

ЛИНГВИСТИКА



LINGUISTICS

Guglielmo CINQUE

(Ca' Foscari University, Venice)

THE HIDDEN RULES OF WORD ORDER VARIATIONS

Abstract. *The present article addresses the question of cross-linguistic word order variation, considering the partial regularities found in the languages of the world and the non-existence of certain expected orders. It proposes that movement is at the basis of a restrictive theory of word order variation.*

Keywords: *word order, cross-linguistic variation, regularities, movement*

Резюме. *В статията се разглежда въпросът за словоредните различия между езиците и се обсъждат частичните съвпадения между словоредните модели, както и причините, поради които някои словоредни редици са теоретично възможни, но никога не се реализират в естествения език. Предлага се обяснение от гледна точка на генеративния синтаксис под формата на рестриктивна теория за словоредните различия, в основата на която стои движението на централния конституент на всяка изреченска проекция.*

Ключови думи: *словоред, междуезикови различия, съвпадения между словоредните модели, движение на централния конституент*

1. Introduction

The principal goal of the present article is to sketch a restrictive theory of the variation in word order that is found across languages.¹

For example, a Japanese sentence like (1)b. appears to differ dramatically in the order of words from the corresponding Bulgarian sentence ((1)a.); and this despite the virtual identity in meaning between the two sentences:

¹ The article reproduces the content of an oral presentation of the same general topic at the Collège de France on June 30 2023. I wish to thank Adriana Belletti, Carmen Dobrovie-Sorin, Iliyana Krapova, Salikoko S. Mufwene and Luigi Rizzi for their questions and observations.

- (1)a. Bulgarian (SVO – Iliyana Krapova, pers. comm.)
 Az kazah [če toj može da napravi tova dobre]
 I said [COMP he can do that well]
- b. Japanese (SOV – Endo Yoshio, pers. comm.)
 Watasi-wa [kare-ga sore-o zyoozuni okona-e-ru to] it-ta.
 I-Top [he-Nom it-Acc well do-Mod-Present COMP] say-Past
 ‘I said that he can do it well.’

Apart from the position of the subjects of the main and the embedded clause, which are initial in their respective clauses in both languages, in Bulgarian the finite complement clause follows the matrix verb, is introduced by an initial subordinator, has a modal which precedes the embedded verb, which in turn precedes its object and the manner adverb, while the Japanese matrix verb follows its finite complement clause, which is introduced by a final subordinator, has the modal following the embedded verb, itself preceded by its object and manner adverb: an almost complete mirror image of Bulgarian.

To give another example of the bewildering variation in word order that is found cross-linguistically without any meaning difference, consider some of the different orders of the four elements demonstrative, numeral, adjective and noun, which will be taken up later in more detail:

- (2) **Demonstrative Numeral Adjective Noun** (Bulgarian)
tezi tri hubavi kotki (Iliyana Krapova: pers. comm.)
 ‘these three nice cats’

- (3) **Noun Adjective Numeral Demonstrative**
 (Mbum – Niger-Congo)
yón im húnáké dóà àí (Hagège 1970: 238)
 calabasse beer big two these
 ‘these two big calabasse beers’

- (4) **Noun Demonstrative Numeral Adjective**
 (Abu’ Arapesh – Papuan)
ba-kuh a-kuha bia-kuh afu-kuhi (Lynch 1998: 171)
 stick-CLASS this-CLASS two-CLASS good-CLASS
 ‘these two good sticks’

- (5) **Demonstrative Noun Adjective Numeral**
 (Burmese – Sino-Tibetan)
dí lú jī θòun yaw? (Jones 1970: 5)

these person big three clf.
 ‘these three big persons’

(6) Demonstrative Noun Numeral Adjective

(Maasai – Nilo-Saharan)
kuló payaní oáre tasát’í (Payne 2020: 330)
 these men two old
 ‘these two old men’

When referring to word order one must distinguish *canonical*, or *neutral*, orders (those which do not require special linguistic contexts and can be uttered ‘out of the blue’) from orders which are used to highlight a specific constituent (to put it in focus, or to make it the topic of the discussion). For example, in Bangla (Syed 2015) and Chinese (Zhang 2015), which have, like Bulgarian, Demonstrative > Numeral > Adjective > Noun as a neutral order, adjectives can also appear before the demonstrative for emphatic/focus reasons. See (7):

- (7)a. [**joghonyok** oi du-To biskut (Bangla – cf. Syed 2015: 337)
 disgusting those two-Cla biscuit
 ‘those **disgusting** two biscuits.’
- b. **hong yanse de** na si ge panzi (Chinese – Zhang 2015: 378)
 red color DE that four CL plate
 ‘those four **red** plates’

These orders should not be considered, as they are not neutral orders. Now, the question is: are there principled limits on the variation of the canonical, neutral orders?

To start, one can observe the existence of partial regularities (see §2 below), and, even more importantly, the non-existence of certain orders (see §3) (both have been prominently pointed out by Joseph Greenberg).

2. Partial regularities

Greenberg (1963), and, after him, Dryer (1992) and Sheehan, Biberauer, Roberts, and Holmberg (2017), among others, have shown the existence of partial correlations between the order of the verb with its complements and the order of other pairs. These are stronger within the same extended projection (see (8) to (10))²

² Cf. Dryer (1992: 94, 100, 103).

- (8) a. OV and V modal: 29 genera
b. VO and modal V: 42 genera
c. OV and modal V: 10 genera
d. VO and V modal: 4 genera

- (9) a. OV and VAux: 36 genera
b. VO and AuxV: 28 genera
c. OV and AuxV: 3 genera
d. VO and VAux: 4 genera

- (10) a. OV and IP Subordinator: 38 genera
b. VO and Subordinator IP: 59 genera
c. OV and Subordinator IP: 17 genera
d. VO and IP Subordinator: 1 genus

and weaker across different extended projections (see, for example, (11)):

- (11) a. OV and GEN N: 434 languages
b. VO and N GEN: 352 languages
c. VO and GEN N: 113 languages
d. OV and N GEN: 30 languages

In certain languages the extended projections of the verb (the clause) and that of the noun (the nominal phrase) behave in opposite ways. In the Papuan language Bargam (Hepner 2006: §4.1) the clause is head-final (AdvP Subj PP Obj V) while the nominal phrase is head-initial (N AP NumP DemP). The Mayan language Tzutujil (Dayley 1981: §8.2.3 and §8.1.1) shows the converse situation. The clause is head-initial (V Obj Subj AdvP PP) while the nominal phrase is head-final (DemP NumP AP N).

One should also add that consistent head-final and head-initial languages are actually a minority among the languages of the world, and never totally consistent. Possibly each language differs from every other language in word order type.

3. Non attested orders

Consider now the non-existence of certain orders. In Greenberg (1963) only languages with the orders degree adverb Adj N ((12)a.) (11 langs), N Adj degree adverb ((12)b.) (8 langs) and N degree adverb Adj ((12)c.) (2 langs) are reported as attested (cf. his table 7 and Universal 21).

The order Adj degree adverb N ((12)d.) is also attested, even if rarely (Davies 2020: §2.2), but the others (12)e. and f. are not attested:

- (12)a. degree adverb **A N** (English,..) (e.g. *very tall men*)
 b. **N A** degree adverb (Daai Chin (Sino-Tibetan),..) *aang-ki boo:k sa:* ‘Lit. shirt white very’) (So-Hartmann 2009: 113)
 c. **N** degree adverb **A** (Italian,..) (*un uomo molto alto* ‘Lit. a man very tall’)
 d. **A** degree adverb **N** (Sakha (Turkic),..)(*bu kirakij bayaji deriebine-tten*) ‘Lit. this tiny very village-ABL’ (Stapert 2013: 244)³

e. *A N degree word

f. *degree word N A

Another case, observed by Greenberg (1972: 185), concerns the attested orders of N(oun), Num(eral), and Numeral Classifier (CLF). Of the six potential orders (factorial of 3 = 6), only four are attested:

- (13)a. Num **CLF N** (Chinese,..) (*san ben shu* ‘three CLF book’ (Her 2017a: ex.(1))
 b. **N CLF Num** (Tetun Dili (Austronesian),..)(*feto na’in neen* ‘Lit. woman CLF.HUM six’) (van Engelenhoven and Williams-van Klinken 2005: 758)
 c. **N Num CLF** (Lii (Tai-Kadai),..) (*paa saam too* ‘Lit. fish three CLF’) (Conklin 1981: 108)
 d. **CLF Num N** (Rongga (Austronesian),..) (*esa zhua mbo* ‘CLF two house’) (Arka 2008: 2)

e. *CLF N Num

f. *Num N CLF

We will see other such cases of non-attested orders.

The task of a theory of word order variation is to derive the possible (attested/attestable) word orders, both when they maximally conform to the “head-final” or “head-initial” types and when they depart from them to

³ Also see Chuluu (1994: 29) for an example from the Mongolic language Monguor, and Davies 2020: §2.2 for few more languages that instantiate it.

varying degrees; **and this without deriving the impossible (unattested/unattestable) ones.**

The value of a theory is not how much it allows, but how much it disallows (without excluding what is actually found).

The first step in developing a theory of word order is to see what could be at the basis of it. I argue that the driving force responsible for the different word orders attested is **movement**.

4. An argument that movement is at the basis of word order variation.

The adjective order in the languages of the world has been argued to respect a particular semantic/cognitive hierarchy: adjectives that denote more **absolute** (Martin 1969, Frawley 1992), or **objective** (Hetzron 1978; Scontras, Degen, & Goodman 2017), or **inherent** (Dryer 2018: 816f, and references cited there), properties occur closer to the noun.⁴

Consider the three adjective classes of **color, size and value** (the term ‘value’ used here corresponds to what others call ‘subjective comment’ – Scott 2002 – or ‘quality’):

- (14) **A_{value} A_{size} A_{color} N**
 English (Plank 2003: 11)
A beautiful big red ball

- (15) **N A_{color} A_{size} A_{value}**
 Indonesian (Plank 2003: 11)
bola merah besar tjantik
 ‘Lit. ball red big beautiful’

These orders, which respect the principle of relative distance of the distinct adjectives (**A_{color} A_{size} A_{value}**) from the noun, pre- or post-nominally, are by far the most frequent (listed here are only few of the many languages instantiating these orders):

A_{value} A_{size} A_{color} N

⁴ Hetzron (1978) phrases it as follows: “the major rule is to place the more objective and undisputable qualifications closer to the noun, and the more subjective, opinion-like ones farther away.” (p. 178).

I abstract away here from the questions of how to establish the precise degree of objectivity/subjectivity of each adjectival class, of why this particular hierarchy/order should hold and what principles it could be made to follow from.

Bangla (Syed 2015:337); **Hungarian** (Hetzron 1978:70,73); the Papuan languages **Awa** (Loving 1973:§4.2), **Gimi** (McBride and McBride 1973:83), **Kewapi** (Yarapea 2006:154); the Tibeto- Burman language **Magar** (Grunow-Hårsta 2008:364), **Turkish** (Bayırlı 2018:§2), etc.


$N A_{color} A_{size} A_{value}$

The Austronesian language **Javanese** (Ishizuka 2008:§3.3); **Basque** (Artiagoitia 2006:§1); the Kwa language **Akan** (Afriyie 2014:§4.2); the Papuan languages **Golin** (Bunun 1974: §2.2.1.1), **Maybrat** (Dol 1999:142f), **Urim** (Hemmilä and Luoma (1987:122); the Semitic languages **Soqotri** (Makhashen, Shuib and Che Lah. 2008:13), **Zahrani Arabic** (Alzahrani 2015:234); the Tibeto-Burman language **Lolo** (Fu 1997:191); the Kam-Tai language **Tai Phake** (Morey 2005:260); the Niger- Congo language **Yorùbá** (Ajíbóyè 2005:16); the language isolate **Trumai** (Guirardello 1999:17), etc.

Nonetheless there are few languages with post-nominal adjectives that seem to violate this principle (see (16)). This was noted by the eminent typologist Frans Plank, who also noted that this principle (which is never violated pre-nominally (see (17)) could be retained if one assumed movement of the noun, as shown in (18):

(16) $N A_{value} A_{size} A_{color}$
 Maltese (Plank 2003:12)⁵
balloon sabiħ kbir aħmar
 ‘Lit. ball beautiful big red’

(17)* $A_{color} A_{size} A_{value} N$ (Plank 2003:12)

(18) $N A_{value} A_{size} A_{color} N$


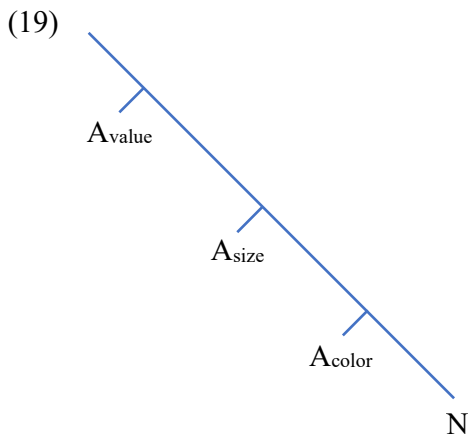
“It is only on such an abstract analysis that the iconic explanation of the relative distance of semantic classes of adjectives from nouns can be upheld.” (Plank 2003:13).

Of course the movement approach involved in the derivation of these different orders must be constrained to derive the actually attested

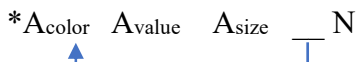
⁵ This order is also found in the Niger-Congo, Kwa, language Leteh (Akrofi Ansah 2014:12).

orders without deriving the unattested ones (such as the *A_{color} A_{size} A_{value} N of (17) above, or other orders, like the unattested pre-nominal orders: *A_{size} A_{color} A_{value} N, or *A_{value} A_{color} A_{size} N.

The restriction on movement which appears to discriminate between the possible (attested) from the impossible (unattested) orders is that only the Head of the constituent, the ‘engine’ of the movement, here the N, can move, along the hierarchy which respects the relative distance from the noun:



If the adjectives could move independently of the N we would get unattested orders, like



If only the Head of the nominal constituent, the N, can move, what are the possible ways it can move?

From *wh*-movement (the movement of a relative or interrogative category) we know that the ‘engine’ of the movement, here the *wh*-category, can move by itself, as in (20)a., or by dragging along a constituent containing non-*wh*-material below the N, as in (20)b., or above it, as in (20)c.:

- (20)a. Aylan, [who] you have certainly seen [pictures of []],..
- b. Aylan, [[whose] pictures] you have certainly seen [],..
- c. Aylan, [pictures of [whom]] you have certainly seen [],..

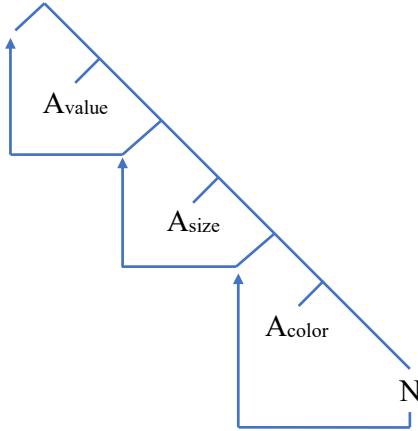
The movements in (20)b. and c. are called pied pipings (recalling the pied piper of Hamelin): *whose-pictures* and *pictures-of-whom* pied piping, respectively.

N can move likewise: by itself, of via one of the pied piping modes.

If it moves to the top of the hierarchy by itself (without pied pipings), as in (21), the order that is obtained is the order of Maltese: N

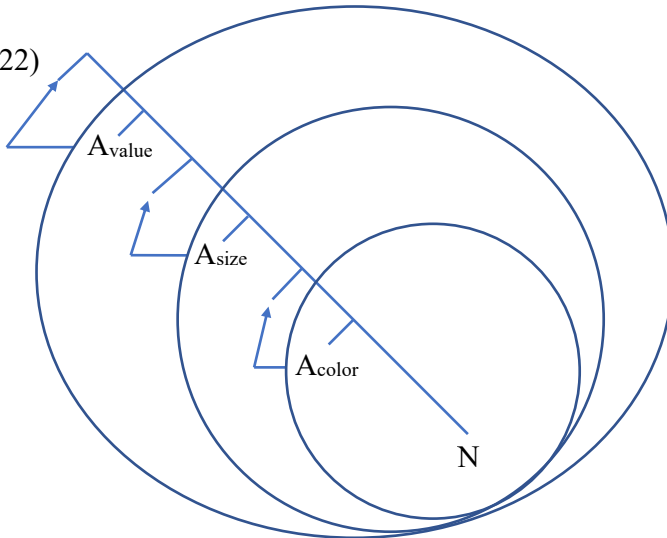
$A_{value} A_{size} A_{color}$:

(21)

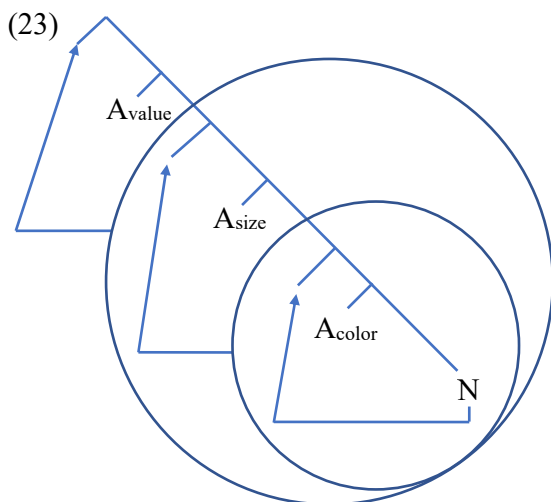


If it moves in the *pictures-of-whom* pied piping mode, whereby the noun drags along material immediately above it (see (22)), the order that is obtained is the order of Bangla and other head-final languages: $A_{value} A_{size} A_{color} N$, in which the relative scope, and order, of the different adjectives is not altered.

(22)



If the noun moves in the *whose-pictures* pied piping mode, as in (23), the order that is obtained is N A_{color} A_{size} A_{value}, the mirror-image of the order of Bangla and other consistent head-final languages, in which the order of the different adjectives is altered (this is apparently the most frequent order):

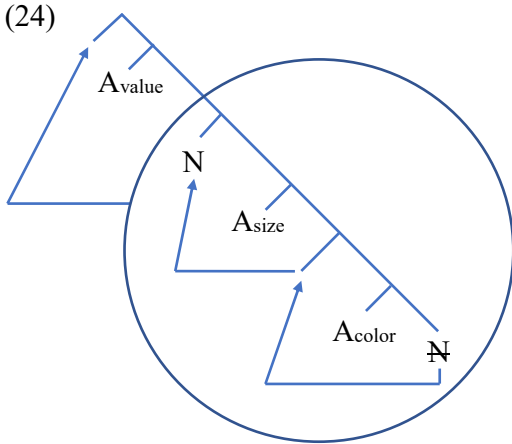


The picture is actually more complex, but in ways that do not affect the conclusion drawn above.

