

## In defense of the Universal hierarchy of adjectival functional projections: Evidence from Slovenian complex NP

Vesna Plesničar

University of Nova Gorica  
vesna.plesnicar@ung.si

**Background.** When multiple attributive adjectives modify a noun in a given language, they are very likely to be ordered in a specific way, regardless of whether they are placed before or after the noun (e.g., Martin 1969, Hetzron 1978, Dixon 1982, Sproat and Shih 1991, Shlonsky 2004, Scontras and Kachakeche 2020). According to a highly prominent model in the domain of noun phrase research—the cartographic model of natural language syntax—crosslinguistic tendencies for relatively rigid adjective ordering can be accounted for if we assume that adjectives are merged in specifiers of various dedicated functional projections merged above the noun (Cinque 1994, 2010, Scott 2002, Shlonsky 2004, Laenzlinger 2005, Ramaglia 2014, Rizzi and Cinque 2016). The most elaborate version of adjective hierarchy within the cartographic model is the Universal hierarchy of adjectival functional projections proposed in Scott (2002), illustrated in (1).

- (1) SUBJECTIVE COMMENT > EVIDENTIAL > SIZE > LENGTH > HEIGHT > SPEED > ?DEPTH > WIDTH > WEIGHT > TEMPERATURE > ?WETNESS > AGE > SHAPE > COLOR > NATIONALITY/ORIGIN > MATERIAL > COMPOUND ELEMENT > NP

**Motivation.** As a highly prominent model in the domain of complex NP research, the cartographic model has also attracted a lot of criticism. Critics largely focus on Scott's (2002) model, which is frequently set as an example of being too detailed and rigid to account for the data observed (e.g., Truswell 2009, Scontras et al. 2017, Leivada and Westergaard 2019, Kotowski and Hartl 2019, Trotzke and Wittenberg 2019, Larson 2021). At the same time, the plausibility of the hypothesis under which the linear order of adjectives is a direct output of the functional hierarchy, where selectional instructions determine merge properties with a complement – e.g., SIZE head selects LENGTH projection, etc. (Rizzi and Cinque 2016) – is also receiving support. Evidence comes from the selective impairment of syntactically relevant semantic constraints on adjective order (Kemmerer 2000, Kemmerer et al. 2007 and Kemmerer et al. 2009), processing difficulties caused by marked adjective order (Kennison 2010), and language acquisition data indicating clear ordering tendencies in the acquisition of concepts denoted by adjectives (Smith 1984 and Pitchford and Mullen 2001). Thus, for example, the acquisition of concepts for COLOR, SHAPE, and SIZE is aligned with Scott's (2002) proposal if we assume that acquisition proceeds bottom-up, as argued for in acquisition studies of the nominal (e.g., Shafer and Villiers 2000), clausal (e.g., Radford 1996) and left-peripheral domains of the clause (Rizzi et al. 2021). Similarly, novel data from a highly detailed corpus examination of the validity of the predictions that Scott's (2002) model makes about language use (Plesničar to appear) also shows that the order of attested adjectives from thirteen semantic categories is overall compatible with the proposal that the order is governed by a hierarchy of adjectival projections. However, the scarcity of relevant data for certain subdomains of the proposed hierarchy in corpora makes it impossible to evaluate Scott's (2002) model as a whole. The purpose of the present paper is to provide a complementary source of evidence to refine the theoretical inquiry of Plesničar (to appear) and investigate the predictions of the adjectival hierarchy proposed by Scott (2002) in even more detail.

**Method.** We designed a two-alternative forced-choice task, using the Gorilla Experiment Builder software (Anwyl-Irvine, 2020), in which, subjects (N=55), while measuring their reaction time, had to choose their preferred variant out of a set of two adjective-adjective-noun strings with the same adjectives in both orders. Since, according to Scott, adjective order restrictions "fall out as a direct consequence of UG" (Scott, 2002: 97), we limited the response time to 5 seconds, to be able to really tap into speakers' intuitions. Our interest was in those combinations that are poorly represented in corpora (Plesničar to appear) or are relatively well-attested in corpora but deviate from Scott's (2002) hierarchy, i.e., SIZE-SPEED, LENGTH-SPEED,

DEPTH-WIDTH, HEIGHT-WEIGHT, WETNESS-AGE, WEIGHT-AGE, WIDTH-AGE, SPEED-AGE. The experiment consisted of 48 test items (6 per combination) and 48 fillers, carefully selected to resemble the task from the structural violation point of view, preventing unwanted relativization of preferred order selection in the case of test examples due to incomparable syntactic violation in the case of fillers (cf. Leivada and Westergaard 2019). We had 2 types of control items; Control 1, combinations SIZE-COLOR previously found to categorically follow Scott's (2002) hierarchy (cf. Truswell 2009, Scontras et al. 2017, Kotowski and Hartl 2019, Plesničar to appear), and Control 2, combinations of adjectives from the same semantic category, COLOR-COLOR, for which we presume an absence of any ordering preferences. Attention to the task was checked with comprehension questions that appeared randomly during the experiment.

