WHAT CAN A WUG TEST TELL US ABOUT ABSTRACT MORPHEME STORAGE?

Shiloh Drake[1]

1 – Department of Linguistics, University of Arizona
sndrake@email.arizona.edu

ABSTRACT

Wug Tests can be used to probe morphological knowledge, from the stages of morphological development in the classic Wug Test [1], to the productivity of morphemes in a human language [6, 21], to testing the acquisition of an artificial grammar [7, 9, 22]. The present study tested three speaker groups with an artificial grammar learning task, and an error analysis provided further evidence for the storage and processing of abstract morphemes in speakers of Maltese and Arabic. Errors from the three groups seem to be based on acceptable abstract morphemes and/or phonological structures found in the native language of participants.

Keywords: wug test, root and pattern morphology, Arabic, Maltese, abstract morphological representation

Acknowledgements: I wish to thank Lauren Ackerman, Elise Bell, Megan Figueroa, Heidi Harley, Diane Ohala, Adam Ussishkin, and Mary-Caitlyn Valentinsson for productive discussions about this data. All errors remain my own.

1. INTRODUCTION

Wug Tests, or novel word elicitation tasks, can be used to probe various forms of metalinguistic knowledge related to both lexical structure and morphological restrictions. In the classic Wug Test, Berko [1] provided evidence for the development of morphological understanding. Later research has used a Wug Test to observe morphological variation and productivity [6, 21] and to assess learning in an artificial grammar [7, 9, 22]. In this study, I use results from a Wug Test to provide further evidence for the presence of an abstract morpheme in the grammars of Arabic and Maltese speakers, and suggest that this might have an effect on morphosyntactic borrowing processes through language contact.

This study was conducted to examine whether an abstract consonant-vowel (CV) skeleton morpheme is stored in a speaker's lexicon, and whether this would affect how an artificial language was learned by three different experimental populations. With additional research, it can inform a wider community of how to best teach novel components of a foreign language, and how the diversity of unconscious expectations that humans have from processing their native language affects how they process a new language.

1.1. Semitic Morphology

Semitic languages, which include Arabic, Hebrew, Maltese, Amharic, and others, are known for utilizing non-concatenative morphology to form words. The particular type of non-concatenative morphology that these languages use is known as root and pattern morphology or templatic morphology to distinguish from other forms of non-concatenative morphology like reduplication. Most Semitic words consist of a 2-4 consonant *root*, which provides a broad lexical meaning, and a *pattern* made up of consonants and vowels interleaved with the root, providing grammatical and prosodic information. For example, the Maltese root *gdm* can be combined with multiple patterns to slightly alter the overall meaning of 'bite': *gdim* 'bite.n', *gidem* 'bite.v.citation', *nigdem* 'I bite'. Contrast this with concatenative morphology, which is the primary word formation system in English and can also be observed in Maltese: *teżor -i* 'treasures', *ip- park -ja* 'to park', *ip-parteċipa -w* 'they are participating'. With neither the root nor the pattern completely contiguous in words formed with root and pattern morphology, they constitute non-adjacent dependencies. Non-adjacent dependencies have been posited as difficult to learn or parse by many researchers (e.g., [2, 11, 14]) in spite of their typological prevalence in phonology, morphology, and syntax.

Researchers have also suggested the existence of abstract morphemes, such as CV skeleta, in speakers of Semitic languages ([3] on Arabic; [10] on Maltese) as one of the components of lexical entries. Further studies have also suggested that the presence of a CV skeleton-type morpheme might influence the computational learnability of non-concatenative broken plurals in

Arabic [5] and Maltese [8, 15]. Thus, it is reasonable to believe that the existence of this CV skeleton morpheme might affect how participants learn and internalize novel words with similar morphological structures.

1.2. Structural Borrowing

Authors in sociolinguistic frameworks are able to show that borrowing of grammatical structures occurs when there is substantial contact between two languages and in highly proficient bilingual populations [4, 12, 13]. This is shown mainly through analyses of codeswitching and codemixing, and occurs at both the morphological and syntactic levels of language. Language dominance is proposed to be one of the driving factors in this type of borrowing [4], and this type of borrowing is suggested to be one of the stages of a shift in the population from a minority language to a majority language [4].