While pre-nominally only one order is possible ($A_{value} > A_{size} > A_{color} > N$), post-nominally more than two orders are actually possible. In addition to $N > A_{value} > A_{size} > A_{color}$ and $N > A_{color} > A_{size} > A_{value}$ other orders are found, which can only be derived by movement if we want to retain the principle of relative distance of the adjectives from the noun, just as we have seen with the Maltese order $N A_{value} A_{size} A_{color}$.

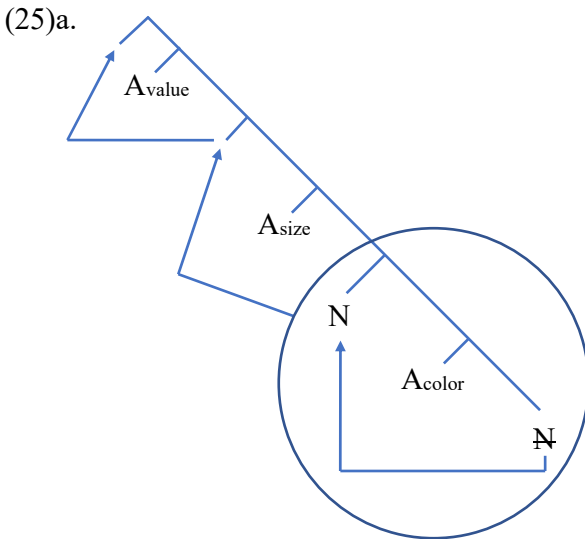
For example the order $N A_{size} A_{color} A_{value}$ of the Austronesian language Tatana' (Dunn and Peck 1988:211)⁶ (as well as of the Niger-Congo, Gur, language Kasem – Danti 2007:122, and Welsh –Flanagan 2014:§8.2.2) can be derived by first moving the noun by itself, without pied piping, above A_{color} and above A_{size} , and then moving the whole constituent [$N A_{size} A_{color}$] above A_{value} via the *whose-pictures* pied piping option:

⁶ “Any change in the order of modifiers was quickly rejected by speakers of the language, indicating that they had a definite order for the adjectives.” (Dunn and Peck 1988:211).



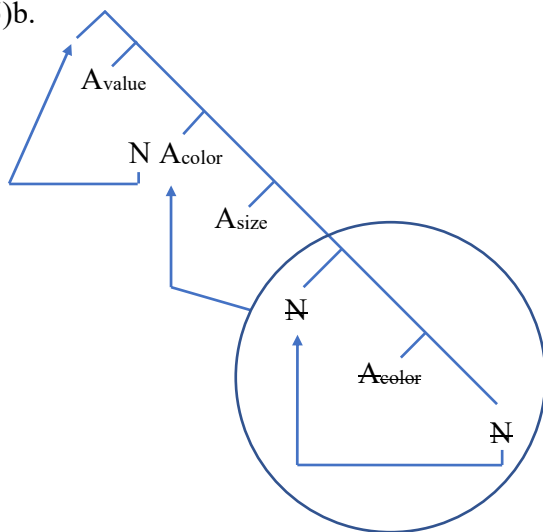
The order **N A_{color} A_{value} A_{size}** of Selepet (Papuan, Finisterre-Huon – McElhanon1972: 81; Dixon 1982: 26fn27) and of the West Caucasian language Abkhaz (Hewitt 1989: 59), as well as the order **N A_{value} A_{color} A_{size}** of Umbu-Ungu (Papuan, Chimbu – Head 1976: 67) are two more orders which must be derived by movement if we want to retain the principle of relative distance of **A_{value} A_{size} A_{color}** from the noun.

For the order **N A_{color} A_{value} A_{size}** the noun moves above **A_{color}** and then the whole constituent **[N A_{color}]** moves by itself, without pied piping, above **A_{size}** and **A_{value}** (see (25)a.):



For the order $N A_{value} A_{color} A_{size}$ the noun must move step-wise with the *whose-pictures* pied piping mode above A_{color} and A_{size} , reversing their order, and then move by itself above A_{value} (see (25)b.):

(25)b.



This approach to word order variation in the cross-linguistic ordering of adjectives may appear not to be the only account possible, but there are several considerations that support it:

First: It allows to derive all of the attested orders from one and the same hierarchy that respects the semantic/cognitive principle of relative distance of modifiers from the Head (Frans Plank's point).

Second: With the important restriction that only the Head of the projection, here the Noun, can move, in one of the possible ways that movement can take place, it allows the discrimination of possible from impossible orders.

Third: The movement options that it involves are, as we have seen, independently motivated by other clear cases of movement (like *wh*-movement).

Fourth (and foremost): The movement approach naturally carries over to the different syntactic domains that display the same pattern of possible and impossible orders (one order to the left and several orders to the right of a Head).

Other domains displaying the same pattern (of a single order to the left of the Head, and (at least) two orders to the right of the Head) are, for

example, the order of adverbs (see (26)), and that of circumstantial PPs (see (27)):

Order of adverbs: (Cinque 1999: 42f; Rackowski and Travis 2000; Pearson 2000)

- (26)a. Adv_{no longer} > Adv_{always} > Adv_{completely} > V (English, Chinese,..)
 b. *Adv_{completely} > Adv_{always} > Adv_{no longer} > V 0
 c. V > Adv_{no longer} > Adv_{always} > Adv_{completely} (Italian, (main clause) German,..)
 d. V > Adv_{completely} > Adv_{always} > Adv_{no longer} (Malagasy, Niuean,..)

Limiting ourselves to circumstantial Time, Place and Manner PPs, whose order has been investigated from a cross-linguistic perspective by Boisson (1981), and Lu and Wen (2022) (also see Hinterhölzl 2001, Schweikert 2005, Cinque 2006), we find exactly the same pattern:⁷

Order of circumstantial PPs

- (27)a. Time > Place > Manner > V (Basque – Lu and Wen 2022:399)⁸
 b. *Manner > Place > Time > V 0
 c. V > Time > Place > Manner (Otomí (Oto-Manguéan) – Boisson 1981:76)⁹
 d. V > Manner > Place > Time (Vietnamese, Yorùbá – Lu and Wen 2022:399)

The same pattern is also attested in the order of Demonstrative, Numeral, Adjective, and Noun (Greenberg 1963; Cinque 1996, 2005, 2023, among many others)¹⁰

- (28)a. Dem > Num > A > N (English, Malayalam,..)
 b. *A > Num > Dem > N 0
 c. N > Dem > Num > A (Abu‘ Arapesh, Kikuyu,..)
 d. N > A > Num > Dem (Gungbe, Thai,..)

⁷ On the irrelevant interference of focus on the canonical order of circumstantial PPs and possible diagnostics for the canonical order, see Schweikert (2005), Cinque (2006), Lu and Wen (2022: fn3).

⁸ Another language displaying this order is Nambikuára (Nambiquaran, Brazil – Kroeker 2001:3)

⁹ Another language displaying this order is Tanga (Melanesian – Bell 1977:x)

¹⁰ For additional domains in which one finds the same pattern I refer to Cinque (2009).

This case is even more telling as we have exhaustive evidence of all the possible and impossible combinations of these four elements (from a sample of over 2200 languages).

Out of the 24 mathematically possible orders (factorial 4) only 14 are actually attested (Cinque 2005, 2023).¹¹ See (29):

(29)	[languages]	[genera]
a. N A Num Dem	631	136
b. Dem Num A N	446	115
c. Dem N A Num	204	89
d. Dem Num N	186	76
e. Num N A Dem	242	50
f. N A Dem Num	103	35
g. N Num A Dem	70	34
h. Dem N Num A	52	29
i. Dem A N Num	50	27
l. N Dem Num A	83	25
m. Num A N Dem	55	21
n. N Dem A Num	33	17
o. A N Num Dem	33	13
p. A N Dem Num	21	9
	Tot:	2.209

The boldfaced orders (29)a., b., l are derived in the same way as the orders seen above of N A_{color} A_{size} A_{value}, A_{value} A_{size} A_{color} N, and N A_{value} A_{size} A_{color}. The remaining 11 orders of the 14 attested ones (those not in boldface in (29)) are derivable via different combinations of the three possible movement options seen above. See Cinque (2005, 2023).

¹¹ Dryer (2018) claims that some of the orders that Cinque's (2005) account of Greenberg's (1963) Universal 20 ruled out are actually attested in at least few languages. But this does not appear to be right as these languages also have orders which are among the 14 allowed in Cinque's (2005) account. A real counterexample would be a language with only one order which is not among the 14 admitted. See Cinque (forthcoming) for relevant discussion.

Crucially, the 10 orders which are unattested (those in (30)) are underivable under the condition that only the Head N(P) can move (by itself or in one of the two pied piping modes):

- (30)a. *Dem A Num N
 b. *Num A Dem N
 c. *A Num Dem N
 d. *Num Dem A N
 e. *A Dem Num N
 f. *A Num N Dem
 g. *A Dem N Num
 h. *Num Dem N A
 i. *Num N Dem A
 l. *N Num Dem A

The restriction that only the Head of the projection can move within its projection thus proves instrumental in deriving the 14 attested orders (out of the 24 potential ones) of demonstrative, numeral, adjective and noun without deriving the 10 unattested ones.¹²

The restriction that **only** the Head of a projection can move within its projection can also account for the two missing orders of Degree Adverb A N and Numeral Classifier N ((12) and (13), seen above, and repeated here without examples and references as (31) and (32)), under the plausible assumptions that Degree Adverb and A and Numeral and Classifier form constituents that modify the N:¹³

- (31)a. [[degree adverb A] N]
 b. [N [A degree adverb]]
 c. [N [degree adverb A]]
 d. [[A degree adverb]N]
-
- e. *A N degree adverb
 f. *degree adverb N A

¹² As already noted, adjectives, numerals and demonstratives cannot move by themselves, except, irrelevantly, for focus reasons (or for scope reasons, as with the movement of English APs to a superlative high position in front of numerals, as in *The black*(est) two dogs that I've (ever) seen* – Kayne 2008:fn15).

¹³ For the case of Numeral and Numeral Classifier see Greenberg (1975:29).

- (32)a. [[Numeral **Classifier**] N]
 b. [N[**Classifier** Numeral]]
 c. [N [Numeral **Classifier**]]
 d. [[**Classifier** Numeral] N]
-

- e. *Classifier N Numeral
 f. *Numeral N Classifier

A movement approach to word order also promises to shed light on the numerosity of certain orders and the rarity of other orders. Recall the three modes in which the noun can move: by itself and via the *whose-pictures* and the *pictures-of-whom* pied pipings.

That the *pictures-of-whom* pied piping mode is more marked than the *whose-pictures* pied piping mode is shown, for example, by the following contrasts in English *wh*-interrogatives and restrictive relatives:

- (33)a. Whose pictures do you keep in your wallet?
 b. *?Pictures of whom do you keep in your wallet?

- (34)a. I wonder whose pictures they published yesterday.
 b. *I wonder pictures of whom they published yesterday.

- (35)a. Here is someone whose pictures have always appealed to her.
 b. *Here is someone pictures of whom have always appealed to her.

This may be at the basis of the fact that orders derived by the *whose-pictures* pied piping mode are more frequent than the orders derived by the *pictures-of-whom* pied piping mode. See for example the fact that the languages and genera of the order N A Num Dem (which is derived via a consistent application of the same *whose-pictures* pied piping mode) outnumber those with the order Dem Num A N (derived with a consistent application of the same *pictures-of-whom* pied piping mode). See (29) above.¹⁴

¹⁴ The orders that mix the three different modes in which N(P) moves (the non-boldfaced orders of (29)) are progressively rarer, possibly depending on the different costs associated with the various types of mixtures (movement without pied piping appears to be the most marked option).

This is plausibly also at the basis of the fact that orders derived through a change from a *whose-pictures* pied piping (the less marked one) in an inner projection to a *pictures-of-whom* pied piping (the more marked one) in an outer projection is more costly (hence rarer) than changing mode in the opposite direction (from a *pictures-of-whom* to a *whose-pictures* pied piping mode).

In the orders of degree adverb Adj and N seen above, there is a decline in numerosity. (36)a., with a consistent application of the *whose-pictures* pied piping and (36)b., with a consistent application of the *pictures-of-whom* pied piping are more frequent; then follows (36)c., which involves a change from a *pictures-of-whom* pied piping to the less marked *whose-pictures* pied piping. (36)d., which involves a switch from a *whose-pictures* pied piping to the more marked *pictures-of-whom* pied piping, is the absolute rarest (recall that Greenberg did not record its existence):

- (36)a. [[N [Adj degree adverb] N] (both Adj and N move with the *whose-pictures* pied piping)
- b. [[degree adverb Adj] N] (both Adj and N move with the *pictures-of-whom* pied piping)
- c. [N [degree adverb Adj]] (Adj moves with the *pictures-of-whom* pied piping in the inner projection and N moves with the *whose-pictures* pied piping in the outer projection)
- d. [[Adj degree adverb] N] (Adj moves in the inner projection with the *whose-pictures* pied piping and in the outer projection N moves with the *pictures-of-whom* pied piping)

The same is true with the orders of Numeral, Classifier, and N:

(37)d. which involves a change from a *whose-pictures* pied piping mode (the unmarked one) to a *pictures-of-whom* pied piping mode (the marked one), is the absolute rarest.

- (37)a. [[Num CLF] N] (both N and CLF move with *pictures-of-whom* pied piping mode)
- b. [N [CLF Num]] (both N and CLF move with *whose-pictures* pied piping mode)
- c. [N [Num CLF]] ((CLF moves in the inner projection with the *pictures-of-whom* pied piping mode and N in the outer projection with the *whose-pictures* pied piping mode)
- d. [[CLF Num] N] ((CLF moves in the inner projection with the *whose-pictures* pied piping mode and N in the outer projection with the *pictures-of-whom* pied piping mode)

Change of mode from the less marked to the more marked pied piping may also explain the extreme rarity of the exceptions to two well-known constraints discussed in the literature, the Final-Over-Final Condition (Holmberg 2000 and Sheehan, Biberauer, Roberts, and Holmberg 2017) and the Head-Final Filter (Williams 1982), which display the exact same pattern.

The Final-Over-Final Condition was proposed to disallow structures where a head-initial phrase is contained in a head-final phrase in the same projection, though certain exceptions exist (see (38)d., (39)d. and fns.15 and 16, 17.

- (38)a. **[[O V] AUX]** (Hindi, Turkish, Yukaghir, etc.)
 b. **[AUX [V O]]** (English, Italian, Niuean, etc.)
 c. **[AUX [O V]]** (Guébie, Tunen, West Flemish, etc.)
 d. **%[[V O] AUX]** (Gumuz, Kokama-Kokamilla)¹⁵
- (39)a. **[[O V] COMP]** (Japanese, Malayalam, Marathi, etc.)
 b. **[COMP [V O]]** (Arabic, English, Italian, etc.)
 c. **[COMP [O V]]** (German, Georgian, Hindi, etc.)
 d. **%[[V O] COMP]** (Hkongso¹⁶; East !Xóõ¹⁷)

The same holds for the Head-Final Filter, which was proposed to rule out cases like (40)d., ungrammatical in English.¹⁸ But that case is not unattested. It is only extremely rare cross-linguistically (note that we have here exactly the same pattern as the one seen so far):

¹⁵ For references and discussion see Cinque (2023:§5.2). The 4 genera which are reported as instantiating VOAux in Dryer (1992:100) should also be looked at.

¹⁶ “The sole instance in my database of a VO language with final complementizer is Hkongso, a Tibeto-Burman language of Burma.” (Dryer 2012:76fn8).

¹⁷ Güldemann (2004:7) reports a sentence that exemplifies the order V O C, confirming in personal communication (January 16, 2010) that “the language indeed is an exception to the supposed universal SVO → initial complementizer”.

¹⁸ Williams’ (1982) original proposal was meant to rule out in English cases like **[a proud of his children] man*, which is “unacceptable because the pronominal modifier (*proud of his children*) does not end in its head (*proud*.” (p.160) [as opposed to the corresponding Dutch case (40)b., which is grammatical]. Other names for the phenomenon are the “recursion restriction” (Emonds1976), the “consistency principle” (Giorgi and Longobardi 1991), the “adjacency requirement” (van Riemsdijk 1998), and the “edge effect” (Haider 2000).

- (40)a. [N [A PP]] (a [man [proud of his children]]) (English)
 b. [[PP A] N] (een [[op zijn vader] trotse] man])
 (Dutch) (Neeleman 1994:233)
 ‘Lit. A of his father proud man’
- c. [N [PP A]] madar-an-e [[be farzand-an-e xod] moftaxar]
 (Persian)
 ‘Lit. mother-PL-LNK [in child-PL-LNK own] proud’
 (from Alexeyenko and Zeijlstra 2021:53 after Zahra Mirrazi, p.c.)
- d. [[A PP] N] [gordiyat [sās svoeto dete]] băsta]
 (Bulgarian) (Iliyana Krapova, p.c.)
 ‘Lit. proud.the with SELF.the child father’
 ‘the father proud of his child’

5. A restrictive theory of word order variation.

In the end, let me sketch how the requirement that only the Head of a projection can move within its projection (by itself or with pied piping) may contribute to build a restrictive theory of (canonical) word order (recall that a restrictive theory is measured in how much it excludes without excluding what is actually found).

Consider the addition of a fifth element, the numeral classifier, to the elements demonstrative, numeral, adjective and noun.

Factorial 5 (demonstrative, numeral, numeral classifier, adjective and noun) is $1 \times 2 \times 3 \times 4 \times 5 = 120$ possible combinations. But if only 14 orders of demonstrative, numeral, adjective and noun are possible (as we saw in (29) and (30) above), and we multiply this number by two, due to the possibility for each of the 14 orders to vary according to whether the numeral classifier Head comes to precede or to follow its modifier (the numeral), we get not 120 but just 28 possible orders.

See (41) below, where only two of the 28 expected orders are not (yet) documented, no other order being found distinct from those in (41), as far as I could ascertain. (I refer to Cinque 2023 for reference to the relevant sources for each language mentioned).

