OCP Effects in Turkish Partial Reduplication: Locality and Feature Specificity

Kevin Tang (Zhejiang University; kevtang@gmail.com) &
Faruk Akkus (University of Pennsylvania; akkusf@sas.upenn.edu)

This paper investigates the *partial (emphatic) reduplication* in Turkish, used with modifiers (adverbs and adjectives), which gives the modifier the meaning of 'fullness' (Demircan 1987).

Base	Gloss	Reduplication	Gloss
dingin	'serene'	*di**p**-dingin*	'very bald'
beyaz	'white'	*be**m**-beyaz*	'very white'
ma:vi	'blue'	*ma**s**-ma:vi*	'fully blue'
temiz	'clean'	*te**r**-temiz*	'completely clean'

The reduplicant, realized as a prefix, has the form $(C_1)VC_2$. C_2 of the reduplicant prefix ends in one of the four *linking consonants* (LC): *-p, -m, -s, -r* (Lewis 1967). While previous studies on the topic approach the issue from different angles with different constraints, they do not take into consideration many factors (Table 1). These data are arguably inconclusive because many relied solely on intuitions (Gibson & Fedorenko 2013), examined a small number of participants and items, used a forced-choice task and not a rating task (Kawahara 2015). This study undertakes a comprehensive data collection and analysis.

Method: 162 real words were selected from these studies to cover a broad range of previously commonly tested items. To enable within participant comparisons, each participant was tested on both a rating task and a forced-choice task (4-AFC). Participants were asked to rate a naturalness scale of 1-7 of each of the 4 reduplicated forms per item as well as to pick one option out of the 4 reduplicated forms. The items were divided into 5 lists, each tested by ≈ 40 participants. This amounts to 208 participants with detailed demographic information.

Discussion: Previous studies converge on the view that the dissimilation in reduplication stems from some kind of OCP, yet questions regarding the nature of the OCP or the extent of its effect have not been explained thoroughly and remain largely at the observational level. Most of the studies reduce the OCP effect to anti-faithfulness constraints between segments, explained at the level of natural classes, e.g. Demircan (1987), Yu (1999), Kelepir (2000). Moreover, the OCP between the reduplicant and the base is assumed to extend to C_2 of the base, and no further (Demircan 1987, Kelepir 2000, Sofu 2005). However, minimal pairs showing that the effect extends to C_3 can be found. For instance, **beyaz** 'white' and **bayat** 'stale' have identical consonants except for C_3 ([z] vs. [t]) (n.b. vowels' identity do not matter (Yu 1999)), but the LC [s] is only dispreferred with *beyaz* (4-AFC: 0%; Rating: 3.6/7), while it is the most preferred LC with *bayat* (4-AFC: 68%; Rating: 5.95/7).

<u>Modelling:</u> Only the rating data were described below because the results mostly converged with those of the 4-AFC data and the rating task was rarely used in previous studies. The trial-level rating data were analysed using mixed-effect regression modelling (*lme4* in *R* (R Core Team 2013)). Focusing on consonant-initial words, the data were divided into three groups by the number of consonants they contain (C_1C_2, $C_1C_2C_3$, $C_1C_2C_3C_4$). Following Graff & Jaeger (2009), to examine the **positional and feature specificity** of the OCP effects, each phonological feature was specified as an independent predictor for each consonant (IndFeat). Other predictors were included such as total identity of consonants (LC = C_{1-4}) (Iden), sum of the matched features (SumFeat), and ease of articulatory (betwee LC and C_1). *Random* effects of items, their LC, and participants were included.

<u>Feature-specific effect:</u> First we focused on the question of specificity by fitting a number of separate models with the following key predictors: (1) Iden, (2) SumFeat, (3) IndFeat, (4) Iden + SumFeat, (5) Iden + IndFeat. The best model based on AIC/BIC was found to be **Iden +**

IndFeat which includes both OCP predictors of total identity as well as individual features. For C_1C_2, the AIC of Iden + IndFeat is 491 lower than the next best model Iden + SumFeat; similarly for $C_1C_2C_3$, the AIC difference is 681 and for $C_1C_2C_3C_4$, it is 556 (n.b. an AIC difference of >2 is considered as significant). The paper therefore shows that the classic factors reported in the literature, e.g. avoid full reduplication or avoid doubling consonants (Demircan 1987), or natural classes such as coronal, labial (Kelepir 2000) are not finer-grained enough to capture the complete pattern. We provide evidence (i) that OCP constraints need to refer to phonological features individually, and (ii) the individual features are not weighted equally.