With the Maltese youths' sentiment of English being an important language for communication around the world [20], the Arabic-speaking participants in this study living in a highly English-dominant environment, and the prevalence of English as a *lingua franca* around the world, it is reasonable to expect that English morphemes might be borrowed into Maltese and Arabic. However, participants should be sensitive to both English and Maltese/Arabic structures, since they are highly proficient in both languages. This potential fluidity of language grammars, linguistic proficiency, and the structures mimicked by the artificial grammar should lead participants to show some of the more abstract parts of their mental grammars by way of the errors they produce in this artificial grammar learning task.

2. METHODS

2.1. Participants

Forty-two monolingual American English speakers (29 female; age 18-40, mean: 20.875), 38 native Maltese speakers (23 female; age 18-35, mean: 21.45), and 20 native Arabic speakers (5 female; age 18-38, mean: 26.1) participated in the task. Participants were randomly assigned to either the concatenative or non-concatenative grammar, and were compensated with course credit, $5, or €5.

2.2. Materials

Eighty phonotactically legal but nonexistent roots, 3 nonexistent patterns, and 3 nonexistent suffixes were used to construct the artificial grammars. All phonemes were common to English, Modern Standard Arabic, and Maltese. Roots and patterns, or roots, random vowels, and affixes were combined to create 80 singular-plural pairs, 10 of which were used during the training phase, and 70 of which were used during the test phase. Each word pair was randomly paired with the image of an imaginary animal (originally appearing in Ohala [16, 17]). All words were presented in the Latin alphabet, or in the case of the Maltese speakers, the modified version of the Latin alphabet used in Malta. Tables 1 and 2 provide examples of the words that participants saw during the experiment. For a complete description of the artificial grammar construction, see Drake [7].

The three morphemes in each grammar were used to mimic morphological variation, and were therefore presented with different degrees of frequency. The most frequent morpheme appeared roughly 60% of the time in both training and test phases, while the other two morphemes each appeared roughly 20% of the time. This variation was based on the phonological class of the second root consonant to ensure consistency. To find out whether participants had consciously observed a pattern, they were asked after completing the experiment whether they felt like they knew when to use each of the possible word formations. Participants reported not knowing how to tell which morpheme was appropriate to use with each word.

Table 1: Example words from the concatenative grammar.

Root	Singular	Affix	Plural
BKD	bikid	-uk	bikiduk
MNT	munut	-in	munutin
LFK	lifuk	-af	lifukaf

Table 2: Example words from the non-concatenative grammar.

Root	Singular	Pattern	Plural
BKD	bikidi	$C_1aC_2C_3u$	bakdu
MNT	munuta	$C_1C_2iC_3a$	mnita
LFK	lifuka	$C_1uC_2C_3iC_3$	lufkik

2.3. Procedure

The task was administered in a quiet room using the PsychoPy2 software [18, 19]. The experiment was self-paced and consisted of a training phase and a test phase. The training phase provided participants with information on how to form the plural in the "alien language" by presenting participants with the image of one animal and the singular form, then a picture of two

animals and the plural form. No words were repeated, nor were participants allowed to go back to view items again. During the test phase, participants were presented with two images of the same animal and the singular form of a word, and were asked to type the plural form using an American QWERTY keyboard. No words were repeated from the training phase. The entire procedure took approximately 20 minutes for each participant.

3. RESULTS

Participant responses were hand-coded for response type. Responses typically fell into 5 categories: correct response, overregularization to one of the three morphemes, adding a Maltese-like, Arabic-like, or English-like plural suffix, changing vowels from the singular to plural but nothing else, or using the correct syllable or CV structure but not the correct vowels.

Participants learning the concatenative grammar had the highest proportion of correct responses, at about 25% in each group. All participants in the concatenative grammar were most likely to use the most common plural allomorph, analogous to the plural *-s* in English.