- (41)a. Dem Num CLF A N (Mandarin and Cantonese - Sinitic)
 a’ Dem CLF Num A N (Upper Necaxa Totonac - Totonacan)
- b. Dem Num CLF N A (Yao - Hmong-Mien)

- b' Dem CLF Num N A (Nêlêmwa and Zuanga - Oceanic)
- c. Dem N Num CLF A (Lahu - Sino-Tibetan)
- c' Dem N CLF Num A (Hakha Chin - Sino-Tibetan)
- d. N Dem Num CLF A (Stiêng - Mon-Khmer)
- d' N Dem CLF Num A (Kiriwina - Oceanic)
- e. A N Dem Num CLF (Yunnan Bai - Sino-Tibetan)
- e' A N Dem CLF Num(?)**
- f. N A Dem Num CLF (Kayan Lahta - Sino-Tibetan)
- f' N A Dem CLF Num (Awara and West Makian - Papuan)
- g. Dem A N Num CLF (Newari and Dulong - Sino-Tibetan)
- g' Dem A N CLF Num (Mising and Nyishi - Sino-Tibetan)
- h. Dem N A Num CLF (Burmese and Maru - Sino-Tibetan)
- h' Dem N A CLF Num (Apatani - Sino-Tibetan)
- i. N Dem A Num CLF (Nias Selatan – Malayo-Polynesian)
- i' N Dem A CLF Num (Diola-Fogny – Atlantic-Congo)
- l. Num CLF A N Dem (Tojolab'al – Mayan)
- l' CLF Num A N Dem (Q'anjob'al – Mayan)
- m. Num CLF N A Dem (Vietnamese – Austro-Asiatic)
- m' CLF Num N A Dem (Rongga – Malayo-Polynesian)
- n. N Num CLF A Dem (Kele and Lele – Oceanic)
- n' N CLF Num A Dem (Buglere and Teribe – Chibchan)
- o. A N Num CLF Dem (?)**
- o' A N CLF Num Dem (Galo and Mising-Sino-Tibetan)
- p. N A Num CLF Dem (Thai and Lao-Tai-Kadai)
- p' N A CLF Num Dem (Abun – Papuan; Sudest-Oceanic)

The addition of another element, by taking Num to be a complex cardinal numeral (e.g. 'three hundred'), composed of a multiplier ('three') and a base ('hundred'), would yield 720 possible combinations (factorial 6). Yet, if the multiplier and base form a constituent (cf. Kayne 2005: §9, and others), and if only the base can move within its constituent as it is the

Head of that constituent, at most 56 (rather than 720) orders are expected to be attested (the 28 possibilities of (41) multiplied by the two possible orders between base and multiplier in each of the 28 orders). This expectation awaits to be documented.

The factorial numbers stemming from the combinations of n elements (which with 10 elements arrive at over three and a half millions of possible orders if no restrictions are introduced) are drastically reduced (to less than one thousand) if word order variation is confined within each projection, and if only the Head of that projection can move within it in one of the possible ways.

6. Conclusion: the hidden rules of word order variation.

The hidden rules of word order variation to which I have referred in the title can thus be summarized in the following three points:

- (42)a. (Canonical) word order variation is a function of movement, in one of the forms that movement can take place (with or without pied piping).
- b. The Head, and only the Head, of each (sub-)projection (N(P), A(P), V(P), CLF(P), Num(P), etc.) can move within its projection.
- c. Kayne's (1994) Linear Correspondence Axiom, which applies to hierarchies (linearizing higher constituents to the left of lower ones).

REFERENCES

- Afriyie 2014: Afriyie, Juliana. 2014. *Syntagmatic Relations of Noun Modifiers in Akan*. MPhil. Thesis, University of Ghana. http://ugspace.ug.edu.gh/bitstream/123456789/7494/1/Juliana%20Afriyie_%20%20Syntagmatic%20Relations%20of%20Noun%20Modifiers%20in%20Akan_2014.pdf
- Ajibóyè 2005: Ajibóyè, Oládíipò Jacob. 2005. *Topics in Yoruba Nominal Expressions*. Ph.D. Dissertation, University of British Columbia. <https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0092961>
- Akrofi Ansah 2014: Akrofi Ansah, Mercy. 2014. The Morphosyntax of the Leteh Simple Noun Phrase. *California Linguistic Notes* 39 https://english.fullerton.edu/_resources/pdfs/CLN_MercyAnsahPAGINATED.pdf
- Alexeyenko and Zeijlstra 2021: Alexeyenko, Sascha and Hedde Zeijlstra. 2021. (Dis)obeying the Head-Final Filter. *Proceedings of the 37th West Coast Conference on Formal Linguistics*. 50 – 58. Somerville, MA: Cascadilla.

- <http://www.lingref.com/cpp/wccfl/37/paper3514.pdf> (ISBN 978-1-57473-477-5)
- Alzaharani 2015: Alzaharani, Salih Jamaan. 2015. *Topics in the Grammar of Zahrani Spoken Arabic*. Ph.D. Dissertation, Newcastle University, Australia. <https://nova.newcastle.edu.au/vital/access/manager/Repository/uon:19763>
- Arka 2008: Arka, I Wayan. 2008. Classifiers in Rongga. *Studies in Philippine Languages and Cultures*. 18: 1 – 17 https://www.sil.org/system/files/reap_data/13/59/13/135913778944838096119610787285170940952/1_Arka.pdf (ISSN 0119-6456)
- Artiagoitia 2006: Artiagoitia, Xabier. 2006. Basque adjectives and the functional structure of the noun phrase. Ms., University of the Basque Country. <https://artxiker.ccsd.cnrs.fr/file/index/docid/87305/filename/adjective-heads.pdf>
- Bayırlı 2018: Bayırlı, İsa Kerem. 2018. Does Turkish have adjective ordering restrictions? *IULC Working Papers* 18 (2): 1 – 26 <https://scholarworks.iu.edu/journals/index.php/iulcwp/article/view/26051> (ISSN 1524-2110)
- Bell 1977: Bell, F. L. S. 1977. *Tanga-English, English-Tanga Dictionary* (Oceania Linguistic Monographs, 21). Sidney: University of Sydney. <https://archive.org/details/tangaenglishengl0000bell>
- Boisson 1981: Boisson, Claude Pierre. 1981. Hiérarchie universelle des spécifications de temps, de lieu, et de manière. *Confluents* (Centre d'études Anglaises et Nord-américaines, Université Lyon 2) 7 (2): 69 – 126.
- Bunn 1974: Bunn, Gordon. 1974. Golin grammar. *Workpapers in Papua New Guinea languages* 5. Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/23729>
- Chuluu 1994: Chuluu, Üjiyediin. 1994. Introduction, Grammar, and Sample Sentences for Monguor. *Sino-Platonic Papers* (Department of East Asian Languages and Civilizations, University of Pennsylvania) 57: 1 – 31 http://sino-platonic.org/complete/spp057_monguor_language.pdf
- Cinque 1996: Cinque, Guglielmo. 1996. The antisymmetric programme: theoretical and typological implications. *Journal of Linguistics* 32:447-464 (ISSN 0022-2267)
- Cinque 1999: Cinque, Guglielmo. 1999. *Adverbs and Functional Heads. A cross-linguistic perspective*. New York: Oxford University Press. (ISBN 9780195115277)
- Cinque 2005: Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry* 36: 315 – 332 (ISSN: 0024-3892) <https://iris.unive.it/handle/10278/30426>
- Cinque 2006: Cinque, Guglielmo. 2006. Complement and Adverbial PPs: Implications for Clause Structure. In G. Cinque *Restructuring and Functional Heads*. 145 – 166. New York: Oxford University Press. (ISBN 9780195179545)

- Cinque 2023: Cinque, Guglielmo. 2023. *On Linearization. Toward a Restrictive Theory*. Cambridge, MA: MIT Press. <https://direct.mit.edu/books/book/5553/On-LinearizationToward-a-Restrictive-Theory> (ISBN 978-026-254-495-5).
- Cinque forthcoming: Cinque, Guglielmo. (forthcoming) *Again on the Order of Demonstrative, Numeral, Adjective, and Noun*.
- Conklin 1981: Conklin, Nancy Faires. 1981. *The Semantics and Syntax of Numeral Classification in Tai and Austronesian*. Ph.D. Dissertation, The University of Michigan.
- Danti 2007: Danti, Alexis L. 2007. The Kasem Adjective Phrase Structure. In M. E. Kropp Dakubu, G. Akanlig-Pare, E. K. Osam and K. K. Saah, eds., *Studies in the Languages of the Volta Basin 4*. 116 – 126. Legon: Linguistics Department, University of Ghana.
- Davies 2020: Davies, Imogen 2020. *The use of 'as' as a post-adjectival intensifier in British English*. M.A. Thesis, University of Cambridge. https://lagb.org.uk/resources/Documents/Davies_Diss.pdf
- Dixon 1982: Dixon, R. M. W. 1982. *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: Mouton. (ISBN: 978-902-793-309-6)
- Dol 1999: Dol, Philomena Hedwig. 1999. *A grammar of Maybrat: a language of the Bird's head, Irian Jaya, Indonesia*. Rijksuniversiteit te Leiden.
- Dryer 1992: Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68: 81 – 138 (ISSN 00978507) <http://www.acsu.buffalo.edu/~dryer/DryerGreenbergian.pdf>
- Dryer 2018: Dryer, Matthew S. 2018. On the Order of Demonstrative, Numeral, Adjective, and Noun. *Language* 94 (4): 798 – 833 (ISSN 00978507)
- Dunn and Peck 1988: Dunn, Phyllis and Charles Peck. 1988. Noun Phrases in Tatana'. In C. Peck, ed., *Borneo Language Studies I: Sabah Syntax Papers*. 206-227. Dallas: Summer Institute of Linguistics.
- Emonds 1976: Emonds, Joseph. 1976. *A Transformational Approach to English Syntax: Root, structure preserving and local transformations*. New York: Academic Press. (ISBN-10 0122385500)
- Evans 2022: Evans, Jonathan P. 2022. Classifiers before numerals. Origin of an atypical morpheme order. *Linguistics of the Tibeto-Burman Area* 45: 1 – 21 (ISSN: 0731-3500) https://www.academia.edu/56600295/CLASSIFIERS_BEFORE_NUMERALS_ORIGIN_OF_AN_ATYPICAL_MORPHEME_ORDER1?auto=download&email_work_card=download-paper
- Flanagan 2014: Flanagan, Paul James. 2014. *A Cross-Linguistic Investigation of the Order of Attributive Adjectives*. Ph.D. Dissertation, Edge Hill University. https://research.edgehill.ac.uk/ws/portalfiles/portal/20511511/Flanagan_Paul_Thesis_Input_Final_2014_03.pdf
- Frawley 1992: Frawley, William. 1992. *Linguistic semantics*. Hillsdale: Lawrence Erlbaum. (ISBN 9780805810752)

- Fu 1997: Fu, Maoji. 1997. A descriptive grammar of Lolo. *Linguistics of the Tibeto-Burman Area* 20 (1): 1 – 242. (ISSN: 0731-3500)
- Giorgi and Longobardi 1991: Giorgi, Alessandra and Giuseppe Longobardi. 1991. *The Syntax of Noun Phrases*. Cambridge: Cambridge University Press. (ISBN: 9780521379021)
- Greenberg 1963: Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed., J.Greenberg, 73 – 113. Cambridge (Mass.): MIT Press. (ISBN 9780262570084)
- Greenberg 1975: Greenberg, Joseph H. 1975. Dynamic aspects of word order in the numeral classifier. In C. N. Li, ed., *Word Order and Word Order Change*. 28 – 45. Austin: University of Texas Press. (ISBN 0292790023)
- Grunow-Hårsta 2008: Grunow-Hårsta, Karen A. 2008. *A Descriptive Grammar of Two Magar Dialects of Nepal: Tanahu and Syangja Magar. Vol. I*. Ph.D. Dissertation, University of Wisconsin-Milwaukee.
- Guirardello 1999: Guirardello, Raquel. 1999. *A reference grammar of Trumai*. Ph.D. Dissertation, University of Houston.
- Hagège 1970: Hagège, Claude. 1970. *La langue Mbum de Nganha (Cameroun)*. *Phonologie-Grammaire II*. Paris: SELAF.
- Haider 2000: Haider, Hubert. 2000. OV is more basic than VO. In P.Svenonius, ed., *Derivation of VO and OV*. 45 – 67. Amsterdam: Benjamins. (ISBN 9789027227522)
- Head 1976: Head, June. 1976. *A grammar of Umbu-Ungu*. Ukarumpa: Summer Institute of Linguistics. <https://pnglanguages.sil.org/resources/archives/42774>
- Hemmilä and Luoma 1987: Hemmilä, Ritva and Pirkko Luoma. 1987. *Urim Grammar*. Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/31244>
- Hetzron 1978: Hetzron, Robert. 1978. On the Relative Order of Adjectives. In H. Seiler, ed., *Language Universals*. 165 – 184. Tübingen, Gunter Narr Verlag. (ISBN-10 3878081111)
- Hewitt 1989: Hewitt, George. (in collaboration with Z. K. Khibia). 1989. *Abkhaz*. London: Routledge.
- Hinterhölzl 2001: Hinterhölzl, Roland. 2001. Event-related Adjuncts and the OV/VO Distinction. In *Proceedings of the 20th West Coast Conference on Formal Linguistics*. 276–289. Somerville, Mass.: Cascadilla Press. (ISBN 978-1-57473-143-9)
- Holmberg 2000: Holmberg, Anders. 2000. Deriving OV order in Finnish. In P.Svenonius, ed. *Derivation of VO and OV*. 123 – 152. Amsterdam: Benjamins. (ISBN 9789027227522)
- Ishizuka 2008: Ishizuka, Tomoko. 2008. Deriving the order of constituents in the Javanese DP. Ms., University of California, Los Angeles. <http://ling.auf.net/lingBuzz/000769>

- Jones 1970: Jones, Robert B. 1970. Classifier constructions in Southeast Asia. *Journal of the American Oriental Society* 90:1 – 12 (ISSN 00030279)
- Kayne 1974: Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press. (ISBN 9780262611077)
- Kayne 2005: Kayne, Richard S. 2005. Some notes on comparative syntax, with special reference to English and French. In G. Cinque and R. Kayne, eds., *Oxford Handbook of Comparative Syntax*. 3 – 69. New York: Oxford University Press. (ISBN: 9780195136517)
- Kayne 2008: Kayne, Richard S. 2008. Some Preliminary Comparative Remarks on French and Italian Definite Articles. In R. Freidin, C. P. Otero, and M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. 291 – 321. Cambridge, Mass.: MIT Press (reprinted in 2010 in R.S.K. *Comparison and Contrasts*. 3 – 28. New York: Oxford University Press). <https://as.nyu.edu/content/dam/nyu-as/linguistics/documents/kayne-publications/Kayne1004DefiniteArticles.pdf> (ISBN 978-026-206-278-7)
- Kroeker 2001: Kroeker, Menno. 2001. A Descriptive Grammar of Nambikuara. *International Journal of American Linguistics* 67: 1 – 87 (ISBN 0020-7071)
- Loving 1973: Loving, Richard. 1973. An Outline of Awa Grammatical Structures. In H. McKaughan, ed., *The Languages of the Eastern Family of the East New Guinea Highland Stock. Vol.1*. 65 – 87. Seattle: University of Washington Press. (ISBN-10 029595132X)
- Lu and Wen 2002: Lu, Bingfu and Zhisheng (Edward) Wen. 2022. Short-term and working memory capacity and the language device: Chunking and parsing complexity. In J. Schwieter and Z. Wen, eds., *The Cambridge Handbook of Working Memory and Language*. 393 – 18. Cambridge: Cambridge University Press. (ISBN 9781108955638)
- Lynch 1998: Lynch, John. 1998. *Pacific Languages. An Introduction*. Honolulu: University of Hawai'i Press. (ISBN 9780824842581)
- Makhashen, Shuib and Lah 2008: Makhashen, Kaled Awad, Munir Shuib and Salasiah Che Lah. 2008. The noun phrase structure in Soqotri. Paper presented at the International Conference on Moribund Languages and Cultures, 14 – 16, October 2008. Kuala Lumpur, University of Malaysia. Ms., University Saine Malaysia.
- Martin 1969: Martin, James E. 1969. Semantic Determinants of Preferred Adjective Order. *Journal of Verbal Learning and Verbal Behavior* 8: 697-704 (ISSN: 0022-5371)
- McBride and McBride 1973: McBride, Sam and Nancy McBride. 1973. *Notes on the Essential Features of the Gimi Language, Eastern Highlands*. Ukarumpa: Summer Institute of Linguistics. <https://www.sil.org/resources/archives/58630>

- McElhanon 1972: McElhanon, Kenneth A. 1972. *Selepet grammar, part I: From root to phrase*. Canberra: Pacific Linguistics. <https://www.sil.org/resources/archives/22828> (ISBN 0 85883 086 8)
- Morey 2005: Morey, Stephen D. 2005. *The Tai Languages of Assam - a grammar and text*. Canberra: Pacific Linguistics (Australian National University). (ISBN-10 0858835495)
- Payne 2020: Payne, Doris L. 2020. Robust argument phrases (DPs) but unruly NPs in Maa. In T. Ono and S. A. Thompson, eds., *The 'Noun Phrase' across Languages. An emergent unit in interaction*. 329-362. Amsterdam: Benjamins. (ISBN 9789027261519)
- Pearson 2000: Pearson, Matthew. 2000. Two types of VO languages. In *The Derivation of VO and OV*, ed. P. Svenonius. 327 – 363. Amsterdam: Benjamins. (ISBN 9789027227522)
- Plank 2003: Plank, Frans. 2003. Noun Phrase Structure: An und für sich, in time and space. In F. Plank, ed., *Noun Phrase Structure in the Languages of Europe*. 3–33. Berlin: Mouton de Gruyter. (ISBN 9783110157482)
- Rackowski and Travis 2000: Rackowski, Andrea and Lisa Travis. 2000. V-initial Languages: X or XP Movement and Adverbial Placement. In *The Syntax of Verb Initial Languages*, eds. A. Carnie and E. Guilfoyle. 117 – 141. New York: Oxford University Press. (ISBN 9780195132229)
- Riemsdijk 1998: Riemsdijk, Henk van. 1998. Head Movement and Adjacency. *Natural Language and Linguistic Theory* 16: 633 – 678 (ISSN 0167-806X)
- Schweikert 2005: Schweikert, Walter. 2005. *The Order of Prepositional Phrases in the Structure of the Clause*. Amsterdam: Benjamins. (ISBN 9789027228079)
- Scontras, Degen and Goodman 2017: Scontras, Gregory, Judith Degen, and Noah D. Goodman. 2017. Subjectivity predicts adjective ordering preferences. *Open mind: discoveries in cognitive science*. 1: 53 – 66 https://alpslab.stanford.edu/papers/2017_ScontrasDegenGoodman.pdf (ISSN 2470-2986)
- Scott 2002: Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In G. Cinque, ed., *Functional structure in DP and IP*. 91 – 120. Oxford: Oxford University Press. (ISBN 9780195148800)
- Sheehan, Biberauer, Roberts, Holmberg 2017: Sheehan, Michelle, Theresa Biberauer, Ian Roberts, Anders Holmberg. 2017. *The Final-Over-Final Condition: A Syntactic Universal*. Cambridge, MA: MIT Press. (ISBN 9780262534161)
- So-Hartmann 2009: So-Hartmann, Helga. 2009. *A Descriptive Grammar of Daai Chin*. Berkeley: University of California (STEDT Monograph 7). https://stedt.berkeley.edu/pubs_and_prods/STEDT_Monograph7_Daai-Chin.pdf (ISBN 0-944613-46-2)
- Stapert 2013: Stapert, Eugénie. 2013. *Contact-Induced Change in Dolgan: An Investigation into the Role of Linguistic Data for the Reconstruction of a People's (Pre)History*. Ph.D. Dissertation, Rijkuniversiteit te Leiden. https://www.lotpublications.nl/Documents/336_fulltext.pdf

