Reevaluating Root Structure Constraints in Proto-Indo-European: the case of **DVD

Adam I. Cooper, Cornell University

The Proto-Indo-European (PIE) root structure constraint **DVD, prohibiting the co-occurrence of voiced unaspirated stops (of which there are five), has received special attention in the literature with respect to its validity. As argued by Iverson and Salmons (1992), and later Barrack (2002), the constraint ought not to be posited for PIE at all: rather, its effects, they claim, can be accounted for by various distributional facts about the voiceless unaspirated stops in particular, and of stops in general. While their explanation is compelling, their reliance on an outdated corpus, the statistical study of PIE roots performed by Jucquois (1966), itself based on Pokorny (1959), leaves a more careful assessment of their claim difficult to develop.

As part of a larger project evaluating the suitability of categorical root structure constraints for PIE from the perspective of modern phonological theory, I consider the behavior of voiced unaspirated stops in the PIE root using a statistically-oriented methodology which draws on Rix et al. (2001)'s corpus of verbal roots. The data set consists of the unambiguous CVC core of 781 verbal roots; though Rix et al. (2001) collect 1195 roots, the other 414 are of questionable PIE date or ambiguous in their reconstruction (or both), and as such have been excluded from consideration. To quantify over-representation and under-representation in the data, I employ the ratio O(bserved) / E(xpected), as used by Pierrehumbert (1993) and more recently, by Frisch, Pierrehumbert and Broe (2004) for their study and similarity avoidance analysis of consonantal co-occurrence in the Arabic root. A formula for the calculation of O/E is given in (1); following past work, a value less than 1.00 suggests a restriction is active against the relevant consonantal pairing, while a value greater than 1.00 suggests no such constraint is active.

O/E values have been calculated for the 781 consonant pairings found in the data set. Specifically with respect to voiced unaspirated stops, we can confirm that instances of their co-occurrence are in fact non-existent (O/E = 0.00), a finding which ostensibly justifies the positing of **DVD. However, if, like Iverson and Salmons (1992) and Barrack (2002), we consider associated information such as the frequency of voiced unaspirated stops, we could conclude that their relative rarity to begin with is to blame for their lack of co-occurrence. Voiced unaspirated stops occur only 46 times prevocalically and 34 times postvocalically, and as such only about 2 roots are expected to occur with a pairing of them (by the denominator portion of the formula in (1)). A difference of this sort is not significant enough to exclude the possibility that the gap is simply due to chance; thus Iverson and Salmons and Barrack appear correct in their explanation.

My findings, however, provide a different motivation for questioning the positing of the specific constraint **DVD – distinct from Iverson and Salmons (1992) and Barrack (2002)'s reliance on frequency counts – based on the broader scope of this study, which examines the behavior of all PIE consonants, not just stops, and thereby offers a more complete perspective. The O/E results presented in (2) are aggregated by manner of articulation, and as they indicate, it is the case that co-occurrence of like manner consonants is in general restricted, indeed across all manner classes. Yet if we focus specifically on the stops, and break down the results by series, as in (3), we see that the voiceless unaspirated stops have the expected restricted co-occurrence, as do the majority of the stop pairings, while the other two series of stops, voiceless unaspirated and voiced aspirated, co-occur freely (O/E = 1.17 and 2.44, respectively). Given this fact, the behavior of voiced unaspirated stops can actually be considered typical, part of a more general trend, while it is the voiceless unaspirated and voiced aspirated stops (and, incidentally, voiced unaspirated / aspirated pairings as well) whose behavior is singular and in need of explanation.

I argue that efforts at reconstruction which have attempted to abide by the purported constraint **DVD have in part led to this result; I also take into consideration the fact that the majority of double-stop roots (and like-manner pairings in general) are found in roots of shape *CVC-, which I argue provides a freer environment for manner co-occurrence than roots of more complex shapes.

(1) **O(bserved)/E(xpected)**

$$\frac{\text{Observed } \{C_1, C_2\} \text{ co-occurrence in roots}}{\frac{\text{Observed } /C_1/ \text{ occurrence in roots} \times \text{Observed } /C_2/ \text{ occurrence in roots}}{\text{Total roots}}}$$

(2) **O/E values for consonantal co-occurrence in the PIE root, aggregated by manner of articulation**

	Stops	**Fricatives**	**Nasals**	**Liquids**	**Glides**
Stops	**0.79**	0.81	1.26	1.44	0.77
Fricatives	1.28	**0.69**	1.01	1.21	0.86
Nasals	0.98	1.11	**0.66**	0.91	1.18
Liquids	1.02	1.25	0.94	**0.00**	1.67
Glides	1.21	1.29	0.71	1.29	**0.59**

($\chi^2 = 110.796$, $p < 0.0001$)

(3) **O/E values for stop co-occurrence in the PIE root, by series**

	Voiceless Unaspirated	**Voiced Unaspirated**	**Voiced Aspirated**
Voiceless Unaspirated	**1.17**	0.53	0.45
Voiced Unaspirated	0.51	**0.00**	0.00
Voiced Aspirated	0.00	1.79	**2.44**

($\chi^2 = 141.558$, $p < 0.0001$ (computed over all 781 roots))

References

Barrack, Charles M. (2002). The Glottalic Theory revisited: a negative appraisal. *Indogermanische Forschungen* 107, 76-95.

Frisch, Stefan A., Janet B. Pierrehumbert and Michael B. Broe. (2004). Similarity Avoidance and the OCP. *Natural Language & Linguistic Theory* 22, 179-228.

Iverson, Gregory K. and Joseph C. Salmons. (1992). The Phonology of the Proto-Indo-European Root Structure Constraints. *Lingua* 87, 293-320.

Jucquois, Guy. (1966). La structure des racines en indo-européen envisage d'un point de vue statistique. *Linguistic Research in Belgium*, ed. by Yvan Lebrun, 57-68. Wetteren: Universa.

Pierrehumbert, Janet. (1993). Dissimilarity in the Arabic verbal roots. *Proceedings of the Annual Meeting of the North Eastern Linguistics Society* 23, 367–381.

Pokorny, Julius. (1959). *Indogermanische etymologisches Wörterbuch*. Bern: Francke.

Rix, Helmut, et al., eds. (2001). *Lexikon der indogermanischen Verben*. Wiesbaden: L. Reichert.