Learning The Categories Underlying Turkish Voicing Alternations

Stem-final Turkish stops alternate between voiced before vowel-initial suffixes and voiceless elsewhere (or [k] ~ ∅) for some **(1a)** but not all **(1b)** nouns. Elicitations of the accusative of known nouns showed that children's alternation errors were exclusively *underapplying* the alternation and that those errors decrease with age [5]. When deciding whether nouns alternate, Turkish adults show sensitivity to some but not all lexical statistical regularities [1]: they appear sensitive to the identity of the stop and the prosodic shape of the noun, but not the quality of stop-preceding vowels, all of which are statistically predictive of (non-)alternation in a corpus [4]. [1] argued that Turkish learners match the frequency of alternation for different categories of words but that the categories are shaped by UG constraints that exclude vowel quality interacting with stop voicing. An alternative, learning-based view, is that learners construct a grammar to sufficiently account for where alternation occurs and thus only track dependencies when generalization is untenable without them [2]. Both proposals accounted for adult Turkish wug responses, though [1]'s proposal requires that UG filter out certain generalizations while [2]'s does not. We argue here that only [2]'s proposal is consistent with Turkish child acquisition results [5].

[2]'s model constructs rules, in the form of a decision tree, to predict which nouns alternate by recursively subdividing words based on their features. At each recursive step, the model splits on the feature shared by the largest number of alternating nouns. Recursion stops when a productive number of nouns at the node do or do not alternate, measured by the Tolerance Principle [7] (see [2, 3] for details). **Figure 1** gives an example; as [2] showed, the alternation is sufficiently predictable without needing to look into the stem to find the vowel quality dependency.

We augmented 14071 TELL [4] nouns with frequencies [6] and simulated learner trajectories by sampling nouns in increments from 1K-14K, weighted by frequency. We ran the model on 30 sampled trajectories and prompted each model on each of [5]'s 48 test words for whether it is predicted to alternate. We implemented a comparison model encapsulating [1]'s proposal, which computes in each of [1]'s proposed categories (e.g. "polysyllabic [k]-final") the fraction that alternate. The expected number of errors from matching those frequencies is compared to [2]'s model errors across all sampled trajectories.

Table 1 shows the results. The proposed model accounts for the fact that children's errors are exclusively underapplications and that errors decrease with training size. This is because it only applies the alternation when it learns the conditions in which it is warranted. Frequency matching fails to match either fact because [1]'s categories contain some non-majority class items and that ratio remains largely constant throughout training. Thus, while both [1] and [2]'s proposal account for [1]'s adult wug test results, [2]'s does not require that UG filter out certain generalizations and is the only one to match child language acquisition.

(500 words)

(1) (a)	[kɑnɑ**t**] ~ [kɑnɑ**d**-ɯ]	'wing'-ACC	(b)	[sy**t**] ~ [sy**t**-y]	'milk'-ACC
	[t͡ʃoɾɑ**p**] ~ [t͡ʃoɾɑ**b**-ɯ]	'sock'-ACC		[to**p**] ~ [to**p**-u]	'ball'-ACC
	[t͡ʃeki**t͡ʃ**] ~ [t͡ʃeki**d͡ʒ**-i]	'hammer'-ACC		[sɑ**t͡ʃ**] ~ [sɑ**t͡ʃ**-ɯ]	'hair'-ACC
	[balɯ**k**] ~ [balɯ-ɯ]	'fish'-ACC		[ke**k**] ~ [ke**k**-i]	'cake'-ACC

Table 1: The % of errors that are underextensions and the decrease in error % from 1K-14K nouns (youngest-oldest age group for humans). The p-values are for Welch's one-sided t-test comparing mean error rates between 1K and 14K.

	% Errors Underapplication	Decrease in Error
Humans [5]	100%	27%
Proposal [2]	100%	32% ($p < 0.05$)
FreqMatch [1]	78%	-3% ($p = 0.57$)

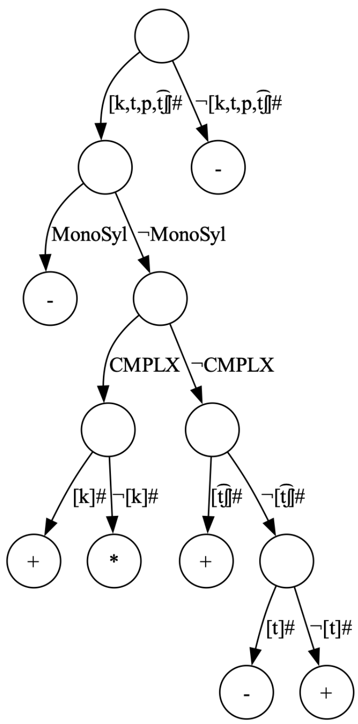

Fig 1: Example tree learned from a sample of 3500 TELL nouns (+/− = +/−alternate; * = not productive)

References

[1] Becker, M., Ketrez, N., & Nevins, A. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*.
[2] Belth, C. (2025). A Learning Based Account of Turkish Laryngeal Alternations. *AMP.*
[3] Belth, C., Payne, S., Beser, D., Kodner, J., & Yang, C. (2021). The greedy and recursive search for morphological productivity. *CogSci*.
[4] Inkelas, S., Küntay, A., Orgun, O., & Sprouse, R. (2000). Turkish electronic living lexicon (TELL). *Turkic Languages*.
[5] Nakipoglu, M., Furman, R., & Üntak, A. (2016). Acquisition of morphophonemic alternations and the role of frequency. *Exploring the Turkish linguistic landscape*.
[6] Wenzek, G., Lachaux, M. A., Conneau, A., Chaudhary, V., Guzmán, F., Joulin, A., & Grave, E. (2019). CCNet: Extracting high quality monolingual datasets from web crawl data. *LREC.*
[7] Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. MIT press.