- Syed 2015: Syed, Saurov. 2015. Focus-movement within the DP: Bangla as a novel case. In U.Steindl, et al., eds. *Proceedings of the 32nd West Coast Conference on Formal Linguistics*. 332 – 341. Somerville, MA: Cascadilla. (ISBN 978-1-57473-466-9)
- Van Engelenhoven and Williams-van Klinken 2005: Van Engelenhoven, Aone, and Catharina Williams-van Klinken. 2005. Tetun and Leti. In S. Adelaar and N. Himmelmann, eds., *The Austronesian languages of Asia and Madagascar*. 735 – 768. London: Routledge. (ISBN 0-7007-1286-0)
- Yarapea 2006: Yarapea, Apoi Mason. 2006. *Morphosyntax of Kewapi*. Ph.D. Dissertation, The Australian National University. <https://digitalcollections.anu.edu.au/handle/1885/13226>
- Zhang 2015: Zhang, Niina Ning. 2015. Nominal-internal phrasal movement in Mandarin Chinese. *The Linguistic Review* 32 (2): 375 – 425 (ISSN 1613-3676)

Prof. Guglielmo Cinque

Ca' Foscari University

Venice, Italy

E-mail: cinque@unive.it

Ваня ЗИДАРОВА

(Пловдивски университет „Паусий Хилендарски“)

КЛАСИФИКАЦИЯ И ВАРИАНТНОСТ НА УСТОЙЧИВИТЕ СЛОВОСЪЧЕТАНИЯ В БЪЛГАРСКИЯ ЕЗИК

Резюме. Устойчивите словосъчетания са голяма, специфична и разнообразна група езикови единици. Това ги прави сериозно теоретично и лексикографско предизвикателство. В статията се прави преглед на досегашните класификации и се предлага нов модел, при който се разграничават ядро и периферия. Разглеждат се и проявите на структурна вариантност в две насоки: в езика и при речева реализация. Наблюденията се базират на корпуса в един от фразеологичните речници.

Ключови думи: устойчиво словосъчетание, фразеологизъм, класификация, вариантност

Vanya ZIDAROVA

(Paisii Hilendarski University of Plovdiv)

CLASSIFICATION AND VARIATION OF FIXED EXPRESSIONS IN BULGARIAN

Abstract. Fixed expressions comprise a large, specific and diverse group of linguistic units which makes them a serious theoretical and lexicographical challenge. This article offers a review of the existing classifications. There is also a new model proposed – one which distinguishes between a core and a periphery. Manifestations of structural variability are also considered in two directions: in language and in speech. The observations are based on the corpus in one of the Bulgarian phraseological dictionaries.

Keywords: fixed expressions, phraseme, classification, variation

Устойчивите словосъчетания представляват голяма и доста разнообразна по структура група, която е важна част от разговорната реч

на първо място. Въпреки своята популярност и сравнително ясната концепция за природата на тези словосъчетания те все още не са ясно очертани като обем и разновидности.

В българската лингвистика вече съществува значително количество изследвания в областта на фразеологията, но все още има редица спорни възгледи и територии, отнасящи се преди всичко до семантичните особености на устойчивите словосъчетания, до тяхната класификация и оттам – до обема на самото понятие „устойчиво словосъчетание“. Класификацията на устойчивите словосъчетания до голяма степен зависи от възгледа на автора по отношение на обема от словосъчетания, приемани за устойчиви, и от разбирането за същността на фразеологичната единица. Съществуват гранични или спорни по същността си устойчиви словосъчетания, които е трудно да бъдат локализиращи и/или приобщени в състава на българския фразеологичен фонд. Единство съществува само що се отнася до схващането, че устойчивите словосъчетания са езикови единици (подобно на лексемите) и имат номинативна (макар и специфична по отношение на номинативната функция при лексемите) функция. Това са социално утвърдени и приложими единици, които не се създават по индивидуален път, а се възпроизвеждат в готов вид, тъй като са част от колективния езиков запас. Както показва самото наименование, това са: 1) словосъчетания, т.е. имат разделна лексикална структура и включват повече от една лексема; 2) връзката между лексикалните единици е устойчива и възпроизводима. Семантиката на тези словосъчетания обаче е доста разнообразна по своята същност, което изисква при евентуална класификация да се вземе предвид и семантичният критерий – (не)задължителна образност и съответно конотация във фразеологичното значение. Ако приемем за определящ единствено критерия „устойчивост и възпроизводимост на словосъчетанието“, в обема на устойчивите словосъчетания ще се окажат различни видове словосъчетания, като разликата между някои от тях е твърде съществена. При едно широко разбиране на понятието „устойчиво словосъчетание“ в обема му ще влязат всички изрази с разделна, но възпроизводима структура. При по-тясно разбиране на понятието, когато се вземе предвид и наличието/отсъствието на образно преосмисляне, обемът очаквано ще се стесни. Ст. Калдиева-Захаријева разглежда признака „устойчивост“ като комплексен феномен, който се изразява чрез три важни стойности: цялостна обвързаност (независимо от степента на споеност), употреба в готов вид като единна цялост, исторически утвърдено присъствие в езика (Калдиева-

Захариева / Kaldieva-Zaharieva 2013: 31). Съгласни сме, че единството на тези три характеристики обединява различните типове устойчиви словосъчетания, чиято класификация и интерпретация обаче варира при различните изследователи. В българската фразеология преобладава тясното схващане за същността на устойчивите словосъчетания, при което значителни групи устойчиви словосъчетания не се включват в понятието „фразеологизъм“ и съответно не се смятат за предмет на фразеологията. За разграничител между същинските фразеологизми и спорните откъм принадлежност други устойчиви словосъчетания Ст. Калдиева-Захариева приема следните признаци, присъщи единствено на фразеологизмите: фразеологично значение (имплициращо оценъчност, експресия, познание), образност и номинативна функция от втори ранг, обединяваща номинация и конотация (Калдиева-Захариева/ Kaldieva-Zaharieva 2013: 32). Имайки предвид тези признаци, авторката основателно изключва от обема на фразеологията сложни съюзи, предлози и частици, чиято устойчивост не е съпроводена от образност.

В първото монографично изследване на устойчивите словосъчетания К. Ничева определено предпочита широкото схващане на понятието „устойчиво словосъчетание“, като вмести в обема му изключително разнообразни по природа изрази (Ничева / Nicheva 1987: 20 – 36). Класифицирайки видовете устойчиви словосъчетания с оглед на това дали те могат да бъдат определени като фразеологизми, или не, авторката обособява три основни групи изрази, излизащи от обема на понятието „фразеологизъм“:

1) устойчиви словесни комплекси, към които се отнасят описателни словосъчетания (перифрази) без семантична трансформация (*превъртам ключа = заключвам, отивам на дъното = потъвам, давам съвет = съветвам, вземам решение = решавам*), някои устойчиви сравнения¹, предложно-именни съчетания (*на вятъра, до шия*),

¹ Възгледите на К. Ничева са доста колебливи по отношение на устойчивите сравнения. Тя не ги приема като определен тип устойчиви словосъчетания, а внася диференциация сред тях, която според нас не е необходима. Например не приема за фразеологизирани сравненията, при които признакът е органично, обективно присъщ на обекта и съответно не е налице семантична образност (*бял като сняг, черен като въглен*). Ако има семантична трансформация, субективно приписване на признак, степента на фразеологизираност може да бъде по-ниска (*лъжа като циганин, ям като ламя*) или по-висока (*бял като на тенджерата дъното, сит като просяк*). При ироничните сравнения експресивността произтича от антонимното отношение *бял – черен, сит – гладен*.

сложни предлози и съюзи, съставни названия и термини (*майски бръмбар, адамова ябълка, момина сълза, плачеца върба*)², фразеосхеми (*град като град, от хубав по-хубав, срам не срам*);

2) устойчиви фрази със затворена структура (*вятър и мъгла, загубена работа, горе главата*);

3) фразеологизми от преходен тип, които заемат междинно положение между поговорките (въпреки че често имат структура на изречение) и фразеологизмите. Например: *плаши Пейо кокошките; ела, Вълчо, изяж ме; облякъл се Илия – пак в тия*. Преходен характер имат и същинските пословици и поговорки³ (*два остри камъка брашно не мелят, гладна мечка хоро не играе*), крилатите фрази (*буря в чаша вода, целта оправдава средствата, жребият е хвърлен*), цитатите и заглавията, добили общоизвестност (*мамино детенце, тежка е царската корона, мила родна картинка, изгубена Станка*). Обособяването на този „преходен тип“ е дело на К. Ничева и към тях тя отнася изрази, които заемат междинно положение и клонят към фразеологизация. Характерна особеност на този преходен тип според К. Ничева е семантизирането им с помощта на изрази от типа „употребява се, казва се когато, като...“.

Посочените дотук три основни групи устойчиви словосъчетания К. Ничева изключва от състава на същинските фразеологизми. Въпреки своята устойчивост и възпроизводимост тези изрази не се приемат за фразеологизми – мнение, което и ние споделяме, въпреки, че имаме някои съмнения относно природата на устойчивите сравнения и поговорките. За нас устойчивите сравнения и поговорките стоят много близо до фразеологизмите, тъй като споделят най-съществения белег на фразеологизмите – образна експресивност. Смятаме, че отношението „устойчиво словосъчетание“ – „фразеологизъм“ е йерархично от родово-видов порядък. Всички фразеологизми са устойчиви словосъчетания, но не всички устойчиви словосъче-

² Образността при тях е почти напълно избледняла и при тях доминира неутралната номинативна функция.

³ Отношението на пословиците и поговорките към фразеологичния фонд е дискусивно до днес. Те биват включвани или изключвани спрямо групата на фразеологизмите в зависимост от становището на конкретния автор. Най-често се изтъква, че тези изрази, принадлежащи по принцип на фолклора, имат структура на завършено изречение и съдържат някаква поука, което ги отдалечава от същинските фразеологизми. Това, което ги сближава, е наличието на образност.

тания са фразеологизми. С други думи – фразеологизмите са вид устойчиво словосъчетание. В зависимост от по-тясното или по-широкото разбиране за науката фразеология тя може да се определи като наука за устойчивите словосъчетания или като наука за фразеологизмите. Българската фразеология прилага по-скоро тясното схващане за предмета на фразеологията, оставяйки извън обема на науката разнообразни по тип устойчиви словосъчетания. Този подход е приемлив, защото позволява да се концентрира вниманието върху сравнително обозрим кръг езикови единици, обединени от общи свойства.

В най-новото изследване върху българската фразеология Ст. Калдиева-Захариева също класифицира устойчивите словосъчетания с оглед на принадлежността им към фразеологията (Калдиева-Захариева / Kaldieva-Zaharieva 2013: 33 – 52). Извън рамките на фразеологизмите авторката поставя следните устойчиви словосъчетания:

1) сложни названия и термини, при които липсва образност или е напълно заличена (*морска сол, лунно затъмнение, научен съвет, попова лъжичка, детска градина*)⁴;

2) пословици и поговорки (паремии). Те са фолклорен жанр и съответно обособена категория изрази със специфични характеристики, невинаги съвпадащи с тези на фразеологизмите (изреченска структура, логическо съждение, наличие на поука или съвет в семантиката). Например: *който не работи, не трябва да яде, бий желязото, докато е горещо, два остри камъка брашно не мелят*. Поради ясно изразената им експресивност обаче те могат да заемат преходно положение и да преминават понякога в групата на фразеологизмите;

3) крилати фрази (*жребият е хвърлен, след нас и потоп*), които авторката поставя на границата между паремии и колокации⁵. И при тях, както при паремии, е възможна фразеологизация в разволя им (*минавам Рубикон*);

4) благословии, пожелания, клетви (*златни ти уста, Бог да те поживи, да спи зло под камък, босилек да ти занесат*), които имат

⁴ Примерите са на авторката във всички разглеждани групи устойчиви словосъчетания.

⁵ За термина „колокация“ виж по-нататък.

старинен фолклорен характер⁶. Повече или по-малко, те съдържат метафоричност в семантиката си и както и при други групи устойчиви словосъчетания е възможно да настъпи фразеологизация. Според наблюденията на Л. Крумова-Цветкова метафоричността е по-широко застъпена при клетвите, при които отрицателната оценъчност намира израз в ярки образи (Крумова-Цветкова / Krumova-Tsvetkova 2010: 129);

5) устойчиви сравнения (*твърд като камък, ясно като бял ден, кротък като агне, мъж като скала*), които са изключително голяма и разнообразна група устойчиви словосъчетания. Те са специфични по структура устойчиви словосъчетания, които могат да се разглеждат като преходна група или като самостоятелен обект на специфични изследвания (Кювлиева-Мишайкова / Kyuvlieva-Mishaykova 1986a). Безспорно устойчивите сравнения притежават всички черти, присъщи на фразеологизмите – разделна оформеност, образно-експресивен характер, устойчив лексикален състав, и при едно по-широко схващане за предмета на фразеологията биха били включени във фразеологичния фонд. Дори и при едно по-тясно разбиране за фразеологизмите някои устойчиви сравнения показват очевидни признаци на фразеологизация. Според В. Кювлиева-Мишайкова това са тези устойчиви сравнения, при които първоначалната ограничена съчетаемост е преодоляна и семантичното поле е разширено (*стоя, слушам, ходя, оставам* като попарен; *говоря, ям, викам, работя* като разпран);

6) колокации, които са нова група устойчиви словосъчетания, въведена от Ст. Калдиева-Захариева, заети като термин от англоезичната и френската лингвистика. Обяснявайки термина, Ст. Калдиева-Захариева го тълкува по следния начин: „...думи, свързани по „съседство“ на разположение, изразяващи минимален свързан текст, които са получили в резултат на честа употреба трайна връзка помежду си и възпроизводимост в речта“ (Калдиева-Захариева / Kaldieva-Zaharieva 2013: 48). Като колокации се разглеждат устойчиви словосъчетания, които в по-ранни изследвания са били отделни групи. Такива са: устойчиви словосъчетания, при които само единият компонент е подложен на семантична трансформация (*мечешка услуга, висок гост, люта зима, дълбока тишина*), перифрази на глаголи (*вземам решение* – реша-

⁶ Подробен и задълбочен анализ на този жанр е направен в изследването на Л. Крумова-Цветкова „Българските благословии, пожелания и клетви – част от националната самобитност и културното наследство“ (2010).

вам, *давам съвет* – съветвам, *вдигам въстание* – въставам), съкратени сравнения (*очи-череша, планина човек*), речеви клишета (*моля за извинение, моля за разрешение, на добър час, за много години*). Както се вижда, в групата на колокациите се поместват доста разнородни по характер словосъчетания. Според нас обединяването им в обща група не е напълно основателно, тъй като те са видимо твърде различни по характер. Например при някои има частична метафоризация, при други има лексикален еквивалент, при трети – исторически установен модел.

Сравнението на двете класификации – на К. Ничева и на Ст. Калдиева-Захаријева, показва, че между вижданията на двете авторки има както сходни позиции (обособяване на едни и същи групи словосъчетания), така и различия. Най-очевидното различие е в отделянето на група „*преходни словосъчетания*“ (К. Ничева) и „*колокации*“ (Ст. Калдиева-Захаријева). Общото е, че и двете авторки допускат възможност за фразеологизация на словосъчетания, които към момента са извън състава на фразеологизмите.

