Modeling Morphological Learning: Tolerance Principle on Turkish past tense -DI

Rüveyda Şahyar[1]

[1]Boğaziçi University, Department of Linguistics

ruveyda.sahyar@std.bogazici.edu.tr

This paper tests Belth et al.'s learning model Abduction of Tolerable Productivity (2021) on the Turkish past tense suffix -DI. The model is a greedy search algorithm that recursively generates a decision tree based on Yang's Tolerance Principle (2016), a proposed predictor of linguistic productivity based on the number of words within the scope of a rule and exceptions to it. Belth et al. note that as the principle is based on the number of exceptions to the proposed rule, it allows for productive rules with limited vocabularies and therefore presents a good model of early acquisition (2021).

Turkish -DI presents an interesting case for a learning model of morphology as it has 8 allomorphs conditioned by voice assimilation and vowel harmony with regards to frontness and roundness (Göksel & Kerslake 2005).

(1)	gel	-di		(3)	ısır	-dı		(5)	oku	-du		(7)	gör	-dü	
	come	-DI			bite	-DI			read	-DI			see	-DI	
They(sg.) came.				They(sg.) bit.				They(sg.) read.				They(sg.) saw.			
(2)	git-ti			(4)	yap	-tı		(6)	somurt	-tu		(8)	düş	-tü	
	go-DI				do	-DI			frown	-DI			fall	-DI	
They(sg.) went.				They(sg.) did.				They(sg.) frowned.				They(sg.) fell			

In the acquisition literature, the morpheme -DI is reported as being productively used by Turkish-acquiring children as early as 1;5 of age (Aksu-Koç & Ketrez, 2003) with very little error (Aksu-Koç & Slobin 1985). This success can be attributed to the fact that vowel harmony and voice assimilation rules apply consistently for all verbal stems. Crucially, Aksu-Koç & Ketrez report a Turkish-acquiring child as young as 1;3 to use the -tü allomorph when the child has less than seven verbs in their speech (2003). This allomorph displays frontness and roundness harmony as well as voice assimilation and is the least frequent form of the morpheme to appear in the combined corpus for this study. We are therefore presented with an interesting challenge where the suffix has many allomorphs that are completely rule-driven, and even the least frequent form is acquired by children very early. Given that TP formulates rules that minimize the number of exceptions, we would predict the model to be successful with these phonologically conditioned allomorphs that exhibit no irregularity. If we account for the high number of allophones, each with different frequency in the input, however, we expect a lower success rate from the model. Furthermore, we must note that although this rule-based allomorphy might be trivially explained using assimilation and vowel harmony, it represents a significant challenge for the model at hand. Belth et al's model tests the final segment of a lemma for a given suffix against the Tolerance Principle, then the final two segments in case it is not productive under TP, and so forth. The allomorph of -DI for a given verb can depend on as much as three final segments (ex. 6), or as little as one (ex. 5). This would mean that in absence of abstraction the model would have to consider, at worst, to estimate roughly disregarding phonotactic constraints, 21 consonants in Turkish orthography + 21 x 21 + 8 vowels in Turkish orthography x 21 x 21 = 3990 possible rules for each allomorph of -DI against TP. In testing the model's performance in Turkish -DI morpheme, we are therefore evaluating its ability to learn complex yet regular rules with limited occurrence, the exact strength of Tolerance Principle as claimed (Belth et al. 2021).

The present study uses a combined corpus of Turkish verbs inflected with -DI to train the ATP model. The dataset comprises of 751 Turkish verbs inflected with -DI. 328 verbs from child-produced and child-directed speech from CHILDES Turkish corpora (280 from Aksu (Slobin, 1982), 192 from Altinkamis (Altinkamis, 2003)) as well as 900 most frequent verbs in Universal Dependencies Turkish Penn 2.10 Treebank (Kuzgun et al., 2020) that were manually checked and cleaned, then combined by removing overlapping instances. Verbs found in Turkish CHILDES corpus were extracted using UDPipe 2.0 (Straka, 2018) tool, while the data from UD-Turkish Penn 2.10 was queried through PML Tree Query (Štěpánek & Pajas, 2010). Both tools were accessed through the LINDAT/CLARIAH-CZ Research Infrastructure (https://lindat.cz), which was supported by the Ministry of Education, Youth and Sports of the Czech Republic (Project No. LM2018101). The verb roots were then inflected with the relevant forms of -DI morpheme, using a Context-Free Grammar with Natural Language Toolkit for Python (Bird et al., 2009). The train/test split was done using the relevant method from the Scikit Learn Python library (Pedregosa et al., 2011) with 563 verbs for training and 188 for testing.

