

## Greenberg's U20: turning Cinque's approach into a model of language change

Nikos Angelopoulos, Christos Zioutis, Elena Anagnostopoulou, David Inman, Dimitris Michelioudakis, Natalia Chousou-Polydouri, Pavlos Pavlidis

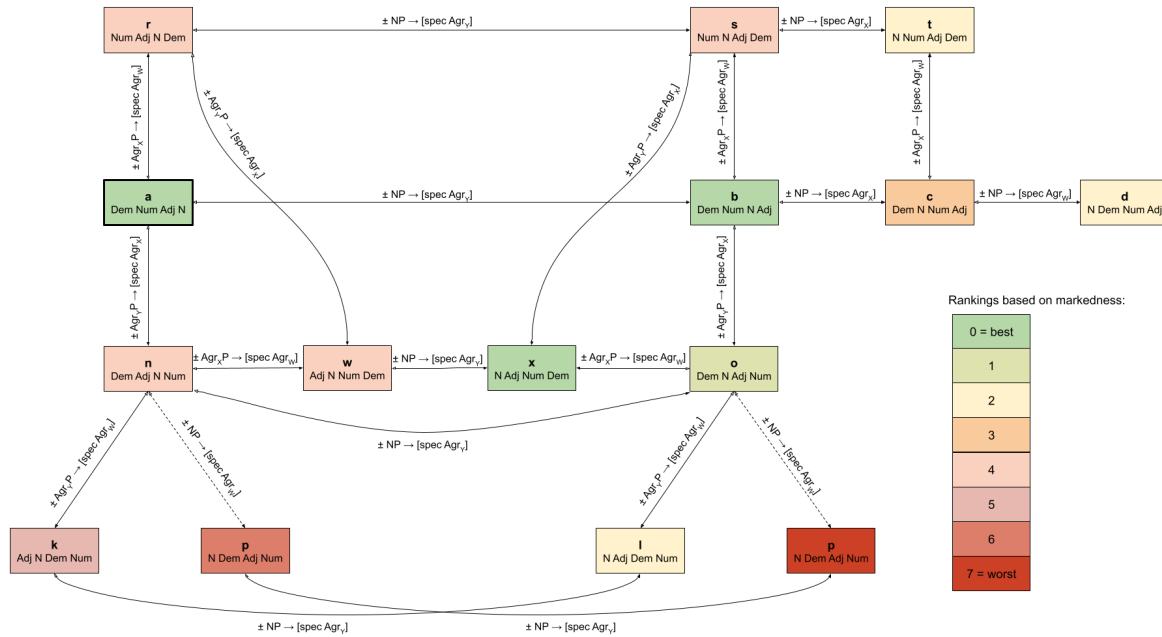
**1. Introduction.** According to Hawkins' (1983) rephrasing, Greenberg's (1963) U20 treats (a) [Dem>Num>Adj>N] as the unmarked order when Dem, Num, Adj are prenominal, and (b) [Dem>Num>Adj] as an unmarked order, as well as its mirror image [Adj>Num>Dem] when postnominal. Crucially, when the noun is final, the order in (a) is fixed, whereas more orders are attested when the noun is not final. Cinque (2005) proposes a movement-based approach to U20, which derives the attested orders and rules out the unattested ones; also, in contrast to Abels & Neeleman's (2009) analysis, it correlates their relative cross-linguistic frequency/rarity with the complexity/markedness of the underlying derivation. However, in order to test whether low markedness correlates with high frequency in a statistically valid way, we need to control for the effects of genealogical relationships among languages. In this talk, we propose a way of turning Cinque's model of synchronic parametric variation into a model of diachronic transitions between grammars and to subsequently test whether the markedness of the derivation indeed correlates with diachronic stability (which is equivalent to a high frequency in a large independent language sample).