Споделяйки характерното за българската фразеология тясно схващане за същността на тези езикови единици, ние приемаме, че обект на фразеологията са само фразеологичните единици (фраземи, фразеологизми). Останалите устойчиви словосъчетания имат собствена специфика, не споделят всички признаци, присъщи на фраземите, поради което трябва да бъдат обособени и евентуално да бъдат обект на отделни научни изследвания, както е направено по отношение на групите устойчиви сравнения, поговорки и пословици, пожелания, благословии и клетви. При структурирането на групата „*устойчиви словосъчетания*“ смятаме за подходящо тази група да се представи като поле съответно с ядро и периферия. Приемаме, че ядро на полето са фразеологизмите, които дефинираме по следния начин: фразеологизмът представлява социално осъзната и възпроизводима в речта езикова единица с разделно оформена структура, устойчива синтактична връзка, относително постоянен лексикален състав и цялостна семантика, съчетаваща номинация и задължителна конотация. Например: *от кол и въже, през просото, с две леви ръце, девета дупка на кавала; от дъжд на вятър; крия си главата в пясък; под път и над път; цепен косъма на две; свиря първа цигулка; вкарвам си автогол; на час по лъжичка; клетя си краката*. Експресивно преосмислената семантика приемаме за задължителен и най-съществен признак по отношение на словосъчетания, определяни като фразеологизми. В периферията поставяме онези устойчиви сло-

восъчетания, по отношение на които не са изпълнени всички от посочените в дефиницията езикови характеристики и най-вече семантичната трансформация. Те съответно не следва да се разглеждат като фразеологизми. Периферията се формира от следните видове устойчиви словосъчетания:

- ✓ Описателни названия (съставни думи, съставни термини), които се отличават с липса на образност или с частична образност (*черна дъска, нощна лампа, вземам решение, вземам участие, имам намерение, пускат тока, вдигам масата, вдигам телефона, пускам телевизора, мечешка услуга, миден месец*). В тази група словосъчетания е налице устойчив синтактичен модел, разделна лексикална структура, но в семантично отношение доминира първичната номинативна функция, а образността е силно избледняла. Границата между словосъчетанията, определяни като съставни думи (*нощна лампа, кабелна телевизия, помощен персонал, черна дъска, маршрутно такси, момина сълза, натриев хлорид, захарна болест*) и като описателни названия (*давам си сметка, развалям пари, ям бой, златна среда*), е изключително тънка и е почти невъзможно да се отнесе категорично даден израз към едната от двете групи;
- ✓ Устойчиви сравнения, които стоят най-близо до фразеологичното ядро (*хитър като лисица, седя като на тръни, повтаря като папагал, катери се като коза, студен като лед, спя като заклан*). Въпреки че в тях една от думите не е преосмислена (*хитър, седя, повтарям, студен, спя, катеря се*), като цяло устойчивите сравнения се отличават с ярко изразена експресивност, дължаща се на образа, заложен при тяхната номинация. Образността на устойчивите сравнения произтича от прехвърлянето на даден признак от един денотат върху друг. В някои случаи образността е доста очевидно мотивирана (*бял като сняг, черен като въглен, червен като домати*), но има и такива сравнения, при които образността се отличава с подчертана експресивност и оригиналност (*точат се като народна песен, обичат се като куче и котка, прилича му като на свинче звънче, навързани като свински черва*);
- ✓ Фразеосхеми, които са особено специфичен вид устойчиви словосъчетания, тъй като при тях устойчивост се наблюдава само по отношение на синтактичната конструкция, но не и

по отношение на лексикалния състав (*от хубав по-хубав, от нов по-нов; ученото си е учено, новото си е ново; къщата му къща, работата му работа; хубав не, ами хубав; кино, та кино, море, та море*);

- ✓ Фрази от речевия етикет, при които всъщност липсва образност, но те са се наложили като устойчиви конструкции в исторически и културен план (*добър вечер, лека нощ, на добър час, за много години, честит празник, живот и здраве*);
- ✓ Пожелания, благословии и клетви (*бял ден да не види, да пукне дано, голям да порасне, да е жив и здрав*). Това са исторически утвърдени словосъчетания, чиято ономаσιологична същност е доста прозрачна, но въпреки това те се употребяват като компактни единици;
- ✓ Поговорки и пословици, които (наред с устойчивите сравнения) формират групата на може би най-спорните устойчиви словосъчетания, тъй като, от една страна, представляват специфична в структурно и семантично отношение група, но от друга страна, притежават присъщата на фразеологизмите образност и експресивност. Конкретно поговорките и пословиците например имат структура на изречение и семантика от типа на съждение, която е свързана с определен тип ситуация и поука от нея (*насила хубост не става; една лястовичка пролет не прави; от всяко дърво свирка не става; кучетата си лаят, керванът си върви; всяка жаба да си знае гъола*);
- ✓ Цитати и крилати фрази, които имат относително познат произход, получили са социална известност и биват включвани в контекст и ситуация, различни от оригиналните (*мила родна картинка; има нещо гнило в Дания; орел, рак и цука; целта оправдава средствата, минавам Рубикон, жребият е хвърлен*). Характерно за тази група е, че фразите или имат конкретен автор, или са родени в конкретна историческа ситуация, но впоследствие са се откъснали от първоначалния контекст (фолклорен, литературнохудожествен, исторически), а авторството им (ако е съществувало) е избледняло и днес тези фрази битуват в колективната говорна практика и се възпроизвеждат като част от общия езиков фонд.

Подобен полеви модел използва и В. Кювлиева, която разграничава *същински фразеологизми*, при които има цялостно преосмисляне на значението, и *периферни фразеологизми*, при които има само частично преосмисляне на значението (Кювлиева / Kyuvlieva: 1986б).

Вторият проблем, на който ще обърнем внимание, е възможността за вариантност при устойчивите словосъчетания, които по подразбиране на тяхната устойчивост не би трябвало да допускат различни варианти на реализация. Все пак, въпреки че устойчивото словосъчетание представлява неразруσιμο единство от форма и значение, а основна негова езикова характеристика е именно установеността, повторемостта на структурния и лексикалния модел, съществува възможност и за вариантност. По тази тема бихме отбелязали студията на Кр. Чолакова „Към въпроса за формата на фразеологичната единица“ (Чолакова/Cholakova 1967) и статията на същата авторка „За някои видоизменения във формалната структура на фразеологичните единици“ (Чолакова/Cholakova 1968). Сериозно внимание на въпроса за вариантността при фразеологизмите отделят и авторките на монографични изследвания К. Ничева (Ничева / Nicheva 1987) и Ст. Калдиева-Захариева (Калдиева-Захариева / Kaldieva-Zaharieva 2013). Разделната структура на устойчивото словосъчетание позволява в структурата му да се вмъкнат или от нея да изпаднат определени лексеми, без това да се отрази на семантиката. Включването в дискурса също води до структурни промени, отнасящи се главно до граматичната адаптация на израза спрямо контекста. Кр. Чолакова и С. Спасова в увода към Фразеологичния речник на българския език говорят за формална вариантност (*като свят – като света*), т.е. замяна на една граматична форма с друга, и за лексикална вариантност, при която бива заменена една лексема с друга, обикновено неин синоним (*седя със скръстени ръце – стоя със скръстени ръце*) (Спасова, Чолакова, Ничева / Spasova, Cholakova / Nicheva 1974 – 1975). В монографията си К. Ничева разграничава „абсолютни варианти“, които са идентични по качество и количество на значенията си, но с разлика в лексикалния състав и словоформите, и „частични варианти“, при които има разлика в броя на значенията и в стилистичната характеристика (Ничева / Nicheva 1987: 143). Наред с това същата авторка различава „формална вариантност“, при която е налице разлика в граматичните форми и която се среща най-често при фразеологизми, включващи глагол, и „лексикална вариантност“, при която се заменя една лексема с неин синоним (Ничева / Nicheva 1987: 145). Ст. Калдиева-Захариева предлага свое виждане за видовете вариантност при фразеологизмите, уместно отбелязвайки, че става дума за своеобразно противоречие: от една страна, фразеологизмът е устойчива, непроменлива единица, а от друга, в плана на изразяване могат да се

осъществяват различни промени. Тези промени са възможни на първо място поради разделната лексикална оформеност на фразеологизма, която позволява замяна на един компонент с друг. Авторката свежда вариантността до три типа: граматична, позиционна и лексикална (Калдиева-Захариева / Kaldieva-Zaharieva 2013: 178 – 180). Граматичната е сравнително ограничена откъм проява, защото Ст. Калдиева-Захариева не приема за такава граматичната адаптация към контекста (например по лице, число, време, вид при глаголите (*хвърлям прах в очите на.../хвърля прах в очите на.../хвърлил прах в очите на...*)), а приема само случаите, в които редуването на форми (предимно именни) е проява на вариантност в езиков план (*оплитам в мрежата си/оплитам в мрежите си*). Според нас всяка вариация във формално отношение (независимо дали се дължи на участие в конкретен контекст, или е езикова вариация) представлява граматична вариантност. Но може да се въведе разграничение по линията речеви варианти (промяна, дължаща се на контекста) и езикови варианти (речникови разновидности на фразеологизма). Лексикалната вариантност у Ст. Калдиева-Захариева се изразява в отнемане (или добавяне) на лексема, като добавеният елемент има факултативен характер (*никнат като гъби <след дъжд>*), или в замяна на една лексема с друга (*падам в ръцете на някого/падам в лапите на някого*). Замяната по принцип включва лексеми, сходни по значение или принадлежащи към едно тематично поле.

Когато разглеждаме формалната страна на фразеологизма, не можем да избегнем т.нар. екзоелементи, които като структурен елемент имат задължителен характер, но като лексикален елемент се реализират под различна форма. Задължително съпровождащи могат да бъдат различни лексеми – съществителни имена, собствени имена, местоимения (*правя нещо до спукване, затягам въжето около шията на някого, правя нещо за четирима, скривам шайбата някому, някой е дялан камък, не цепя басма никому, хвърлям камъни в нечия градина, слагам рога на някого, в нечи уста съм, бия дупката на някого*). Екзоелементи има при фразеологизмите с отворена структура, които се адаптират към условията на речевия акт, за разлика от фразеологизмите със затворена структура, които имат завършен, установен вид и се възпроизвеждат по един и същ начин независимо от контекста (*през просото, от дъжд на вятър, девета дупка на кавала*). Екзоелементите са лексикографско предизвикателство, защото, от една страна, те са продукт на контекстовата употреба, но от друга страна, без тях фразеологизмът няма завърше-

ност. Като екзоелемент се разглежда понякога и *tertium comparationis* (признак на сравняването) в устойчивите сравнения, който е задължително съпровождащ елемент в устойчивото сравнение: *(остър) като четина, (глава) като шиник, (слаб) като щиглец, (вреция) като яре* (Анкова-Ничева / Ankova-Nicheva 1993).

В българската лингвистична литература вариантността се теоретизира като вид структурна трансформация, присъща на фразеологизмите. В широк смисъл тя е разглеждана като всяка промяна във формата на утвърдени фразеологизми – граматична, лексикална, позиционна, количествена, структурна. В по-тесен смисъл се приема само лексикалната вариантност, замяна на един лексикален компонент с друг без промяна в семантиката и образа. Тук ще разгледаме вариантността в структурата (граматична и лексикална) на два типа устойчиви словосъчетания – фразеологизми и устойчиви сравнения. За удобство ще употребяваме родовото (устойчиви словосъчетания) и видовото (фразеологизми) понятие като синоними. От периферията на фразеологичното поле избираме именно устойчивите сравнения, защото те стоят най-близо до фразеологизмите, като споделят основните признаци – образност, устойчив синтактичен строеж, възпроизводимост в речта. Като база за ексцерпция сме използвали „Нов фразеологичен речник на българския език“ (Анкова-Ничева / Ankova-Nicheva 1993). Всъщност този речник към момента не би могъл да се определи като „нов“, защото датира от преди тридесет години. Но по-нов фразеологичен речник не е изготвен, а и по-голямата част от материала е исторически утвърден и в този смисъл е актуален и към днешна дата. К. Анкова-Ничева разграничава и съответно включва в корпуса два вида устойчиви словосъчетания – фразеологизми и т.нар. от авторката „фразеологизми от преходен тип“, които имат структура на изречение и стоят близо до поговорките и пословиците. В речника обаче са включени и поговорки, и устойчиви сравнения, които авторката не отбелязва специално, от което съдим, че тя ги третира като фразеологизми въпреки спорния им характер. Съдържанието на речника включва устойчиви словосъчетания, взети от тълковни речници, двутомния фразеологичен речник, лични архиви, диалектни ресурси, монографични изследвания, фолклорни сборници. В увода към речника К. Анкова-Ничева представя начина, по който са обработени вариантите на устойчивите словосъчетания, като различава съответно следните видове варианти: структурни (*вари го, печи го/вари го, печи го, с ориз тъпчи го, пак е рак*), позиционни (*на брашното етвин, на триците скъп/на*

триците скъп, на брашното евтин), лексикални (*показвам/разкривам истинското си лице*).

Ще продължим с наблюдения над вариантите при фразеологизмите и устойчивите сравнения, представени в речника, като ще обособим два основни вида вариантност – речева и езикова.

В речника поради неговия системно езиков характер не се отбелязва граматичната вариантност, която се проявява при включване на фразеологизма в определен дискурс, изискващ граматическа съгласуваност на глаголните и местоименните лексеми. В речев план устойчивите словосъчетания се променят по граматичните категории род, лице, число (местоименията), време, лице, число (глаголите). Тази граматична адаптация не е същинска вариантност, тъй като не засяга на практика лексикалния състав на ФЕ и фразеологичното значение – *пепел (му, ти, ѝ) на езика, бия (му, ѝ) дузпата, падам (паднах) в очите на някого, ще пия (ще пием, тих, тихме) една студена вода, лапам (лапа, лапат) мухи, надувам (надух, наду, надул) гайдата, свиря (свири) втора цигулка*. В езиков и лексикографски план тази вариантност не се отбелязва, поради което, както посочихме по-горе, някои изследователи не я приемат за същинска. Но с оглед на функционирането на фразеологизма в речта граматичните промени не може да не бъдат отбелязани. Специфично е положението по отношение на категорията вид на глагола, която има лексикално-граматичен характер. Видовото редуване е включено в речника като вариант, което подсказва, че аспектиалните форми се приемат за различни езикови единици (*вдигам/вдигна мерника на някого, падам/падна в очите на някого, виждам/видя бели вълци по пладне, пращам/пратя за зелен хайвер, предавам/предам фронта*), *влизам/вляза под кожата на някого*.

Елиптичните варианти също имат речев характер и се реализират при включване на фразеологизма в конкретен речев акт. При утвърдени и разпознаваеми фразеологични единици може да настъпи елипса на част от конструкцията. Това се дължи на стремеж към икономия и когато изразът е добре познат, част от него се изпуска, като се подразбира. Особено характерна е тази елипса при поговорките, които са общоизвестни изречения. *Всичко Мара втасала (само се не сресала), Прав ти път (крива ти пътека), Казали на лисицата (тя на опашката си), Думам ти, дъще (сещай се, снахо), Роди ме, мамо, с късмет (па ме хвърли на смет), Вода газии (жаден ходи)*. Елипса се среща и при някои устойчиви сравнения: *като сардели, като на шегга, като мухи на мед, като изваден от възглавница* (Калдиева-

Захариева / Kaldieva-Zaharieva 2013: 44). Като проява на елипса могат да се разглеждат и примерите, определяни като структурна (количествена) вариантност, от типа *меля като (празна) воденица, по никое (си) време, пада като (зряла) круша*.

В системата на езика фразеологичната вариантност се изразява преди всичко в редуването на лексеми и бива отбелязвана в почти всички дефиниции и изследвания на устойчивите словосъчетания. Редуващите се лексеми най-често са синоними, но са възможни и други редувания.

- ✓ Редуване на книжовен и диалектен вариант (или два книжовни/диалектни варианта) – *армасвам за църноземско юначе/годя се за черноземско юначе; бел като арапин/бел като егюптин; дрънкам като празна воденица (жерка, мелница); захванал от аз, че не изкарал до буки/забрътвил от аз, че не изкарал до буки; захлупвам под връшник (връшина); жълт като каун (типон); бия като котка (мачка) о дирек; мъж искам, днес го искам/мъж сакам, сега го сакам; обичаме се като куче и котка/обичаме се като мачка и поганец; спори ми като на костенурка (желка) през угар;*
- ✓ Редуване на лексеми от общо семантично гнездо (синоними) – *умът (акълът) ми е в краката; аман съм казал (викнал/рекъл); белея се (лъция/светя) като тиган на месечина; вземам (тегля) на буксир; грея (светя) като ясно слънце; дяля се като крастава (шугава) овца от стадото; вадя думите с кука (с ченгел) от устата; залепям се като лепка (репей); излезе на ачик (наяве, на бял свят); изскочи ми (изхвъркна ми, изпари ми се, изфиряса ми) от главата; бия тъпана (барабана); вървя (ходя) като през турски гробища; къща като палат (дворец, сарай); меря (тегля) на своя кантар; наредих се като куче в кола (каруца); шуква ми (изпарява ми се, изскача ми) от главата;*
- ✓ Редуване на лексеми от общо понятийно гнездо – *виж му акъла, па му крой шапка (калпак, капа, гугла); до стотинка (до петак); като патка (пате, гъска) в мъгла; черен гологан (грош) не се губи; кротък като муха (божа кравичка); викам като говедар (овчар, поляк, пъдар); кой те пита бръсне ли се кадията (попът, владиката); не се знае от кой храст (от коя трънка) ще изскочи заек;*

- ✓ Редуване на лексеми от общо словообразователно гнездо – *вдигам (повдигам, свивам) рамене; държа яката (яко); като ударен от гръм (като гръмнат); захвърлям (изхвърлям, хвърлям) на боклука; златни ти ръце (златна ти ръчица); и аз съм от майка роден (и мен ме е майка раждала); лепвам се (залепвам се) като гербова марка; проливам (лея) сълзи; и кучетата (кучките) не ме лаят; като змии в змиярник (змиярница);*
- ✓ Редуване на напълно различни лексеми – *бия вода да пусне мляко/бия вода да пусне масло; болен Божко за кокошка/ болен Тошко за кокошка; виждам бял вълк по пладне/виждам бял бивол по пладне; вързан в устата/вързан в езика; гладя с паче перо/гладя с пауново перо; глупости на търкалета/красавици на търкалета; гракнали на него като на бяла врана/гракнали връз него като връз бял гарван; (дъжд вали, слънце грее) мечка (лисицата, заекът, таралежът) се жени; оплачи се на владиката (на арменския поп); женски Пейо (Дойно, Райно); за една бълха цяла черга (юрган) изгарям; връзвам парата в три възела/една пара в девет възела връзвам; какъвто гостът, такава и чорбата (млинът); калесал съм с червена (жълта) бъклица; псува като хамалин (каруцар); кълна като мащеха (свекърва); кметът (царят) дава, пъдарят не дава; кога върба роди круши (грозде, дренки); лежа като паша (селски чорбаджия); лъв с миша (овча) опашка; гладък (мазен) като таралеж; работата не е заек (сърна) да избяга; само на едно куче не викат Шаро (Караман); търся под дърво и камък/търся под трън и габер.*

Според нас различните лексикални варианти са възникнали или по различно време, или на различно място. Те съществуват паралелно, а не са производни един от друг. Особено често се проявява ролята на диалекта – в различни диалектни ареали се употребяват различни варианти. При устното предаване на устойчиви словосъчетания от фолклорен тип е възможна различна интерпретация при възпроизвеждането. Въпреки че вариантността, макар и посочвана, никога не е изследвана системно, наблюденията ни над материала в „Нов фразеологичен речник“ сочат, че вариантността е доста широко застъпена във фразеологичния фонд. От съдържащите се в лексикографския корпус над 9000 устойчиви словосъчетания почти при половината се

наблюдава някакъв вид вариантност. Най-често промяната засяга глаголи, съществителни имена и прилагателни имена. Въпреки че при различните варианти метафоричният образ (съответно и фразеологичното значение) се запазва един и същ, остава открит въпросът дали вариантите трябва да се разглеждат като синоними, или като фразеологични дублети. А конкуренцията между вариантите следва да бъде обект на специално изследване.