<u>Locality effect:</u> Second we focused on the question of locality by dropping the predictors related to each of the base consonants in bulk, e.g. for C_1C_2 words, two models were fitted with only predictors concerning one of the two consonants. Model comparisons were made to evaluate the relative importance of each consonant. Our analysis makes a significant contribution by demonstrating that the OCP effect is observed at the whole word-level, not just restricted to C_1 or C_2. We show that in words consisting of four consonants, the OCP effect extends all the way to C_4. Thus our finding supports the view that OCP is not solely an adjacency phenomenon, but the OCP constraints need to be able to refer to feature matches between consonants separated by another consonant (e.g., Pierrehumbert 1993, Frisch & Zawaydeh 2001, Frisch et al. 2004, Graff & Jaeger 2009). Relatedly, the previous studies predict that the effect size should decrease from left to right, or disappear after some point (Zymet 2014). However, our findings show that this is not entirely correct for Turkish, as we also note that the degree a certain consonant plays a role varies, e.g. C_2 outweighs C_1 in C_1C_2 words.

Conclusion: This paper presents a comprehensive analysis of Turkish Particial Reduplication. We highlight that the OCP constraints are more graded than they have been previously proposed. Our findings support previous work on the formulations of OCP constraints that treat individual features as free parameters in the similarity computation (e.g., Graff & Jaeger 2009). The surprising finding with regard to locality is that not only the OCP effect extends all the way from C_1 to C_4, but it is not weighted linearly from high to low.

Table 1: A summary of the data examined in 10 previous studies

Sources	Intuition	Experiment	Type of Experiment	# of Participants	# of Items	Item Types
Kelepir (2000)	Yes	No	-	-	89	Real
Yu (1999)	Yes	No	-	-	152	Real
Hatiboğlu (1973)	Yes	No	-	-	142	Real
Demircan (1987)	Yes	Yes	Forced-Choice (FC)	100	130	Real (110), Nonce (20)
Sofu (2005)	-	Yes	FC	25 adults, 89 children	38	Nonce
Sofu & Altan (2008)	-	Yes + Corpus	FC	80	132	Real
Dobrovolsky (1987)	Yes	No	-	-	9	Real
Wedel (1999)	Yes	Yes	FC	-	125 + 80	Real
Taneri (1990)	-	Yes	FC	32	300	Real
Kaufman (2014)	-	Yes	1-FC, 2-Rating	1-16, 2-50	1-44, 2-45	1-Nonce, 2-Real

Demircan, Ozgur. 1987. Emphatic reduplication in Turkish. In *Studies on modern Turkish: Proceedings of the 3rd conference on Turkish linguistics*, 24–41.
Dobrovolsky, Michael. 1987. Why CVC in Turkish reduplication. *The Nordic languages and modern linguistics* 6. 131–146.
Frisch, Stefan A, Janet B Pierrehumbert & Michael B Broe. 2004. Similarity avoidance and the ocp. *Natural Language & Linguistic Theory* 22(1). 179–228.
Frisch, Stefan A & Bushra Adnan Zawaydeh. 2001. The psychological reality of ocp-place in arabic. *Language*. 91–106.
Gibson, Edward & Evelina Fedorenko. 2013. The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28(1-2). 88–124.
Graff, Peter & T Jaeger. 2009. Locality and feature specificity in ocp effects: evidence from aymara, dutch, and javanese. In *Proceedings from the annual meeting of the chicago linguistic society*, vol. 45, 127–141.
Hatiboğlu, V. 1973. *Pekiştirme ve kuralları*. Türk Dil Kurumu Tanıtım Yayınları.
Kaufman, Brianna Danielle. 2014. *Learning an unproductive process: Turkish emphatic reduplication*. UC Santa Cruz MA thesis.
Kawahara, Shigeto. 2015. Comparing a forced-choice wug test and a naturalness rating test: an exploration using rendaku. *Language Sciences* 48. 42–47. http://dx.doi.org/10.1016/j.langsci.2014.12.001.
Kelepir, Meltem. 2000. To be or not to be faithful. In *Studies on Turkish and Turkic Languages: Proceedings of the Ninth International Conference on Turkish Linguistics*, 11–18.
Lewis, Geoffrey L. 1967. *Turkish grammar*. Oxford University Press.
Pierrehumbert, Janet. 1993. Dissimilarity in the arabic verbal roots. In *Proceedings of nels*, vol. 23, 367–381.
R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/.
Sofu, Hatice. 2005. Acquisition of reduplication in Turkish. In *Studies on reduplication*, 493–509. Mouton de Gruyter.
Sofu, Hatice & Aslı Altan. 2008. Partial reduplication: revisited. In *Essays on Turkish Linguistics: Proceedings of the 14th International Conference on Turkish Linguistics*, 63–73. Otto Harrassowitz Verlag.
Taneri, Mübeccel. 1990. A type of reduplication in turkish. In *Kansas working papers in linguistics*, 93–126.
Wedel, Andrew. 1999. Turkish emphatic reduplication. *Linguistics Research Center*.
Yu, Alan. 1999. Dissimilation and allomorphy: the case of Turkish emphatic reduplication. *UC-Berkeley ms*.
Zymet, Jesse. 2014. Distance-based decay in long-distance phonological processes. In *The proceedings of the 32nd west coast conference on formal linguistics*, 72–81.