	Features in experiment	Precision	Recall	F1
Experiment 1	[+ / - VOICE]	1.0	1.0	1.0
Experiment 2	[+ / - BACK]	0.955539	0.934803	0.943099
Experiment 3	[+ / - ROUND]	0.734524	0.650497	0.675638
Experiment 4	[+ / - VOICE] [+ / - BACK]	0.951042	0.942859	0.946500
Experiment 5	[+ / - VOICE] [+ / - ROUND]	0.867888	0.777437	0.805699
Experiment 6	[+ / - BACK] [+ / - ROUND]	0.906071	0.891674	0.893532
Experiment 7 **(Turkish forms)**	[+ / - VOICE] [+ / - BACK]	0.883886	0.888727	0.880219

To help clarify the effect of rule complexity and frequency of each allomorph on model performance, seven experiments are carried out to test the model's performance with regards to each phonological feature. Each experiment is evaluated with regards to its decision trees as well as precision, recall and F1 analyses for the test data. Additionally, a wug-test (following Berko, 1958) of 8 nonce words is performed to evaluate performance on each allomorph. As decision trees explicitly demonstrate the rules formulated by the model, they allow for an in-depth discussion of the factors effecting performance. The results are compared with findings from acquisition literature both in terms of acquisition of morphology and of voice assimilation and vowel harmony. Considering the input for the model only reflects type frequency, the results are further discussed with reference to the effect of type versus token frequency in children's acquisition.

References

Aksu-Koç, A. & Slobin, D.I. (1985). The acquisition of Turkish. In *The Crosslinguistic Study of Language Acquisition*, Vol. 1, D.I. Slobin (ed.), 839–878. Hillsdale NJ: Lawrence Erlbaum Associates.

Aksu-Koç, A. & Ketrez, F.N. (2003). Early verbal morphology in Turkish: Emergence of inflections. In *Mini-paradigms and the Emergence of Verb Morphology*, D. Bittner, W.U. Dressler & M. Kilani-Schoch (eds), 27–52. Berlin: Mouton de Gruyter.

Altinkamis, F. (2003). CHILDES Database Turkish Altinkamis Corpus. DOI: 10.21415/T5H89W

Belth, C., Payne, S., Beser, D., Kodner, J., & Yang, C. (2021). The Greedy and Recursive Search for Morphological Productivity. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43(43).

Berko, J. (1958). The Child's Learning of English Morphology, *WORD*, 14(2-3), 150-177, DOI: 10.1080/00437956.1958.11659661.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. O'Reilly. www.nltk.org/book/

Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. London: Routledge.

Kuzgun, A., Cesur, N., Arıcan, B. N., Özçelik, M., Marşan, B., Kara, N., Aslan, D. B., & Yıldız, O. T. (2020). On building the largest and cross-linguistic Turkish dependency corpus. In *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 1-6. IEEE.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. In *Journal of Machine Learning Research* 12(85), 2825−2830.

Slobin, D. (1982). Universal and particular in the acquisition of language. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*, 128-172. New York: Cambridge University Press. DOI: 10.21415/T5SP6C

Štěpánek, J. & Pajas, P. (2010). Querying Diverse Treebanks in a Uniform Way, in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, European Language Resources Association (ELRA), 1828-1835. Valletta: Malta.

Straka, M. (2018). UDPipe 2.0 Prototype at CoNLL 2018 UD Shared Task. In *Proceedings of CoNLL 2018: The SIGNLL Conference on Computational Natural Language Learning*, 197-207, Association for Computational Linguistics, Stroudsburg, PA: USA.

Yang, C. (2016). The price of linguistic productivity: How children learn to break the rules of language. MIT Press.