ЦИТИРАНА ЛИТЕРАТУРА

- Анкова-Ничева 1993: Анкова-Ничева, К. *Нов фразеологичен речник на българския език*. София: УИ „Св. Климент Охридски“, 1993. [Ankova-Nicheva 1993: Ankova-Nicheva, K. *Nov frazeologichen rechnik na balgarskiya ezik*. Sofia: UI Sv. Kliment Ohridski.]
- Калдиева-Захаријева 2013: Калдиева-Захаријева, Ст. *Българска фразеология*. София: АИ „Марин Дринов“, 2013. [Kaldieva-Zaharieva 2013: Kaldieva-Zaharieva, St. *Balgarska frazeologiya*. Sofia: Marin Drinov.]. ISBN 978-954-322-699-3.
- Кримова-Цветкова 2010: Кримова-Цветкова, Л. *Българските благословии, пожелания и клетви – част от националната самобитност и културното наследство*. София: Емас, 2010. [Krumova-Tsvetkova 2010: Krumova-Tsvetkova, L. *Balgarskite blagoslovii, pozhelaniya i kletvi – chast ot natsionalnata samobitnost i kulturnoto nasledstvo*. Sofia: Emas.]. ISBN: 978-954-357-199-4.
- Кювлиева-Мишайкова 1986а: Кювлиева-Мишайкова, В. *Устойчивите сравнения в българския език*. София: БАН, 1986. [Kyuvlieva-Mishaykova 1986a: Kyuvlieva-Mishaykova, V. *Ustoychivite sravneniya v balgarskiya ezik*. Sofia: BAN.].
- Кювлиева-Мишайкова 1986б: Кювлиева-Мишайкова, В. *Фразеологизмите в българския език*. София: Народна просвета, 1986. [Kyuvlieva-Mishaykova 1986b: Kyuvlieva-Mishaykova, V. *Frazeologizmite v balgarskiya ezik*. Sofia: Narodna prosveta.].
- Ничева 1987: Ничева, К. *Българска фразеология*. София: Наука и изкуство, 1987. [Nicheva 1987: Nicheva, K. *Balgarska frazeologiya*. Sofia: Nauka i izkustvo.].
- Спасова, Чолакова, Ничева 1974 – 1975: Спасова, С., Чолакова, Кр., Ничева, К. *Фразеологичен речник на българския език*. София: БАН. [Spasova, S., Cholakova, Kr., Nicheva, K. *Frazeologichen rechnik na balgarskiya ezik*. Sofia: BAN.].
- Чолакова 1967: Чолакова, Кр. Към въпроса за формата на фразеологичните единици. – *Известия на Института за български език*, 15 (1967), 139 –

171. [Cholakova 1967: Cholakova, Kr. Kam vaprosa za formatata na frazeologichnite edinitsi. – *Izvestiya na Institutata za balgarski ezik*, 15 (1967), 139 – 171.]. ISSN 0323-9934 (print).

Чолакова 1968: Чолакова, Кр. За някои видоизменения във формалната структура на фразеологичните единици. – *Известия на Института за български език*, 16 (1968), 703 – 711. [Cholakova 1968: Cholakova, Kr. Za nyakoi vidoizmeneniya vav formalnata struktura na frazeologichnite edinitsi. – *Izvestiya na Institutata za balgarski ezik*, 16 (1968), 703 – 711.]. ISSN 0323-9934 (print).

Prof. Vanya Zidarova DSc

Paisii Hilendarski University of Plovdiv

Plovdiv, Bulgaria

e-mail: zidarova_v@uni-plovdiv.bg

Христо САЛДЖИЕВ

(Тракийски университет – Стара Загора)

ЗА ЕТИМОЛОГИЯТА НА СТАРОБЪЛГАРСКАТА ТИТЛА ЦЪСАРЬ/ЦЪСАРЬ

Резюме. Статията разглежда етимологията на старобългарската титла *цъсарь/цъсарь*, която до този момент е била обект на различни хипотези и коментари, но през втората половина на ХХ век преобладава мнението за нейния готски произход. На първо място са представени критичните аргументи в съвременните езиковедски изследвания, които оспорват тази хипотеза. Представена е нова хипотеза, свързваща произхода на титлата в двата ѝ фонетични варианта, познати в старобългарските паметници, с късноантичното латиноезично население от Мизия, съвременна Западна България и Източна Сърбия. Неговият език е относително добре познат и проучен благодарение на големия брой късноантични надписи от тези райони, датирани към V – VI в. Разгледани са голям брой антропоними и топоними, както и някои романски заемки в старобългарския език, чиято фонетична структура издава съществуването на езикови контакти между югоизточните славяни и късноантичното латиноезично население на Балканите. Двата фонетични варианта на титлата са обяснени в контекста на фонетичните промени в късните латински говори.

Ключови думи: *цъсарь/цъсарь*, старобългарски, късни латински балкански говори

Hristo SALDZHIEV

(Trakia University – Stara Zagora)

THE ETYMOLOGY OF THE OLD BULGARIAN TITLE ЦЪСАРЬ/ЦЪСАРЬ

Abstract. This article deals with the etymology of the Old Bulgarian (Slavonic) title *цъсарь/цъсарь* which has been a matter of discussion and different hypotheses for a long time. In the second half of the 20th century most authors accepted the Gothic hypothesis according to which the title appears to be a Gothic loan – from *Káisar*.

It is attested in the Gothic transcription of the New Testament made by the Gothic bishop Ulfilla in the 4th century in Moesia. However, this hypothesis is problematic too – it cannot explain the different phonetic variants of the title in the Old Bulgarian (Slavonic) records and modern Slavic languages as well as the lack of labialization of the final short -a in Káisar. The article introduces a new hypothesis insisting on the connection of the title and its variants with the Late Latin vernaculars spread in Moesia and modern Western Bulgaria and Eastern Serbia and attested in a relatively big number of inscriptions dated back to the 5th and 6th century. A large number of anthroponyms, toponyms and some of the Romance loanwords in the Old Bulgarian language, whose phonetic structure indicates contacts between South Eastern Slavs and the Late Antique Latin speaking population in the Balkans are examined. The Old Bulgarian (Slavonic) variants of the title are explained in the context of the phonetic changes in the Late Latin vernaculars.

Keywords: *цѣсарь/цѣсарь, Old Bulgarian (Slavonic), Late Latin Balkan vernaculars*

I. Постановка на проблема

В своя етимологичен и правописен речник на българския език Стефан Младенов (ЕПРБКЕ / ЕРРВКЕ 1941: 674) дава доста ограничена информация върху етимологията на старобългарската титла цѣсарь. Както и всички останали езиковеди преди и след него, той не поставя под съмнение латинската протоформа Caesar, но оставя без коментар междинните етапи до цѣсарь. Единственото, което отбелязва, е, че титлата се явява в старобългарските паметници във вариантите цѣсарь и цѣсарь и че старобългарската кесарь произхожда от ст. вис. немското keisar и гръцкото kaíсар. Последното само по себе си не дава отговор за етимологията на цѣсарь.

Доста по-обстоятелствен в това отношение е Макс Фасмер в своя етимологичен речник на руския език. Той последователно отхвърля опитите за етимологизация на гръцка и народнолатинска основа, както и извеждането на титлата от хипотетичната готска форма *Kaisāreis. Фасмер се спира на засвидетелствания в готските евангелски преводи на епископ Улфила вариант Káisar като евентуалния източник на цѣсарь. Същевременно трябва да се отбележи, че единствената причина, заради която Фасмер отхвърля директен произход от латинското Caesar, е отсъствието на такава дума в румънски и албански (ЭСРЯ / ESRY 1987: 290 – 291).

През втората половина на XX век мнението на езиковедите по този въпрос се консолидира именно около засвидетелстваната в превода на Улфила мизийско-готска форма Kaisar. В това отношение

показателна е „Грамматика на старобългарския език“, в която усвояването на мизийско-готското *Kaisar* в епохата на късния праславянски не се поставя под съмнение и дори се дава като пример за втори тип палатализация **k > ts* след монофтонгизирането на праславянския дифтонг **ai* към *ě* (Илчев / Пчев 1991: 116). Едно излязло през 2013 г. изследване върху германските заемки в славянските езици представя доста подробно проблемите пред тази етимологизация, макар и авторката (С. Пронк-Тиетхоф) да не предлага алтернатива на готската и най-общо на германската хипотеза. Тя намира за проблематична връзката с *Kaisar* най-вече с оглед на вокализма на началната сричка: *цѣсарь* не е единственият фонетичен славянски вариант на титлата. Според нея в „староцърковнославянските“ източници тя се появява в два варианта: с *-ě* и *-e*, като от първия произлизат съвременните чешки и украински *-i* форми, а сръбските, хърватските и словенските *-e* форми произлизат от начален *-e* вариант, тъй като *-ě* дава в сръбски и хърватски *-je*. Отделен проблем представлява адаптацията на последната гласна *-a* от готската *Kaisar*, доколкото в протославянски адаптацията на краткото *-a* става посредством *-o*. Въз основа на това противоречие С. Пронк-Тиетхоф заключава, че е възможно титлата да е усвоена след VII век, при това в три различни варианта от отделните славянски племена: **cěsarjъ*, **cesarjъ* и **cъsarjъ*. За последното тя намира основание не само в срещаните в църковнославянския записи *цѣсарь*, но и съкратената форма *царь* в български и руски, която тя очевидно приема като резултат от отпадането на еровите гласни в неударена позиция. Отново според С. Пронк-Тиетхоф северозападните славяни са възприели варианта с *-ě*, предшествениците на днешните сърби, хървати и словенци – този с *-e*, а тези на българи, сърби¹ и руси – с *-ъ*. Тя също така заключава, че готската основа на славянската титла е съмнителна, но и че останалите познати германски форми също не могат да се считат за нейни донори и че произходът на титлата е неизвестен (Pronk-Tiethoff 2013: 100 – 102).

Без съмнение възраженията на С. Пронк-Тиетхоф имат основания, доколкото тезата за първоначална адаптация на *Kaisar* посредством протославянския дифтонг **ai* впоследствие монофтонгизиран до

¹ С. Пронк-Тиетхоф на два пъти споменава сърбите – веднъж като приели варианта с „*e*“ и втори път – варианта с „*ъ*“. Съдейки от логиката на нейните разсъждения, споменаването на сърбите при варианта с „*ъ*“ или е грешка, или е резултат от това, че и в сръбския съществува думата „*цар*“ – т.е. в езика на средновековните сърби трябва да са съществували две дублетни форми.

-ѐ, не е в състояние да обясни вариантите с -е и с -ъ. Тук обаче трябва да се отбележат и някои несъответствия между сведенията на С. Пронк-Тиетхоф и другите източници: т. напр. речникът на старобългарския език дава данни основно за вариант с -ѐ (СБ / SB 2009: 1179 – 1184) и един пример от Супрасълския сборник с -ъ: цъсароу (СБ / SB 2009: 1179 – 1184), който трябва да е развитие от форма с -ъ вследствие на регресивна асимилация $\text{ъ} > \text{ѐ}$. Последното се потвърждава и от съобщението на Стефан Младенов за втора форма в старобългарски с -ъ, а не с -е (ЕПРБКЕ / EPRBKE 1941: 674). Очевидно -е вариантът е изведен по-скоро по дедуктивен път от сръбско-хърватско-словенските форми.

Допълнително объркване предизвиква и тезата на С. Пронк-Тиетхоф за адаптирането на думата от различни славянски общности. Така например според нея славянските предци на руснаците и българите са адаптирали титлата с -ъ, на украинците и чехите – с -ѐ, а на сърбите, хърватите и словенците – с -е. Подобно разпределение едва ли е възможно с оглед на добре известните разделения в славянската езикова общност, които не предполагат подобно „прескачащо“ заимстване – т. напр. трудно може да се обясни украинският „клин“ между руснаци и българи, особено като се има предвид, че предците и на руснаците, и на украинците през Ранното средновековие са едни и същи източнославянски племена.

В настоящата статия ще се опитам да представя една различна хипотеза относно езика донор на цъсарь / цъсарь и географския регион на нейното заимстване. За такъв език донор приемам народния латински и неговите диалекти от Мизия и Централните Балкани, отразени в редица антични надписи от IV – VI век. Този език по редица свои белези съществено се е различавал от проторумънския и е взаимодействал с преселилите се на юг от Дунав славянски племена от т. нар. „българска група“.

II. Средновековни български антропоними, топоними и лексеми, издаващи контакти с балкански романоезични групи от непроторумънски произход

Исходният модел, по който ще се опитам да анализирам наличните езикови следи, се базира на фонетичните различия между езика на латински надписи от Мизия и проторумънския съгласно с общоприетите реконструкции на последния.

1. Особенности на проторумънския език

Проторумънският се отнася към т. нар. апенино-балканска група на романските езици, към която се числят още далматинският и италианските диалекти/езици от централните и южните части на Апенинския полуостров, като всички те произхождат от т. нар. източен латински (Íllyés 1992: 125). Негови особености във вокалната система са съхраняването на дифтонга -au в ударена позиция и редуцирането му до -a, ако е следван от сричка, съдържаща -a или -u, и съхраняването на краткото латинско -ŭ, т.е. характерният за западните романски езици преход -ŭ > o не се наблюдава в проторумънския. В консонантната система като обща особеност на източния латински се отбелязва отпадането на крайното -s, неозвучаването на интервокалните беззвучни оклузивни -p, -t, -k, както и промяната на консонантните групи – -ct към -pt или -ft и -cs към -ps или -fs (Íllyés 1992: 198 – 200).

При прехода от източен народнолатински към проторумънски – около V – VI век, се реализира серия от фонетични промени, които обуславят трансформацията на проторумънския в отделен романски език, ясно разграничим от останалите романски езици.

В областта на вокалите се наблюдава дифтонгизация на краткото латинско -e под ударение към -ie и на неудареното кратко -i към -e (Íllyés 1992: 203). Всъщност последната промяна е на практика повсеместна и е регистрирана в народния латински, говорен в различни райони на Късната римска империя (Herman 2000: 31, 34).

При консонантите се наблюдава поредица от палатализации и асибиляции пред предните гласни -e/-i: -c + e/i > č, но в арумънски и мегелнорумънски -ts; -t + e/i > ts; -d + e/i > z (в румънски) и dz (в арумънски), -g + e/i > dž. Други характерни черти на проторумънската консонантна система е ротацизмът на -l (l > r) в интервокална позиция, промяната на -v към -b и изчезването на интервокалните -v и -b в някои случаи (Íllyés 1992: 203 – 204).

Проторумънският език е споделял и някои общи фонетични черти с албанския, приписвани обикновено на спорен балкански субстратен език, който може да се обозначи и като протоалбански и притежаващ предполагаеми, но неизяснени връзки с тракийския, мизийския или илирийския – -lv > lb, -rv > rb (Íllyés 1992: 227).

2. Регистрирани фонетични особености на латинските надписи от Мизия

Латинските надписи от Мизия представят някои сходства, но и съществени различия с гореописаните промени.

Вокали: добре застъпена е промяната на краткото *-ĭ* към *-e* в ударено и неударено положение (Козарова / Kozarova 1987: 172 – 173). Както подчертахме по-горе, това явление е широко застъпено на територията на целия латиноезичен свят. Съществуват данни за поява на особен звук между *-e* и *-i* на мястото на старото дълго *-ē* след изчезването на квантитета – т.е. след сливането му с краткото *-e* (Козарова / Kozarova 1990: 178 – 179). Наблюдава се и елизия на *-i* преди и след гласни под ударение: *legionis* > *legon*, *Antoniae* > *Antonae*, *filiis* > *filis* (Козарова / Kozarova 1994: 171 – 173).

В късните латински надписи се регистрира и една типична за западния латински черта – преход на краткото ударено *ĭ* към *o* (Бояджиев / Boyadzhiev 2008: 141). Съществуването на „*o*“ латински диалект на Балканите се доказва и от някои латински заемки в албанския, съдържащи необичайно *-o* вместо *-ĭ* (Orel 2000: 29). Тази черта обаче е неприсъща за проторумънския и неговите разклонения. При късните надписи се наблюдава пълно изчезване на старите дифтонги и очевиден преход към монофтонги, както и изчезване на квантитета (Бояджиев / Boyadzhiev 2008: 199).

Сред особеностите на консонантната система на надписите трябва да се споменат съхраняването на лабиала *-b* и липсата на регистрирани преходи *-b* > *v*. Вместо това са налице случаи на поява на *-b* на мястото на *-v* (Козарова / Kozarova 1990: 179 – 180). В покъсните надписи обаче от V – VI век са налични и редица обратни случаи на замяна на *-b* с *-v*, като дори според Д. Бояджиев „двете съгласни са станали неразличими в обикновения говор“ (Бояджиев / Boyadzhiev 2008: 179). Беззвучният дентал *-t* в изгласна позиция изпада, а в отделни примери пред *-i* се палатализира, за което свидетелства изписването му като *-tz*, но тези случаи не са повсеместни. При това в регистрираните примери *-tz* „все още няма статута на фонема, а е позиционен вариант на фонемата *t*“ (Козарова / Kozarova 1996: 118). Наблюдава се промяна *-ct* > *t* (Козарова / Kozarova 1996: 118), както и отпадане на *-n* пред *-s* в късните надписи (Бояджиев / Boyadzhiev 2008: 120 – 121). Езикът на тези надписи споделя и редица общи черти с народния латински, като отпадане на началното *-h* и

акузативното окончание по първо склонение -m (Бояджиев / Boyadzhiev 2008: 117, 200).

Предвид ограниченията в обема на статията ще се спра само на малка част от топонимите и антропонимите, които издават контакти между югоизточните славяни и носителите на езика, отразен в късните мизиолатински надписи.

3. Топоними и антропоними с преход на кратко ѝ към о

- **Ѓгоста** – среднобългарски запис и название на река Огоста. Хидронимът произхожда от името на римската крепост Augustae, построена на устието на тази река в Дунав и спомената няколко пъти в античните извори (Бешевлиев / Beshevliev 1955: 284). В румънските диалекти е съхранена думата *agust*, произхождаща от *augustus* (İlliyés 1992: 199). Очевидно е, че българският хидроним не произлиза от проторумънски. В него е отразена именно промяната -ѝ > -о, регистрирана в късните латински надписи по нашите земи. По отношение на началната гласна може да се предположи типичната за източния латински свят промяна *au* > *a* и последвала лабиализация на -а при усвояването на хидронима в славянски.