Participants learning the non-concatenative grammar had a very low percentage of correct responses, even for the two speaker populations that have productive non-concatenative morphology in their native language. Percentage of correct responses ranged from 0.5% to 5.5% correct.

Accuracy differences were not statistically significant between groups according to a linear mixed effects model ($\chi^2(1) = 2.9$, $p > .05$), but were significant between grammars ($\chi^2(1) = 48.0$, $p > .001$). The model included native language (Arabic, English, Maltese) and grammar type (concatenative, non-concatenative) as fixed effects, and subjects and items as random effects. The full model specification can be found in the Supplementary Materials.

Table 3: Results for fixed effects.

Predictor	Estimate	SE	t Value	Pr(> |z|)
(intercept)	.70	.04	16.29	
English	.05	.05	1.04	.30
Maltese	.08	.05	1.17	.10
NC Gram.	-.27	.04	-7.67	>.01

Table 4: Goodness-of-fit measures for main model.

AIC	BIC	log likelihood	Deviance
-5113.3	-5065.4	2563.6	-5127.3

However, the types of errors committed differed based on the native language of the participants, summarized in Tables 5-7 in the Supplementary Materials. Native Arabic speakers and native English speakers were most likely to change vowels and nothing else in the non-concatenative grammar, while native Maltese speakers were most likely to use a Maltese-like concatenative plural (*-ijiet* or *-i*), overregularize to a word-initial CC cluster, and then to change the vowels and nothing else. Roughly 15% of the Arabic speakers' responses were either English-like (*-s*) or Arabic-like (*-oon*, *-aat*, or *-iin*) concatenative suffixes, and about 6% of the English speakers' responses were the English concatenative suffix.

4. DISCUSSION

Error types differed based on the native language of the participants. This is expected based on the design of the experiment: a very short exposure to the artificial grammar should result in participants relying on their previous linguistic and morphological knowledge to make hypotheses about the correct form of a word in a novel language. The results from this experiment show that a CV skeleton, or possibly another abstract morpheme, is one of the components used by participants at a very early stage in learning a new language.

Previous research suggests that Arabic and Maltese speakers utilize an abstract CV skeleton as an additional component of word formation [3, 10]. The present research adds to this theory, as even the overregularizations produced by the Maltese speakers and Arabic speakers were qualitatively different. Maltese speakers were the only group that was more likely to overregularize to the plural form with a word-initial consonant cluster. This occurred in some cases even when the consonant cluster is not phonotactically licit in Maltese, which suggests that adhering to the CV skeleton in word formation is a very strong bias and can override phonotactic constraints.

Native Arabic speakers were overwhelmingly more likely than the other groups to change vowels and/or syllable structure, which again can be exemplified by singular/plural changes in Arabic. For instance, *kitaab* 'book' changes to *kutub* in the plural, and *Sadiiq* 'friend' changes to *'aSdiqa'* in the plural.

Further, in both languages, CV skeleta of singular forms can be used to predict plural forms [5, 8, 15]. In Arabic, for example, singular words with the skeleton CVCCVVC are likely to take a plural with the skeleton

CVCVVCV(V)C: *miftaaH* ‘key’ → *mafaatiH* ‘keys’; *sikkiin* ‘knife’ → *sakaakiin* ‘knives’.

In a language like Maltese with considerable structural variation, some type of abstract morpheme might contribute to the selection among concatenative and non-concatenative morphemes. In the present study, the nonsense word *binixa* received possible plural responses as both *binixi*, the regular concatenative Maltese plural, and *bnienex*, which is a common non-concatenative structure in Maltese analogous to *ġakketta* ‘jacket’ → *ġkieket* ‘jackets’. Further, another study [8] elicited both *kxejjex* and *kaxxi* as plurals to *kaxxa* ‘box’, *lpup* and *lupi* as plurals to *lupu* ‘wolf’, and *vlup* and *volpijiet* as plurals to *volpi* ‘fox’. *Kaxxi*, *lpup*, and *volpijiet* are the attested dictionary plurals of those particular words, yet speakers not only accept but also generate non-concatenative (or concatenative) plurals for each of them.

