Jennifer Kuo. AFLA 27 at NUS, 2020

Evidence for base-driven alternation in Tgdaya Seediq

Standard approaches to learning the phonology of inflectional paradigms require positing URs from which surface contrasts can be derived (Chomsky and Halle, 1968). In many cases, URs are 'cobbled', meaning that they combine information from multiple forms of the paradigm, and do not not correspond to an existing surface form. In contrast, Albright (2002, et seq.) argues for a **surface base approach**, where learners must base the UR on a single existing surface form. In this study, we present evidence from Tgdaya Seediq in support of the Albrightian surface base hypothesis. In particular, a survey of the Seediq lexicon reveals asymmetries which cannot be explained under the traditional UR analysis, but are predicted under a surface-base model.

In Seediq, processes of vowel reduction and final consonant neutralization result in the neutralization of contrasts in both the suffixed and non-suffixed forms of verb paradigms.

Seediq vowel reduction depends on stress, which is always penultimate. Pretonically, vowels always reduce to [u], such as in example (1a). This results in neutralization of contrasts in the suffixed forms. For example, the two stems in (1a) and (1b) are contrastive in the non-suffixed stems, but become homophonous when suffixed with /-an/.

Post-tonically, a similar but more restricted process occurs, where /e, o, u/ reduce to [u]. This results in neutralization of vowel contrasts in the non-suffixed forms of a paradigm. For example, in (1c) and (1d), the contrast between the final vowel of the stem is lost in the isolation stem.

(1) *Examples of vowel reduction*

	STEM	SUFFIXED			STEM	SUFFIXED	
a.	ˈbarah	buˈrahan	'reduce'	c.	ˈcebuw	cuˈbuwan	'shoot'
b.	ˈberah	buˈrahan	'advance'	d.	ˈrebuw	ruˈbewan	'soak'

Neutralization in the non-suffixed forms also arises from extensive final consonant neutralization. For example, /p, b, k/ neutralize to [k] word-finally. As a result, as shown in (2), stem-final [k] could surface as [k], [p], or [b] in the suffixed form of a paradigm. Other processes of final consonant neutralization, listed in (3), lead to similar alternations for stem-final [ŋ, t͡s, n].

(2) *Examples:* /p, b, k/ → [k] *neutralization*

	STEM	SUFFIXED	
a.	ˈpiyuk	puˈyupan	'blow'
b.	ˈgemuk	guˈmekan	'cover'
c.	ˈpeluk	puˈleban	'close'

(3) *Final consonant neutralizations*

- a. /p, b, k/ → [k]
- b. /m, ŋ/ → [ŋ]
- c. /d, t, t͡s/ → [t͡s]
- d. /l, n/ → [n]

Due to these processes, all forms of a paradigm suffer from some type of neutralization. Prior analyses of Seediq have accounted for this extensive neutralization by forming 'cobbled' URs (Yang, 1976). For a verb such as *rebuw~rubewan* in (1d), the stem UR would get its first vowel from the non-suffixed form, and its second vowel from the suffixed from. The resulting UR, /rebew/, can then be used to derive the surface forms via rules for pre- and post-tonic vowel reduction.

However, asymmetries in the Seediq lexicon support the alternative Albrightian surface-base approach. In the current study, we conducted a quantitative analysis of 342 verb paradigms, collected from online databases and fieldwork. Results of this analysis suggest that suffixed forms are highly predictable from the non-suffixed forms (e.g. the isolation stem) of the paradigm.

For example, because of vowel reduction, the post-tonic [u] in a stem of the form CVCuC surfaces as [e], [o], or [u] when it is stressed in the suffixed form. Crucially, which vowel surfaces
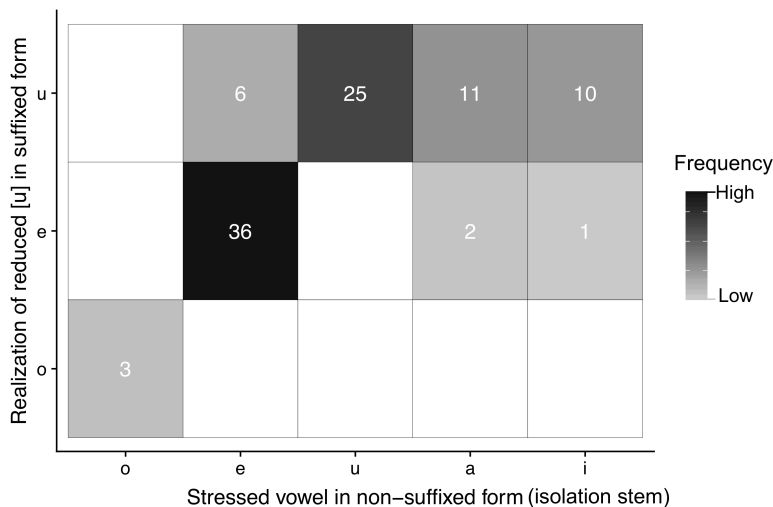
Figure 1: Frequency counts of CVCuC stems
(White boxes indicate a frequency of zero)

in the suffixed forms is actually strongly correlated with the identity of the stressed vowel in the non-suffixed stem. This is demonstrated in Fig. 1, which summarizes the distribution of CVCuC stems. If the stem stressed vowel is [o], the reduced vowel always surfaces as [o] in the suffixed form. Similarly, if the stem stressed vowel is /e/, the reduced vowel tends to surface as [e] in suffixed forms. Otherwise, the reduced vowel usually surfaces as [u].

As a result of these statistical regularities, whether or not the stem final vowel alternates is largely predictable from just the isolation stem. Although not discussed here, similar asymmetries were found for the final consonant alternations. In other words, whether or not a stem-final consonant alternates is predictable to a large degree from the isolation stem.

These findings were confirmed in a computational model which predicted the inflected forms of 342 Seediq verbs. This rule-based model, based on the Minimal Generalization Learner (Albright and Hayes, 2003), takes a surface variant as a base and uses it to predict other forms of the paradigm.

Two separate models were built, respectively using the non-suffixed and suffixed forms as the base. Each model was evaluated by the **scope** and **accuracy** of its rules, where scope is the number of forms a rule is applicable to, and accuracy is the number of forms for which the rule derives the correct output. Modelling results showed that the isolation stem (which is non-suffixed) correctly predicts the other forms of the paradigm at a higher rate than other forms. This asymmetry cannot be easily explained under the view that learners establish URs using all available surface allomorphs. On the other hand, it falls out naturally from a surface-base approach.

References

Adam Albright. *The identification of bases in morphological paradigms*. PhD thesis, University of California, Los Angeles, 2002.

Adam Albright and Bruce Hayes. Rules vs. analogy in english past tenses: A computational/experimental study. *Cognition*, 90(2):119–161, 2003.

Noam Chomsky and Morris Halle. *The sound pattern of English*. ERIC, 1968.

Hsiu-fang Yang. The phonological structure of the paran dialect of sediq. *Bulletin of the Institute of History and Philology Academia Sinica*, 47(4):611–706, 1976.