- **Маторне горн**: средновековното българско название на централната част на Стара планина (БЕР / BER 1986: 686). В района на Севлиево е открит и среднобългарски надпис, съобщаващ за построяването на манастир, посветен на Света Богородица Маторска (Иванова / Ivanova 1946: 136). Първият компонент очевидно произхожда от народнолатинското **mator/*matora* от *mātūrus/mātūra* посредством промените -ū > ѝ > о. Думата „матор“ е добре позната в българските диалекти със значение на „стар, възрастен“ (Младенов / Mladenov 2008: 251). Тя присъства и в старобългарския речников фонд посредством глагола *зматорѣти* – „остарявам“ (СБ / SB 1999: 523).

- **Роснта** – среднобългарски запис на името на река Росица. Античното име на реката не е запазено в изворите, но според изричното сведение на Владислав Граматик през XV век местното население нарича реката Росита (Ивановъ / Ivanov 1936: 80). Всъщност този хидроним се среща в писмените извори дори и в епохата на Възраждането. Хидронимът се състои от два компонента – корен -*ros* и наставка -*nta*, които насочват към късна народнолатинска адаптация. Коренът най-вероятно има палеобалкански произход, подобно на хидронима *Rhusion*, чието значение се обяснява като „бавно течение“ (Дуриданов / Duridanov 1976: 47). В този случай имаме пример за нов

преход на -u > -o и вероятно народно преосмисляне посредством латинската дума *ros* – „роса“, но също и „влага, вода“ (ЛБР / LBR 1971: 615).

- **Потѣ** – название на местност, известно със своите минерални и топли извори, близо до Кюстендил. Записано е от Йордан Иванов в края на XIX век, по време на неговите теренни проучвания в Северна Македония (Ивановъ / Ivanov 1906: 9). Най-вероятно названието произхожда от латинското *puteus* – „кладенец, цистерна“ (ЛБР / LBR 1971: 574). В румънски и арумънски същата дума дава съответно *puț* и *puts* (пуц).

- **Плодѣвъ** – една от старобългарските адаптации на мизийския и къснолатинския ойконим *Pulpudeva* (Пловдив). Този ойконим е регистриран в изворите през VI век и се обяснява като мизийско калкиране на гръцкото *Φιλίππουπόλις* (Дуриданов / Duridanov 1976: 115). Вариантът **Плодѣвъ** е засвидетелстван в апокрифната българска летопис от XIII век, като мястото на нейното съставяне се определя някъде около София (Петканова / Petkanova 1992: 333). Промяната -u > -o и съкращаването на втората сричка насочва към народнолатински диалект, говорен вероятно в днешна Западна България.

- **Таворъ** – име на селище близо до Скопие, засвидетелствано във *Виргинската грамота на Константин Асен* – средата на XIII век (Даскалова & Райкова / Daskalova & Raykova 2005: 373). Името произхожда от латинското *Tauresium* и по-точно от негов вариант, претърпял промяната -u > -o. За сравнение латинското *taur* дава в румънски *taur* (Șliués 1992: 199).

- **Скопѣ** – среднобългарски запис във *Виргинската грамота* (Даскалова & Райкова / Daskalova & Raykova 2005: 352), от латинското *Scupi* през промяна u > o.

4. Топоними и антропоними с отпадане на -i пред гласна

Както се спомена по-горе, отпадането на -i в хиатус след ударена гласна се регистрира няколко пъти в латинските надписи от Мизия:

- **Арчаръ** – топонимът е споменат за пръв път в житието на свети Теодосий Търновски, съставено през втората половина на XIV век, като название на манастир във *Видинско* (Патриарх Калист / Patriarh Kalist 1986: 445 – 446), а понастоящем се явява име на река и селище от същия район. Няма никакво съмнение, че Арчар произхожда от латинското *Ratiaria*. Топонимът носи следите на дълго развитие в

романо-славянска среда. Синкопата на гласната -i пред -a в *Ratiaria* > *Ирчаръ* съвпада с описаните от Ст. Козарова идентични явления в мизийските латински надписи. Отделно преходът -t > щ, който трябва да се е реализирал преди изпадането на -i, напомня за съобщеното от авторката изписване на -tz вместо -t пред гласна -i (вж. по-горе), което е индикация за започнал процес на палатализация на -t.

- *Доџетѣ* (Домета) – име на български управител през IX в. в днешна Югозападна Македония и Южна Албания (*Theophylactus Achridensis* 1994: 31). Антропонимът произхожда от латинското име *Domitianus*, многократно срещано в латинските надписи от Мизия. Преходът -i > -e, както и отпадането на -i пред a свидетелстват за произход от народния латински. Вариантът Домета трябва да е славянизирана съкратена форма на **Dometan(us)*, по модела на Гавра от Гавраил от X – XI век и вероятно Дристра от началото на X век.

- **Лола* – името се възстановява от битов преславски надпис, датиран към X век: *Лолни праслен* (Гошев / *Goshev* 1961: 80). Най-вероятен източник на този антропоним се явява женският вариант на латинското име *Lollius* – *Lollia* с изпадане на -i.

5. Отпадане на -n пред s

Както бе посочено по-горе, според Д. Бояджиев това е един от типичните белези на езика на късните латински надписи от Мизия. Като илюстративен пример в това отношение той дава антропонима *Sōstantīnus* (Бояджиев / *Boyadzhiev* 2008: 80).

Костадинъ – името е регистрирано няколко пъти в българските средновековни паметници от XIII – XV век, в т.ч. и като народно име на българския цар Константин Асен (Писарев / *Pisarev* 1981: 200). Далечната връзка с езика на късните латински надписи от Мизия проличава не само от отпадането на -n пред -s, но и от озвучаването на интервокалното -t в позиция между две гласни, което е добре засвидетелствано в същите тези надписи (Козарова / *Kozarova* 1996: 114).

Представените примери са само ограничена част от многобройните топоними и антропоними, засвидетелствани в средновековните български паметници, във византийските и в османските извори, както и в българската топонимия и народната антропонимия от къснола-

тински или латинизиран балкански произход², които дават основание произходът на титлата цѣсарь и нейният вариант цѣсарь да се търсят именно в езиковите контакти между носителите на югоизточните славянски говори, преселили се на Балканите през VI – VII в., и представителите на късноантичното латинско или латинизирано население на полуострова³.

III. Народнолатинската хипотеза за произхода на цѣсарь

В късните латински надписи от Мизия титлата *caesar* многократно се използва за обозначаване на владетеля или владетелите в Римската империя – т.е. в значение на император (Бояджиев / Boyadzhiev 2008: 168, 169, 174, 176, 184, 190, 196, 197), като в семантично отношение тя напълно съвпада с по-късното значение на цѣсарь в старобългарския език.

1. Произходът на -ѣ в цѣсарь:

В късния латински език дифтонгът *-ae* претърпява монофтонгизация до отворено *-e* (Herman 2000: 31) – т.е. класическото латинско *Caesar* напълно закономерно би могло да даде формата **Ceasar* в някои или във всички от тези говори. Всъщност тази промяна е засвидетелствана в късните мизийски надписи посредством дативната форма в мн. ч. – *Cesaribus*, в която старият дифтонг *-ae* е заместен от *-e* (Бояджиев / Boyadzhiev 2008: 191 – 192). В този контекст усвояването на отвореното *-e* посредством ѣ изглежда напълно възможно, още повече че има и конкретни примери, в които дори неотвореното *-e* в топоними и заемки от гръцки и латински се адаптира на старобългарски посредством ѣ: Σέρρες > Сѣръѣ > Сяр; *Serdica* > Сѣрдѣицѣ, Сѣрдѣицѣ, Сѣрдѣицѣ; κλιβερτίτѣ – „слуга, роб“ (СБ / SB 1999: 724) < *collibertus*

² Голяма част от тях съм разгледал в отделни публикации на английски език, издадени у нас и в Германия.

³ Трябва да се спомене, че съгласно с някои съвременни интерпретации на ранносредновековни некрополи от Северозападна България от езическия период в погребенията с инхумация се отчита определено влияние на християнския ритуал, което заедно с широката употреба на амфорна керамика, дава основания за хипотезата, че в тях са погребани потомци на късноантичното население (Стайков / Staykov 2020: 38 – 43), а то в този регион определено е било латиноезично. Континуитет между Ранното средновековие и Късната античност в погребалните практики е силно изразен и в некрополите от днешна Югозападна България (Грозданова / Grozdanova 2018: 153 – 155, 161 – 162).

(„другар по освобождение“ – за освободени роби), *μρῆγορνα* – „тържище“ (СБ / SB 1999: 872), от латинското *mercog* („търговия“, старобългарското окончание издава гръцко посредничество) и др.

2. Произходът на крайното -ъ

Не може да се изключи стандартното обяснение за уподобяване с наставката за *nomina agentis* -арь, но трябва да се вземе под внимание и обстоятелството, че в късния латински, засвидетелстван в мизийските надписи, класическата латинска падежна система се разпада и в първа деklinация винителният падеж след отпадането на -m поема функциите на общ падеж. Във втора деklinация също се наблюдава съвпадение между акузатива и датива вследствие на фонетични промени при акузатива и поемането от тази нова форма на ролята на общ падеж (Бояджиев / Boyadzhiev 2008: 150 – 154, 200 – 201, 209, 126, 163). За трета деklinация, към която се отнася и *Caesar*, положението не е толкова ясно, но и тук не може да се изключи развитието към обща форма **Cesare* от акузатива *Caesarem* или **Cesari* от датива *Caesari*, което да обясни и появата на ь в ауслаута на цѣсарь.

3. Произход на интервокалното -с

За разлика от езиците на проторумънската група, в които -s пред е/і обикновено се изяснява към -щ, в езика на късните латински надписи от Мизия етимологичният латински сибилант -s пред предна гласна и в интервокална позиция не показва промени. В този контекст усвояването на -s от **Cesar*/**Cesari* посредством -с в цѣсарь не буди никакви възражения, още повече че в български съществуват и други примери в това отношение – напр. босилек < *basilicum* (БЕР / BER 1971: 69), споменатият по-горе хидроним Росита и др.

4. Произход на началното -ц

На пръв поглед появата на -ц в начална позиция може да се отдаде на втората палатализация в праславянския⁴. Трябва да се има

⁴ Известно съмнение върху тази теза може да хвърли паралелната гръцка заемка Кесарь от гръцкото *Καῖσαρ* (СБ / SB 1999: 721). В „Грамматика на старобългарския език“ тази заемка изрично е представена като книжовна (Илчев / Ilchev 1991: 115). Първата гласна -е обаче е индикация, че думата в старобългарски се явява не книжовна заемка от гръцки, а адаптация от говоримия гръцки, в който дифтонгът *αῖ* вече е бил дал -е. В противен случай би следвало да се очаква изписване *кансарь* по

предвид обаче, че палатализацията на -с пред предна гласна е добре познато явление и в романските езици. В езиците от проторумънски произход съществува известно разминаване в това отношение: в румънски и истрорумънски -с пред -е/ -і дава редовно -щ, което се счита за първичното състояние в проторумънски. В арумънски обаче на място на румънското -щ се регистрира -ts/-ц (İlliés 1992: 203), което засега трудно може да се обясни. Интересен и дори уникален е случаят на далматинския език. В него се наблюдава съхраняване на -с пред -е, докато пред -і той преминава в -щ. Съществуват индикации, че в балканските романски говори от непроторумънски произход също се е наблюдавало известно разнообразие в това отношение. Средновековни български антропоними и топоними дават основание да се счита, че в някои романски говори от Централна Македония -с пред предна гласна -е претърпял асибиляция, т.е. е преминал в -s – една особеност, която се наблюдава в някои западни романски езици.

- Αλουσίανος (Алусиан) – син на цар Иван Владислав – без съмнение от латинското Lucian(us) с протетично -а и преход -с > -s. Името Lucianus е добре засвидетелствано през античния период в надписи от Горна Мизия (Mirković, Dušanić, 1976: 163; Petrović 1979: 136).

- Васа̀нѣтъ – име, регистрирано в надпис от Търново, датиран към XIII век (Алексиев / Aleksiev 1984: 237), и явяващо се най-вероятно вторична балканска народнолатинска адаптация на гръцкото ὑάκινθος – с преход на -к > -s и -ι > -е.

- Сетина / Σέτινα – селище в съвременна Северозападна Гърция, чието име е променено на Σκολός през 1926 г. В края на X и началото на XI век в него се намира една от резиденциите на последната българска царска династия от Първото царство (Cedrenus & Skylitzes 1965: 289). Българският ойконим е вариант на добре известните от района на Западните Балкани селищни и речни названия „Цетина“. Сред тях са и бившата столица на Черна гора – Cetinje, която на гръцки се произнася като Κετίυνη, и река Cetina (Tsetina) в Южна Далмация. Топонимът се среща и на север от Балканите в Централна Европа, където също се явява с начално -к. Хърватските и черногорските „ц“ варианти трябва да се отдадат на процесите на

подобие на канѣ от Κάιν (СБ / SB 1999: 708). Още Бешевлиев в монографията си върху прабългарските надписи на гръцки език изрично отбелязва, че този дифтонг редовно се заменя с -е, което следва разволя в говоримия гръцки (Бешевлиев / Beshevliev 1992: 49).

палатализация на -к пред предни гласни най-вече в късния протославянски, доколкото далматинският в подобни случаи съхранява старото произношение. Българският „s“ вариант обаче издава влиянието на романския говор, донор на антропонимите Алусиан и Васенит.

Наред с описаните индикации за асибиляция съчинението на Прокопий Кесарийски „За строежите“ дава немалко примери за преход -с > -ts в района на днешна Източна Сърбия и Северозападна България още преди идването на славяните: Τζερτζενοούτζας, Μουτζιανικάστειλλον, Βιτζιαίαις и др. (Procopius Caesariensis 1958: 161 – 162).

Един топоним и един антропоним в български и старобългарски също дават известни основания да се счита, че в късния латински език в Мизия е била реализирана промяна -с > -ts преди идването на славяните:

- Цибър – име на река в Северозападна България. Античното име на реката е известно от името на станцията, построена на нейните брегове – Cebro (Itinerarium Antonini Augusti 1958: 31).

- Потръць – името на военен командир от края на XIV век. Името е записано в Синодика на цар Борил, параграф 168 (Попруженко / Popruzhenko 1928: 94). В зависимост от характера на еровата гласна -ъ оригиналният гласеж на името може да се реконструира като Потруц – в случай че наистина въпросната гласна е правилно поставена, т.е. тя е голям ер, или Потриц, в случай че еровите гласни са объркани, което е често явление в писмените паметници. В първия случай името би могло да има проторумънски произход и може да се свърже с румънското Rătruț, което е засвидетелствано и в съвременната българска антропонимия (Янев / Yanev 2009: 315). Лабиализацията на началната гласна обаче насочва към един ранен период на адаптиране на антропонима – преди XI век, и в такъв случай името би било единственият пример за толкова ранно заимстване на румънски антропоним, което като цяло изглежда съмнително. Освен това Rătruț най-вероятно е късно развитие на стандартното Petruț, като не е ясно дали през Ранното средновековие такова име изобщо е съществувало в антропонимната система на (прото)румънците, тъй като негови записи е трудно да се открият дори в османските и влахо-българските документи от XIV – XVI век. При втория вариант името лесно може да се етимологизира въз основа на латинското patricius, преминало през прехода -с > -ts пред -i. В този случай обаче не може да се изключи и влияние на третата палатализация в южнославянските езици.

В заключение появата на началното -ц в цѣсарь може да се отдаде както на действието на втората палатализация в славянски, така

и на вътрешно развитие в къснолатинските диалекти, евентуални донори на титлата.

IV. Народнолатинската хипотеза за цесарь

Ойконимът *Плодівъ* поставя един интересен въпрос – именно за прехода на крайната гласна -е към -і, който очевидно е настъпил от началната форма *Pulpudeva*. И. Дуриданов защитава тезата за тракийската основа на този преход, като счита, че той идва от тракийското -і, което се явява развитие от първоначално тясно -е (Дуриданов / Duridanov 1976: 115). Трябва обаче да се отбележи, че преход -е към -і се наблюдава и в българската дума тикла от латинската *tegula* (Младенов 2008: 261), в споменатия хидроним Цибър от *Sebro*, в оронима Им от *Naemus* (Дуриданов / Duridanov 1976: 37). За това, че Им произхожда от народен латински говор, ясно свидетелства отпадането на началния консонант -h – явление, характерно за народния латински и добре засвидетелствано в късните латински надписи от българските земи. В същото време имаме изрично свидетелство, че през VI век в езика на оцелялото тракийско население началното -h е било съхранено. Когато Йорданес говори за река Дунав, той изрично пише „*Qui lingua Bessorum Hister vocatur*“ – „която на езика на бесите се нарича Хистер“ (Sturm 2018: 60).

Допълнителна светлина върху проблема могат да хвърлят и няколко латински заемки в албанския, където вместо очакваната адаптация на краткото латинско *ĕ* посредством дифтонга -je или -e се явява -i, така както се усвоява краткото латинско -i, т.е. най-вероятно думите са били преминали през промяната -e > -i още в рамките на езика донор. Преценката на Владимир Орел – автор на едно подробно изследване върху развитието на протоалбанския, е, че тези думи са заимствани от „различен тип староромански“ (Orel 2000: 26). Всъщност този „различен тип староромански“ трябва да е някой романски говор от района на днешна Западна България, Източна Сърбия и северните части на Македония. Основания за такова заключение дават както семантиката, така и фонетичният облик на заемките: „*grigji* < лат. *gregem*, *gjinde* < лат. *gentem*, *mëshirë* < лат. *miseria*, *nip* < лат. *perōs*, *prifit* < лат. *presbyter*, *print* < лат. *parentem*“ (Orel 2000: 27). Думите *grigji*, *gjinde*, *print* са заимствани от техните акузативни форми след отпадането на акузативното окончание -m, а при *prifit* има очевиден преход -b > -v в изходната латинска форма – т.е. заимстването трябва да се е случило в периода на Късната античност. Към това ни

насочва и значението на две от думите: *prifit* – „презвитер“, и *grigji* – „стадо“, но също и „паство“ в религиозния/християнския смисъл на думата. Очевидно е, че думите са заимствани от християнско население, което е имало и своя църковна организация. Последното е времевата индикация за V – VI век, когато са създадени митрополията на Сердика и подчинената ѝ епископия на Ремисиана. Двете се намирили под властта на Рим, използвали са латинския за богослужебен език и са изиграли определена роля за християнизирването на траките. Не е изключено въпросните две думи да са се появили в албанския език като резултат от някакви ранни опити за християнизация на протоалбанците, чиито първоначални територии са се намирали или в диоцеза на Ремисиана и Сердика, или в непосредствена близост до тях.