**Results.** Table 1 presents the data from Plesničar's (to appear) corpus study merged with the bold-bordered data from our experiment. Depicted by higher percentage values in the upper-right part of the table, the responses given in the experiment are consistent with Scott's (2002) proposal. In the case of combinations DEPTH-WIDTH and WETNESS-AGE, for which Plesničar's (to appear) corpus results anticipate the reverse ordering, such ordering is also confirmed by our forced-choice data. We suggest that in these two pairs, the order of the categories is

Table 1: The percentage values of occurrences per combination of two semantic categories

	size	length	height	speed	width	depth	weight	temperature	age	wetness	shape	color	material
size	N/A	83%	93%	71%	96%	96%	91%	88%	87%	96%	98%	100%	99%
length	17%	N/A	74%	57%	79%	100%	58%	92%	54%	94%	89%	91%	99%
height	*7%	26%	N/A	N/A	59%	**67%	53%	*75%	54%	86%	79%	85%	99%
speed	29%	43%	100%	N/A	N/A	*60%	86%	*100%	59%	***100%	***100%	100%	100%
width	*4%	21%	41%	*100%	N/A	57%	N/A	N/A	53%	100%	100%	94%	99%
depth	4%	N/A	***33%	**40%	43%	N/A	**50%	85%	N/A	**67%	75%	95%	95%
weight	9%	42%	47%	***14%	*100%	**50%	N/A	*60%	62%	63%	100%	100%	100%
temperature	13%	8%	***25%	N/A	***100%	**15%	**40%	N/A	83%	74%	N/A	98%	100%
age	13%	46%	46%	41%	47%	*100%	38%	17%	N/A	54%	86%	95%	100%
wetness	*4%	**6%	14%	N/A	N/A	***33%	38%	26%	46%	N/A	80%	100%	100%
shape	2%	***11%	21%	N/A	N/A	25%	N/A	***100%	*14%	***20%	N/A	68%	78%
color	0%	9%	15%	N/A	*6%	*5%	N/A	***2%	*5%	N/A	33%	N/A	87%
material	1%	1%	1%	N/A	***1%	**5%	N/A	N/A	*3%	N/A	22%	13%	N/A

\* less than 5 observations  
 \*\* less than 3 observations  
 \*\*\* only 1 observation  
 N/A no observation

actually flipped compared to the one in Scott (2002); notice that these are exactly the points on the hierarchy where Scott expresses uncertainty about the correct ordering, marked with a question mark in the hierarchy he proposes, see (1) above. Given that the attributive adjective string is generally considered extremely difficult to define within a strictly syntactic model, it is striking that the distance between the two categories on Scott's (2002) hierarchy can be correlated to the reaction time required to intuitively select a more appropriate ordering. The strong negative correlation between average reaction time per combination and the distance of these same categories on Scott's adjectival hierarchy was found with the Pearson correlation coefficient ( $R = -.64, p = .045$ ) plotted in Figure 1.

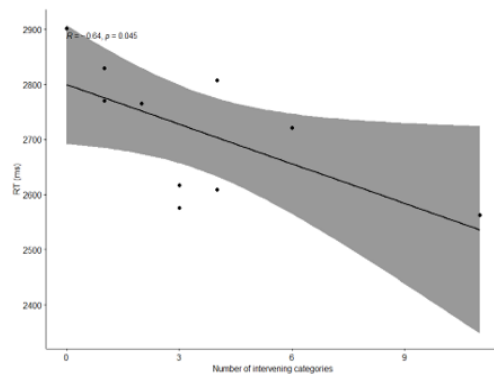


Figure 1: Correlation between average reaction time per combination of two semantic categories and the distance of these same categories on the adjectival hierarchy

**Conclusion.** The results of our Slovenian-based experiment and the results of Plesničar's (to appear) Slovenian corpus study converge in providing support for the adequacy of a strictly syntactic approach to the attributive adjective ordering. Assuming the reversal of the categories DEPTH-WIDTH and WETNESS-AGE in Scott's ordering, as suggested by the results of both of these studies, we also find support for the specific, most-elaborate hierarchy of adjective phrase-related functional projections proposed by Scott (2002).

**Selected references:** • Cinque, G. (2010). *The syntax of adjectives. A comparative study.* MIT Press. • Larson, R. K. (2021). Rethinking cartography. *Language* 97:2, 245-268. • Plesničar, V. (to appear). A corpus investigation of the ordering of selected attributive adjectives in Slovenian. Accepted for publication in *Selected papers from the International Conference of the Department of General Linguistics (BeLiDa 2021)*. Belgrade, Serbia. • Rizzi, L. and G. Cinque. (2016). Functional categories and syntactic theory. *Annual Review of Linguistics* 2, 139-163. • Scott, G.-J. (2002). Stacked adjectival modification and the structure of nominal phrases. In: *The cartography of syntactic structures 1*, (G. Cinque ed.), 91-120.