A Diachronic Counter-example to the Subset Principle:
The Case of Anatolian Reduplication

Anthony D. Yates (UCLA) & Sam Zukoff (Princeton)

INTRODUCTION: The "Subset Principle" (cf. Prince & Tesar 2004) states that, when learners are choosing between multiple possible grammars consistent with the positive evidence, they ought to select the grammar that is most restrictive (i.e., allows the fewest possible unseen forms), because doing otherwise has the potential to overgenerate relative to the target language and make it impossible to later arrive at the more restrictive target language. In phonology, this reduces mainly to a preference for the higher ranking of markedness constraints than faithfulness constraints, as implemented in, for example, Biased Constraint Demotion (BCD; Prince & Tesar 2004) and Low Faithfulness Constraint Demotion (LFCD; Hayes 2004). It is standardly assumed that it is a desideratum of a phonological learning procedure for it to capture the Subset Principle. Capturing the Subset Principle is thus taken as one of the key arguments in favor of BCD and LFCD over simple Recursive Constraint Demotion (RCD; Tesar 1995, Tesar & Smolensky 1998, 2000).

In this paper, we present a case which challenges the universality of Subset Principle-based phonological learning. In the diachronic development of the reduplicative system from Proto-Anatolian into Hittite, speakers evidently learned a superset grammar. By this we mean that they learned a grammar that tolerated violations of a markedness constraint which was surface-true at an earlier stage, despite never having evidence that it could be violated; this is evidenced by the emergence of a reduplication pattern which violates the previously inviolable constraint.

We propose that this sort of non-Subset Principle-compliant learning can be explained based on the relative *informativity* of the constraints involved in the phenomena. We implement this with a slightly modified version of RCD (also compatible with BCD), termed "Maximally Informative Recursive Constraint Demotion" (MIRCD), which is biased towards winner-preferring constraints that can account for the greatest amount of data possible.

ANATOLIAN DATA & ANALYSIS: The reduplication patterns of Hittite and its (reconstructed) proximate ancestor Proto-Anatolian (PA) are shown in (1) (following Yates & Zukoff 2016a,b, fthc., Zukoff 2017; cf. Dempsey 2015).

	Base Shape	Proto-Anatolian (PA)	>	Hittite	cf. Luwian
	CVX–	*<u>CV</u>-CVX–	>	<u>CV</u>-CVX–	<u>CV</u>-CVX–
(1)	*TRVX–*	*<u>TV</u>-TRVX–	>	<u>TRV</u>-TRVX–	<u>TV</u>-TRVX–
	STVX–	*<u>STV</u>-STVX–	>	*i*<u>STV</u>-STVX–	[<u>TV</u>-STVX–]
	VRTX–	*does not exist yet*	>	<u>VR</u>-VRTX–	<u>VR</u>-VRTX–

Note first the distinction in copying patterns in PA between *obstruent-sonorant* (TRVX–) bases and *s-obstruent* (STVX–) bases. This variation mirrors that seen in many of the other ancient Indo-European languages (Steriade 1988, Fleischhacker 2005, Keydana 2006, 2012, DeLisi 2015, Zukoff 2017). This variation can be analyzed with a constraint against certain types of consonant repetitions, termed the NO POORLY-CUED REPETITIONS constraint (Zukoff 2017):

(2) **NO POORLY-CUED REPETITIONS (*PCR) [$\approx$ *$C_\alpha V C_\alpha$ / _$C_{[\text{-sonorant}]}$]**
For each sequence of repeated identical consonants separated by a vowel ($C_\alpha V C_\alpha$), assign a violation mark * if that sequence immediately precedes an obstruent.

When this markedness constraint outranks the constraint against consonant clusters (*CC), it can cause STVX– bases to divert away from the default reduplication pattern (C_1-copying, as seen in TRVX– bases; cf. (4.i) below) to the cluster-copying pattern, as shown in (3.i) for PA.

(3)

i. Proto-Anatolian > ii. Hittite

/RED, stu-/	***PCR**	*CC	CNTG$_{BR}$
a. su-stu–	*!	*	*
b. ☞ stu-stu–		**	

/RED, ark-/	CNTG$_{BR}$	*CC	***PCR**
a. ☞ ar-ark–		*	*
b. ark-ark–		**!	

Within PA, there is no evidence that *PCR can be violated in reduplication. Nevertheless, the Hittite VR-VRTX– pattern — which arises only after Proto-Anatolian, when vowel-initial roots first enter the language (Yates & Zukoff fthc.) — expressly violates *PCR, and requires that it be ranked at the bottom of the grammar, below both *CC and CONTIGUITY-BR [CNTG$_{BR}$] (McCarthy & Prince 1995), as shown in (3.ii). This constitutes **a diachronic counter-example to the Subset Principle**, because speakers evidently learned a less restrictive language (with a low-ranked markedness constraint, *PCR) than was warranted by the positive evidence.

DIACHRONY & LEARNING: What could have led learners to fail to obey the Subset Principle in this case? We hypothesize that the answer lies in changes in the relative *informativity* of the different constraints involved the relevant phenomena.

Prior to the change in ranking of *PCR, there was an independent change in the reduplicative behavior of TRVX– bases: PA *TV-TRVX– > Hitt. TRV-TRVX–. The new cluster-copying pattern for Hittite TRVX– bases entails the new ranking of CNTG$_{BR}$ above *CC, as seen in (4). Given this ranking, a high-ranking *PCR is no longer required in order to explain the STV-STVX– pattern: CNTG$_{BR}$ $\gg$ *CC means that cluster-copying will be the default behavior for all cluster types.

(4)

i. Proto-Anatolian > ii. Hittite

/RED, pri-/	*PCR	*CC	**CNTG$_{BR}$**
a. ☞ pi-pri–		*	*
b. pri-pri–		**!	

/RED, pri-/	**CNTG$_{BR}$**	*CC	*PCR
a. pi-pri–	*!	*	
b. ☞ pri-pri–		**	

The independent promotion of CNTG$_{BR}$ essentially saps *PCR of its informativity, in that it is no longer uniquely required in order to explain any data points. That is to say, for each *Winner∼Loser* pair for which *PCR correctly favors the winner, CNTG$_{BR}$ also correctly favors the winner; but there are additional pairs for which CNTG$_{BR}$ favors the winner but *PCR treats both equally. While both CNTG$_{BR}$ and *PCR are now uniquely winner-preferring constraints, CNTG$_{BR}$ prefers **more winners**.

We can thus explain the total demotion of *PCR if we have a learning algorithm that preferentially installs the constraints that favor the most winners first (cf. Becker 2009). In this case, once CNTG$_{BR}$ is installed, all *Winner∼Loser* pairs which *PCR could have explained are removed from the support. *PCR is not an active winner preferrer within the remaining support (it treats all remaining *Winner∼Loser* pairs the same), and thus will never be installed above another constraint. We implement this strategy as a modification to RCD, which we term "Maximally Informative Recursive Constraint Demotion" (MIRCD). This approach is also consistent with BCD, as long as the informativity bias is enacted prior to the markedness bias.

MIRCD retains Subset Principle-based learning in all situations other than that represented by the current case. What distinguishes this situation is the superset-subset relationship between CNTG$_{BR}$ and *PCR: *PCR explains a proper subset of the data which CNTG$_{BR}$ explains. Under these specific conditions, MIRCD is thus shown to produce the non-Subset Principle-compliant learning necessary to capture the Anatolian facts, without predicting non-Subset Principle-compliant learning in the general case.