**2. Cinque's (2005) model.** Cinque (2005) proposes that the following sequence of functional projections underlies the derivation of all attested orders (1):

(1) [<sub>AgrWP</sub> AgrW [<sub>WP</sub> DemP W [<sub>AgrXP</sub> AgrX [<sub>XP</sub> NumP X [<sub>AgrYP</sub> AgrY [<sub>YP</sub> AdjP Y [<sub>NP</sub>]]]]]]]]]

The order in (a) is a direct reflection of (1) without any further movement steps. All movement steps allowed involve constituents that contain NP: either the NP itself can be moved or bigger constituents with pied-piping. Pied-piping can take place in either of the following two ways: (i) movement of *picture-of-whose*-type constituents, i.e. without roll-up movement, or (ii) movement of *whose picture*-type constituents, formed after roll-up movement inside them. Cinque further posits a set of movement parameters that differ in markedness. In particular, (1) is unmarked because it does not involve any movement. Unmarked are also orders that involve pied-piping of the *whose-picture* type, i.e., the mirror image of (1), N>Adj>Num>Dem. Turning to the marked movement steps, these include: (i) NP movement without pied-piping (marked); (ii) partial movement of the NP (marked); (iii) NP movement with pied-piping of the *picture-of-who* type (more marked). The relative frequency of a given order is therefore expected to correlate with the number and degree of marked movement steps involved in its derivation. Orders derived via one or more marked movements should be less frequent than those requiring only unmarked operations, or none at all, as in the base order Dem>Num>Adj>N.

**3. Creating a model of language change.** We constructed a model of language change that derives transitions between attested word orders. Since unattested orders are not predicted by Cinque's (2005) analysis, they are excluded from the model. Our model operates on the core assumption that language change is not unconstrained; rather, it is movement-constrained. Transitions occur only between orders separated by the addition or removal of a *single* movement step, representing the minimal diachronic change. Consequently, certain typological shifts are predicted to be direct, while others indirect, depending on the number of movements required to be added or removed. This transition network is illustrated in the figure above. A language starting in state (a), with the order Dem>Num>Adj>N, can



transition to only three specific orders through a single movement step: (r), (n), or (b). These represent the most direct diachronic shifts. To reach state (n), a single movement step is required of the Agr<sub>v</sub>P constituent. In this instance, the Noun pied-pipes the AdjP to the Spec,Agr<sub>x</sub>P position, thereby deriving the new surface order Dem>Adj>N>Num. In the figure above, the specific movement step required is indicated above the arrow connecting the two orders. All attested orders are ranked and color-coded, reflecting Cinque’s hierarchy, from less to more marked (and therefore expected to be more to less stable diachronically) on a scale of 0-6. Mathematically, we implemented this model as a modified F81 (Felsenstein 1981) substitution process, which allows unequal stationary frequencies but assumes equal substitution rates across states. Specifically, to accommodate transitions predicted to be theoretically impossible, we constrained such transitions to have a rate of zero.

In order to test the predictions of the suggested model in terms of diachronic stability, we mapped Cinque’s U20 dataset of ~2000 languages onto glottocodes and subsequently on the global phylogeny by Bouckaert et al. (2022). By fitting our transition model on this tree, we can estimate the stationary frequencies of each order (a proxy for their diachronic stability). We use both Bayesian inference and Maximum Likelihood methods and we compare the performance of the inferred models with posterior predictive simulations. During a preliminary investigation multiple stationary frequency optima were inferred, some of which are in line with Cinque’s predictions. We will assess which ones are most reliable and discuss to what extent a movement-based approach can serve as a model of language variation and change.

**Selected References:** Abels, K., & Neeleman, A. (2012). Linear asymmetries and the LCA. *Syntax*, 15(1), 25-74. Bouckaert, R., D. Redding, O. Sheehan, et al. (2022). “Global Language Diversification Is Linked to Socio-Ecology and Threat Status.” In SocArXiv. July 20. Cinque, G. (2005). Deriving Greenberg’s Universal 20 and its exceptions. *Linguistic inquiry*, 36(3), 315-332. Hawkins, J. (1983). *Word order universals*. New York: Academic Press.