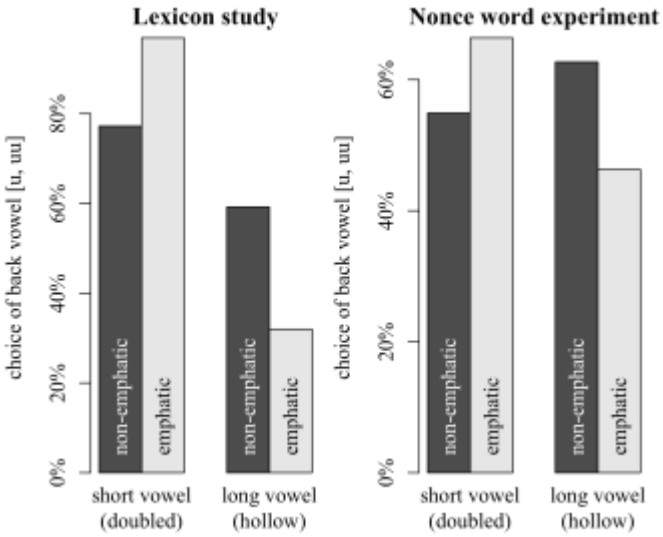
The predictability of vowel alternations in Urban Hijazi Arabic imperfective nonce forms

In Urban Hijazi Arabic verbs that have two consonants on the surface (e.g. [tˤaar] "he flew", [tˤaχ] "he shot"), the vowel is always low in the perfective. In the imperfective, the vowel can be the front [i, ii] or the back [u, uu] (also [aa] in a few verbs). We found that the choice of vowel is partially predictable: the presence of an emphatic (pharyngealized consonant) is correlated with [ii] in "hollow" verbs (long vowel verbs), but the presence of an emphatic is correlated with [u] in "doubled" verbs (short vowel verbs). This fact is true in the existing verbs of the language, and is also apparent in the treatment of nonce verbs, as we establish here experimentally. Existing analyses of hollow and doubled verbs (Brame 1970, Gafos 2003, Shekeri & Scheer 2004, 2005, Rosenthall 2006) assume that the vowel choice depends on a lexical representation, and therefore these analyses do not make predictions about the vowel choice in nonce words. We offer an analysis based on the Minimal Generalization Learner (Albright & Hayes 2002, 2003, 2006) that learns the distribution of vowels in the lexicon and extends them to nonce words. The predictions of our analysis are highly correlated with the experimental results.

Lexicon Study: We compiled an Urban Hijazi Arabic lexicon of 101 long vowel verbs and 133 short vowel verbs, of which about a quarter contain an emphatic. We see that the back vowel is chosen *more* often in the presence of emphatics with short vowel verbs, but *less* often in the presence of emphatics with long vowel verbs.

	emphatic (pharyngealized)			non-emphatic (plain)		
doubled verbs (short vowel)	dˤal ~ ji-dˤil tˤaχ ~ ji- tˤuχ	"to get lost" "to shoot"	97% [u]	ʕal ~ ji-ʕil baχ ~ ji-buχ	"to sicken" "to spray"	77% [u]
hollow verbs (long vowel)	tˤaar ~ ji-tˤiir sˤaam ~ ji-sˤuum	"to fly" "to fast"	32% [uu]	ʕaad ~ ji-ʕiid ʃaaf ~ ji-ʃuuf	"to repeat" "to see"	59% [uu]

Experiment: 104 native speakers of Urban Hijazi Arabic listened to 28 nonce words ("wug test", Berko 1958) put in the perfective, each with two choices for the imperfective (e.g. [tˤak] → [ji-tˤik, ji-tˤuk]). Participants chose vowels that follow the lexical distribution of the emphatics, as seen in the figure, and confirmed with high significance in a fully crossed mixed-effects logistic regression model.



Learning Model: We trained the Minimal Generalization Learner (Albright & Hayes 2002, 2003, 2006) on the real words of the language, and tested it on the same nonce perfective verbs that were given to the participants. The model generated imperfective forms and confidence scores that were strongly correlated with the choices of the native speakers.

Discussion: Speakers have implicit knowledge about the choice of vowel in the imperfective based on the vowel's length and the presence of an emphatic. The trends do not seem phonetically motivated; rather, speakers simply extend the distribution of vowels from their lexicon, as predicted by the Minimal Generalization Learner. Additionally, our results lend support for the view that speakers of Semitic languages can learn trends in their lexicon at a level of representation in which both vowels and consonants are present, not purely consonantal roots (as argued by Gafos 2003, Berent et al. 2007 a.o.).

References

Albright, Adam & Hayes, Bruce. 2002. Modeling English past tense intuitions with minimal generalization. In M. Maxwell (Ed.) *Proceedings of the Sixth Meeting of ACL Special Interest Group in Computational Phonology* (pp. 58-69). Philadelphia: ACL.

Albright, Adam & Hayes, Bruce. 2002. Modeling English past tense intuitions with minimal generalization. In M. Maxwell (Ed.) *Proceedings of the Sixth Meeting of ACL Special Interest Group in Computational Phonology* (pp. 58-69). Philadelphia: ACL.

Albright, Adam, and Bruce Hayes. 2006. Modeling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations. In *Gradience in grammar: Generative perspectives*, eds. Gisbert Fanselow, Caroline Fery, Matthias Schlesewsky, and Ralf Vogel, 185–204. Oxford: Oxford University Press.

Berent, Iris, Vaknin, Vered, & Marcus, Gary F. 2007. Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition* 104:254–286.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14: 150–177.

Brame, Michael. 1970. Arabic phonology: Implications for phonological theory and general Semitic. Doctoral Dissertation, MIT.

Chekayri, Abdellah & Scheer, Tobias. 2004. The non-appearance of glides in the hollow verbs of Classical Arabic. *Paper presented at the first conference on Arabic language and linguistics*, Oxford 30-31 July.

Chekayri, Abdellah & Scheer, Tobias. 2005. Biliteral approach to weak verbs in Arabic. *Paper presented at the 19th Arabic Linguistics Symposium*, University of Illinois 1-3 April.

Gafos, Adamantios. 2003. Greenberg's asymmetry in Arabic: A consequence of stems in paradigms. *Language 79 (2),* pp. 317-355.

Rosenthall, Sam 2006. Glide distribution in Clasical Arabic verb stems. *Linguistic Inquiry* 37, pp. 405-440.

(Phonology, Morphology, language learning)