40, receiving payment for their participation. They were all native English speakers with no learning, hearing, or language impairments.

Experiment Procedure:





- Linguistic ASL
- Non-linguistic ASL
- Author task
- Perceived English likeness
- Picture-Vocabulary task



Results

Table 1: Mean accuracy of statistical learning performance in the current study and Arnon (2018) (SD).

Task	Hebrew native speakers	English native speakers
Linguistic ASL	71% (13%)	57% (13%)
Non-linguistic ASL	71% (15%)	62% (10%)

Table 2: Inferential statistics of English native speakers' performance compared to chance and Hebrew native speakers.

	Compare to chance (50%)	Compare to Arnon (2018)
Linguistic ASL	t(30) = 3.00, p = 0.02	t(81) = 4.75, p = 0.02
Non-linguistic ASL	t(28) = 6.46, p = 0.02	t(79) = 2.89, p = 0.03

Table 3: Correlation matrix showing all individual difference measures in native English speakers (p-value marked in asterisks)

	Linguistic ASL	Non-Linguistic ASL	Author	Vocab	English likeness
Linguistic ASL	1	**	**	**	***
Non-linguistic ASL	0.488	1			
Author	0.486	0.277	1	**	
Vocab	0.462	0.043	0.491	1	
English likeness	0.577	0.117	0.155	0.093	1

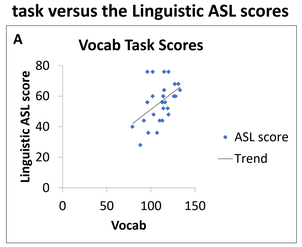
* p < 0.05;** p< 0.01; *** p < .001

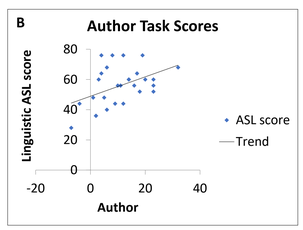
Table 4: Multiple linear regression model predicting Linguistic ASL performance

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	13.58186125	16.20041531	0.838365	0.410098
Vocab	0.286127524	0.155241692	1.84311	0.077698
Author	0.376361066	0.232744313	1.617058	0.118936
Rating of English likeness of the ASL	5.333741243	1.637855687	3.256539	0.003348

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	2285.086027	761.6953	8.152496	0.000646494
Residual	24	2242.342544	93.43094		
Total	27	4527.428571			

Figure 1: Scatter plots of A) vocabulary task scores, B) author task scores, and the C) rating of English likeness of the linguistic task versus the Linguistic ASL scores





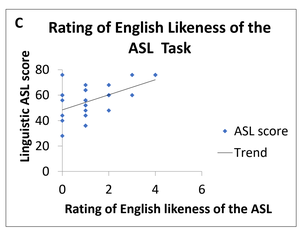
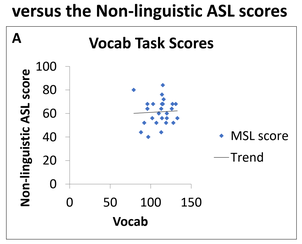


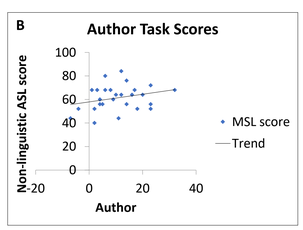
Table 5: Multiple linear regression model predicting Non-linguistic ASL performance

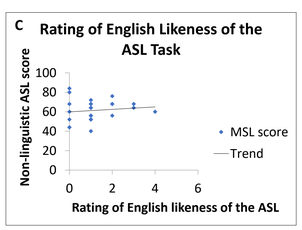
	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	68.08948625	18.60988889	3.65878	0.001307
Vocab	-0.115239837	0.181510002	-0.6349	0.531762
Author	0.408883704	0.270554497	1.51128	0.144336
Rating of English likeness of the ASL	1.46247008	2.102891871	0.695457	0.493737

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	314.0084569	104.6695	0.893618	0.459346258
Residual	23	2693.991543	117.1301		
Total	26	3008			

Figure 2: Scatter plots of A) vocabulary scores, B) author task scores, and the C) rating of English likeness of the linguistic task versus the Non-linguistic ASL scores







Results (continued)

- Table 1 shows the means of English vs. Hebrew native speakers. The English native speakers performance was worse than that of the Hebrew native speakers in both tasks.
- Table 2 shows that English native speakers performed significantly above chance (50%) for both the linguistic and non-linguistic tasks. However, they preformed extensively worse than the Hebrew participants. Without the raw data from the Arnon (2018), it is not possible to determine whether the group difference is greater in linguistic than the non-linguistic task.
- Table 3 shows the correlation matrix of all the individual difference measures.
- Table 4 & Figure 1 show the multiple linear regression of the linguistic data. For A) there a marginal contribution to the task, B) shows no significant relationship between the linguistic asl score and author task, and C) has a significant and unique data trend
- Table 5 & Figure 2 shows the multiple linear regression of the non-linguistic data. For A), B), and C) there is no relation to the individual variabilities of statistical learning performance. The data trends show no relationship between the non-linguistic scores and the vocab scores, author task scores, or rating of English likeness.

Conclusion

1. Familiarity (or perceived familiarity) of language is related to statistical learning success from the speech input.
   - English native speakers perform worse than Hebrew native speakers in learning artificial language consisting of Hebrew syllables.
   - Perceived English likeness of the artificial language is associated with linguistic statistical learning success in English native speakers.
2. Individuals' vocabulary is only associated to linguistic statistical learning performance, but not the non-linguistic statistical learning performance, suggesting dissociable learning mechanisms between linguistic and non-linguistic domains.

References

1. Arnon, i. (2018, May 7). Do current statistical learning capture stable individual differences in children? An investigation of task reliability across modalities. https://doi.org/10.31234/osf.io/9pa8t
2. Romberg, A. and Saffran, J. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), pp.906-914.
3. Saffran, J., Aslin, R., & Newport, E. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926-1928. doi: 10.1126/science.274.5294.1926

Acknowledgements

Thanks to the members of the Language Acquisition and Brain Laboratory for their support in data collection and analysis. I would also like to acknowledge Dr. Inbal Arnon for generously sharing her stimuli for this project.

This poster was supported by the Delaware INBRE program, with a grant from the National Institute of General Medicine Sciences – NIGMS (8 P20 GM103446-16) from the National Institutes of Health.