Същевременно явлението $e > i$ донякъде е отразено и в латинските надписи от Мизия. Ст. Козарова открива специфична гласна между $-i$ и $-e$, която се явява на мястото на старото дълго $-ē$ след изчезването на квантитета и сливането с краткото $-ĕ$ (Козарова / Kozarova 1990: 178 – 179). Д. Бояджиев също говори за „смесвания на i -звук и e -звук“, като отбелязва наличието на подобни паралелни явления и в тракийския, но се въздържа от мнение доколко тези прояви в езика на късните латински надписи на Балканите са с тракийски произход (Бояджиев / Boyadzhiev 2008: 200). Един от изследваните от него надписи от Варна съдържа пример за промяната $-e$ към $-i$: *vixillatione* вместо *vexillatione*, но той изрично отбелязва и крайно неправилния език на надписа (Бояджиев / Boyadzhiev 2008: 156).

Прави впечатление, че както при Им, така и при албанската *prifit* $-i$ се появява на мястото на стария дифтонг $-ae$ (*Naemus / praesbiter*) – това навежда на мисълта, че преходът $-e > -i$ е обхванал и новото отворено $-e$, резултат от монофтонгизацията на $-ae$. Последното дава основания за хипотезата, че в Късната античност в Мизия и вероятно в други части на Централните Балкани по отношение на $-e$ са съществували два типа народен латински: един, който е стоял близо до развитието в другите части на латиноезичния свят, и друг, в който $-e$ се е редуцирало до $-i$. Възможно е зад втория наистина да стои местен балкански език, който е бил различен от протоалбански, евентуално тракийски. В този контекст не може да се изключи вероятността старобългарското **цѣсарь** да има за основа вариант на засвидетелстваното в мизийските надписи *Cesar* със силно редуцирано $-e$.

Що се отнася до сърбо-хърватско-словенския „ $-e$ “ вариант на титлата, тук трябва да се вземат предвид и да се проучат допълнително редица други фактори, като различните далматински диалекти, панонския

латински, влиянието на църковния латински върху хърватския и словенския. Под внимание трябва да се вземе и изричното сведение от началото на XV век на Константин Костенечки, че сръбските книжовници не разбирали смисъла на буквата „Ѣ“ и я подменяли с ѳ или с ю (Константин Костенечки / Konstantin Kostenechki 1993: 32 – 33).

Проблемът за произхода на титлата е само елемент от по-широкия проблем за съдбата на късноантичното население по българските земи през Ранното средновековие, за езиковите взаимодействия между югоизточните славянски говори и езиците, разпространени в късноантичните Балкани. Същевременно тя е елемент и от „влашкия проблем“, който засяга историята на Второто българско царство, доколкото редица топоними и антропоними с романски произход от XIII век, както и съхранените в Норовския псалтир 12 глоси от „неизвестен език“ също издават далечни връзки с езика на късните мизиолатински надписи, а не с езиците, произлезли от проторумънския. Моето убеждение е, че без задълбочено изследване на тези връзки редица въпроси, засягащи историята на средновековна България и историята на българския език (например процесите на неговата „балканизация“), ще останат неизяснени.

ЦИТИРАНА ЛИТЕРАТУРА

- Cedrenus & Skylitzes 1965: Cedrenus, G. & Skylitzes, I. *Historiarum compendium*. In: Tapkova, Z. (Ed.). *Fontes Graeci Historiae Bulgaricae IX*. Sofia: BAN, 198 – 340.
- Herman 2000: Herman, J. *Vulgar Latin*. Pennsylvania: The Pennsylvania State University Press. ISBN: 978-0-271-02001-3.
- İllyés 1992: İllyés, E. *Ethnic Continuity in the Carpatho-Danubian Area*. Hunyadi ÖKS, Hamilton ON: Structura Press. ISBN 0-88033-146-1.
- Itinerarium Antonini Augusti 1958: *Itinerarium Antonini Augusti*. In: Gerov, B. (red.). *Fontes Latini Historiae Bulgaricae I*. Sofia: BAN, 28 – 38.
- Mirković & Dušanić 1976: Mirković, M., Dušanić, S. *Inscriptions de la Mésie Supérieure; vol. I, Singidunum et le nord-ouest de la province*. Beograd: Centre d'études épigraphiques et numismatiques.
- Orel 2000: Orel, V. *A Concise Historical Grammar of the Albanian Language, Reconstruction of Proto-Albanian*. Brill. ISBN: 978-90-04-11647-4.
- Petrović 1979: Petrović, P. *Inscriptions de la Mésie supérieure IV. Naisus - Remesiana – Horrerrum Margi*. Beograd: Centre d'études épigraphiques et numismatiques.

- Procopius Caesariensis 1958: Procopius Caesariensis. De aedificiis. In: Duychev, I. (Eds.). *Fontes Graeci Historiae Bulgaricae* II. Sofia: BAN, 154 – 178.
- Pronk-Tiethoff 2013: Pronk-Tiethoff, S. *The Germanic Loanwords in Proto-Slavic*. Amsterdam / New York: NY. ISBN: 978-90-420-3732-8.
- Sturm 2018: Sturm, R. *Jordanes: Die Geschichte der Goten: Übersetzt und eingeleitet*. Buch.
- Theophylactus Achridensis 1994: Theophylactus Achridensis. Vita S. Clementis Achridensis. In: Iliiev, I. (red.). *Fontes Graeci Historiae Bulgaricae* IX. Sofia: BAN, 10 – 41.
- Алексиев 1984: Алексиев, Й. Среднобългарски надписи и графити от Царевец (XIV). В: Гюзелев, В. (ред.). *Сборник в памет на проф. Станчо Ваклинов*. София: БАН, 236 – 238. [Aleksiev 1984: Aleksiev, Y. *Srednobałgarski nadpisi i grafiti ot Tsarevets (XIV)*. V: Gyuzelev, V. (red.). *Sbornik v pamet na prof. Stancho Vaklinov*. Sofia: BAN, 236 – 238.].
- БЕР = *Български етимологичен речник*. I. 1971. София: БАН [BER = *Balgarski etimologichen rechnik*. I. 1971. Sofia: BAN.].
- БЕР = *Български етимологичен речник*. III. 1986. София: БАН. [BER 1986 = *Balgarski etimologichen rechnik*. III. 1986. Sofia: BAN.].
- Бешевлиев 1955: Бешевлиев, В. Латинските местни имена в Мизия и Тракия. – *Известия на Археологическия институт. Сборник Гавраил Кацаров*. 19. БАН, 279 – 303. [Beshevliev 1975: Beshevliev, V. *Latinskite mestni imena v Mizia i Trakia*. – *Izvestiya na Arheologicheskiya institut. Sbornik Gavrail Katsarov*. 19. BAN, 279 – 303.].
- Бешевлиев 1975: Бешевлиев, В. Първобългарски и късни гръцки и латински надписи. – *Археология*, 1, 15 – 24 [Beshevliev 1975: Beshevliev, V. *Parvobałgarski i kasni gratski i latinski nadpisi – Arheologiya*, 1, 15 – 24.].
- Бешевлиев 1992: Бешевлиев, В. Първобългарски надписи. София: БАН [Beshevliev 1992: Beshevliev, V. *Parvobałgarski nadpisi*. Sofia: BAN.].
- Бояджиев 2008: Бояджиев, Д. Езикът на късните латински надписи в България. В: Славова, М. (съст.) *Opera Minora Selecta*. София: Университетско издателство „Свети Климент Охридски“, 105 – 209. [Boyadzhiev 2008: Boyadzhiev, D. *Ezikat na kasnite latinski nadpisi v Balgaria*. V: Slavova, M. (sast.), *Opera Minora Selecta*. Sofia: Universitetsko izdatelstvo „Sveti Kliment Ohridski“, 105 – 209.]. ISBN: 978-954-07-2717-2.
- Гошев 1961: Гошев, И. *Старобългарски глаголически и кирилски надписи от IX и X в.* София: БАН. [Goshev 1961: Goshev, I. *Starobałgarski glagolicheski i kirilski nadpisi ot IX i X v.* Sofia: BAN.].
- Грозданова 2018: Грозданова, Г. Населението на Южна България VI – IX в. (по археологически данни). Дисертации. Том 12. София: Национален археологически институт и музей – БАН. [Grozdanova 2018: Grozdanova, G. *Naselenieto na Yuzhna Balgaria VI – IX v. (po arheologicheski dannii)*. Disertatsii. Tom 12. Sofia: Natsionalen arheologicheski institut i muzey – BAN.]. ISSN 2603-3100 (online); ISSN

- 2603-3097 (print); ISBN 978-954-9472-66-0 (print); ISBN 978-954-9472-67-7 (online).
- Даскалова & Райкова, 2005: Даскалова, А. Райкова, М. *Грамоти на българските царе*. София: АИ „Проф. Марин Дринов“. [Daskalova & Raykova, 2005: Daskalova, A., Raykova, M. *Gramoti na balgarskite tsare*. Sofia: AI „Prof. Marin Drinov“]. ISBN 9543220344.
- Дуриданов 1976: Дуриданов, И. *Езикът на траките*. София: Наука и изкуство. [Duridanov 1976: Duridanov, I. *Ezikat na trakite*. Sofia: Nauka i izkustvo.].
- ЕПРБКЕ = Младенов, С. *Етимологически и правописен речник на българската книжовен език*. 1941. София: Христо Г. Дановъ. [EPRBKE = Mladenov, S. *Etimologicheski i pravopisen rechnik na balgarskiya knzhoven ezik*. 1941. Sofia: Hristo G. Danov.].
- ЭСРЯ = Фасмер, М. *Этимологический словарь русского языка*. IV. 1986. Москва: Прогресс. [ESRY Fasmer, M. *Étimologicheskij slovar' russkogo yazyka*. IV. 1986. Moskva: Progress.].
- Ивановъ 1906: Ивановъ, Й. *Съверна Македония. Исторически издирвания*. София. [Ivanov 1906: Ivanov, Y. *Severna Makedonia. Istoricheski izdirvaniya*. Sofia.].
- Иванова, 1946: Иванова, В. *Два надписа от Асеновци – Батошевският и Врачанският*. – *Известия на Българския археологически институт*, 1946, XV, 135 – 137. [Ivanova, 1946: Ivanova, V. *Dva nadpisa ot Asenovtsi – Batoshevskiyat i Vrachanskiyat*. – *Izvestiya na Balgarskiya arheologicheski institut*, 1946, XV, 135 – 137.].
- Илчев 1991: Илчев, П. Влияние на предните гласни върху веларните съгласни: трите палатализации в праславянски и тяхното отражение в старобългарския език. В: Дуриданов, И. (ред.). *Грамматика на старобългарския език. Фонетика, морфология, синтаксис*, 113 – 119. София: БАН. [Ilchev 1991: Ilchev, P. Vliyanie na prednite glasni varhu velarnite saglasni: trite palatalizatsii v praslavyanski i tyahnoto otrazhenie v starobalgarskiya ezik. V: Duridanov, I. (red.). *Gramatika na starobalgraskiya ezik. Fonetika, morfologiya, sintaksis*, 113 – 119. Sofia: BAN.].
- Козарова 1987: Козарова, С. За някои промени на вокала і в латинските надписи от територията на Мизия (между р. Янтра и р. Искър). – *Научни трудове на Пловдивския университет „Паусий Хилендарски“*, том 25, кн. 1. Пловдив, 171 – 180. [Kozarova 1987: Kozarova, S. Za nyakoi promeni na vokala i v latinskite nadpisi ot teritoriyata na Mizia (mezhdur. Yantra i r. Iskar). – *Nauchni trudove na Plovdivskiya universitet „Paisii Hilendarski“*, том 25, кн. 1. Plovdiv, 171 – 180.].
- Козарова 1990: Козарова, С. За някои вокални промени в народния латински от нашите земи (върху епиграфски материал). – *Научни трудове на Пловдивския университет „Паусий Хилендарски“*, том 28, кн. 1. Пловдив, 178 – 182. [Kozarova 1990: Kozarova, S. Za nyakoi vokalni promeni v narodniya latinski ot nashite zemi (varhu epigrafski material). – *Nauchni trudove na Plovdivskiya universitet „Paisii Hilendarski“*, том 28, кн. 1. Plovdiv, 178 – 182.].

- promeni v narodniya latinski ot nashite zemi (varhu epigrafski material. – *Nauchni trudove na Plovdivskiya universitet „Paisii Hilendarski“*, tom 28, kn. 1. Plovdiv, 178 – 182.].
- Козарова 1994: Козарова, С. Синкопа и апокопа на предни гласни в латинските надписи от областта между реките Искър и Янтра. – *Научни трудове на Пловдивския университет „Паусий Хилендарски“*, том 32, кн. 1. Пловдив, 169 – 174. [Kozarova 1994: Kozarova, S. Sinkopa i apokopa na predni glasni v latinskite nadpisi ot oblastta mezhdru rekite Iskar i Yantra. – *Nauchni trudove na Plovdivskiya universitet „Paisii Hilendarski“*, tom 32, kn. 1. Plovdiv, 169 – 174.].
- Козарова 1995: Козарова, С. Вокалът а в латинския език по нашите земи (върху епиграфски материал от областта между реките Искър и Янтра). – *Научни трудове на Пловдивския университет „Паусий Хилендарски“*, том 33, кн. 1. Пловдив, 35 – 39. [Kozarova 1994: Kozarova, S. Vokalata a v latinskiya ezik po nashite zemi (varhu epigrafski material ot oblastta mezhdru rekite Iskar i Yantra). – *Nauchni trudove na Plovdivskiya universitet „Paisii Hilendarski“*, tom 33, kn. 1. Plovdiv, 35 – 39.].
- Козарова 1996: Козарова, С. Особености в консонантизма на латинските надписи от територията между реките Искър и Янтра. – *Научни трудове на Пловдивския университет „Паусий Хилендарски“*, том 34, кн. 1. Пловдив, 111 – 119. [Kozarova 1996: Kozarova, S. Osobenosti na konsonantizma na latinskite nadpisi ot teritoriyata mezhdru rekite Iskar i Yantra. – *Nauchni trudove na Plovdivskiya universitet „Paisii Hilendarski“*, tom 34, kn. 1. Plovdiv, 111 – 119.].
- Константин Костенечки 1993: Константин Костенечки. *Съчинения*. София: Славика. [Konstantin Kostenechki 1993: Konstantin Kostenechki. *Sachineniya*. Sofia: Slavika.].
- ЛБР = Войнов, М. & Милев, А. *Латинско-български речник*. 1971. София: Наука и изкуство. [LBR = Voynov, M. & Milev, A. *Latinsko-balgarski rechnik*. 1971. Sofia: Nauka i izkustvo.].
- Младенов 2008: Младенов, М. *Ареална характеристика на романски елементи в българските диалекти. Диалектология. Балканистика. Етнолингвистика*. София: Университетско издателство „Св. Климент Охридски“. [Mladenov 2008: Mladenov, M. *Arealna harakteristika na romanski elementi v balgarskite dialekti. Dialektologiya. Balkanistika. Etnolingvistika*. Sofia: Universitetsko izdatelstvo „Sv. Kliment Ohridski“.]. ISBN: 978-954-07-2684-7.
- Патриарх Калист 1986: Патриарх Калист. *Пространно житие на Теодосий Търновски*. В: Иванова, К. (ред.). *Стара българска литература, IV. Житиеписни творби*. София: Български писател, 443 – 467. [Patriarh Kalist 1986: Patriarh Kalist. *Prostranno zhitie na Teodosiy Tarnovski*. V: Ivanova, K. (red.). *Stara balgarska literatura, IV, Zhitiepisni tvorbi*. Sofia: Balgarski pisatel, 443 – 467.].

- Петканова 1992: Петканова, Д. *Старобългарската литература IX – XVII в.* София: Университетско издателство „Св. Климент Охридски“ [Petkanova 1992: Petkanova, D. *Starobalgarskata literatura IX – XVII v.* Sofia: Universitetsko izdatelstvo „Sv. Kliment Ohridski“].
- Писарев 1980: Писарев, А. Още за новооткрития занаятчийски квартал на хълма Царевец. В: Петров, П. (ред.). *Средновековният български град.* София: Българско историческо дружество, 205 – 213. [Pisarev 1980: Pisarev, A. Oshte za novootkritiya zanayatchiyski kvartal na halma Tsarevets. V: Petrov, P. *Srednovekovniyat balgarski grad.* Sofia: Balgarsko istorichesko druzhestvo, 205 – 213.].
- Попруженко 1928: Попруженко, Г. *Синодик царя Борила.* Български старини. Книга VIII. София: Държавна печатница. [Popruzhenko 1928: Popruzhenko, G. *Sinodik tsarya Borila.* Balgarski starini. Kniga VIII. Sofia: Darzhavna pechatnitsa.].
- Стайков 2020: Стайков, В. Етно-културни връзки по Долен и Среден Дунав през Ранното средновековие, VII – IX в. (по данни на некрополите). *Автореферат на дисертация.* София, 2020. [Staykov 2020: Staykov, V. Etno-kulturni vrazki po Dolen I Sreden Dunav prez Rannoto srednovekovie, VII – IX v. (po dannii na nekropolite). *Avtoreferat na disertatsiya.* Sofia, 2020.].
- СР = Старобългарски речник. I. София: Валентин Траянов. 1999 [SR = Starobalgarski rechnik. I. Sofia: Valentin Trayanov. 1999.]. ISBN: 9549928039.
- СР = Старобългарски речник. II. София: Валентин Траянов, 2009 [SR = Starobalgarski rechnik. II. Sofia: Valentin Trayanov, 2009.]. ISBN: 9789549928389.
- Янев 2009: Янев, Б. *Система на личните имена в българския и немския език.* Пловдив: Университетско издателство „Паисий Хилендарски“. [Yanev 2009: Yanev, B. *Sistema na lichnite imena v balgarskiya i nemskiya ezik.* Plovdiv: Universitetsko izdatelstvo „Paisii Hilendarski“].

Assoc. Prof. Hristo Ivanov Saldzhiev, PhD

Trakia University

Stara Zagora, Bulgaria

e-mail: hristosaldzhiev@yahoo.com