Given these results and the context of previous literature, it is necessary to conduct more research on how an abstract morpheme might be stored and applied productively. Not only would this add to our knowledge of the nature of the mental lexicon in speakers of Maltese and Arabic, but also could lead to further hypotheses of how this type of morphological processing interacts with borrowing and mixing in highly proficient and multilingual communities. In situations of language contact resulting in structural borrowing processes, both new and established structures may be used variably among speakers [4, 12, 13]. Future research in this area can be used to explain how language processing constraints in tandem with language contact affect both morphological change and morphological variation.

REFERENCES

1. Berko, J. (1958). The child’s learning of English morphology. *Word*, 150–177.
2. Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science, 16*(6), 451–459.
3. Boudelaa, S. & Marslen-Wilson, W. D. (2004). Abstract morphemes and lexical representation: The CV-skeleton in Arabic. *Cognition, 92*, 271–303. doi: 10.1016/j.cognition.2003.08.003
4. Brown, B. (2003). Code-convergent borrowing in Louisiana French. *Journal of Sociolinguistics*, 7(1), 3–23.
5. Dawdy-Hesterberg, L. G. & Pierrehumbert, J. B. (2014). Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience, 29*(10), 1268–1282. doi: 10.1080/23273798.2014.899377
6. Drake, S. (2018). The form and productivity of the Maltese morphological diminutive. *Morphology*, 28(3), 297–323. doi: 10.1007/s11525-018-9328-0
7. Drake, S. N. (2018). *L1 biases in learning root-and-pattern morphology* (Ph.D. Dissertation). The University of Arizona, Tucson, AZ.
8. Drake, S. & Sharp, R. (2017). Productivity of the broken plural in Maltese. Presented at GĦILM6, Bratislava, SK.
9. Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. C. M. (2016). The relationship between artificial and second language learning. *Cognitive Science, 40*, 822–847. doi: 10.1111/cogs.12257
10. Galea, L. (2011). *Mental Representations of Maltese Semitic Verbs* (Master’s Thesis). University of Essex, Essex, UK.
11. Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431–436.
12. Myers-Scotton, C. (1993). *Duelling Languages: Grammatical Structure in Codeswitching*. Clarendon Press, Oxford, UK.
13. Myers-Scotton, C. (1998). A way to dusty death: The Matrix Language Turnover Hypothesis. In L. A. Grenoble & L. J. Whaley (eds.). *Endangered Languages: Language Loss and Community Response*. Cambridge University Press, Cambridge, UK. pp 289–316.
14. Newport, E. L. & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.
15. Nieder, J. & van de Vijver, R. (2017). Investigating the plural formation in Maltese with naive discriminative learning. Presented at GĦILM6, Bratislava, SK.
16. Ohala, D. K. (1996). *Cluster reduction and constraints in acquisition* (Ph.D. Dissertation). The University of Arizona, Tucson, AZ.
17. Ohala, D. K. (1999). The influence of sonority on children’s cluster reductions. *Journal of Communication Disorders, 32*, 397–422.
18. Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8–13.
19. Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*(10). doi: 10.3389/neuro.11.010.2008
20. Sciriha, L. (2017). The visibility and salience of Maltese in bilingual Malta’s linguistic landscape. In B. Saade & M. Tosco (eds.). *Advances in Maltese Linguistics*. De Gruyter Mouton, Berlin/Boston. pp 225–244.
21. Twist, A. E. (2006). *A psycholinguistic investigation of the verbal morphology of Maltese* (Ph.D. Dissertation). The University of Arizona, Tucson, AZ.
22. Wong, P. C. M., Ettlinger, M., Zheng, J. (2013). Linguistic Grammar Learning and DRD2-TAQ-IA Polymorphism. *PLOS ONE, 8*, e64983. doi: 10.1371/journal.pone.0064983