

From rigid to flexible verb-finality: A prosodically motivated information-structural account of word order change

Erika Asztalos¹ Lena Borise¹ Katalin Gugán¹ Nikolett Mus¹
Andreas Schmidt² Balázs Surányi^{1,3}

¹Hungarian Research Centre for Linguistics

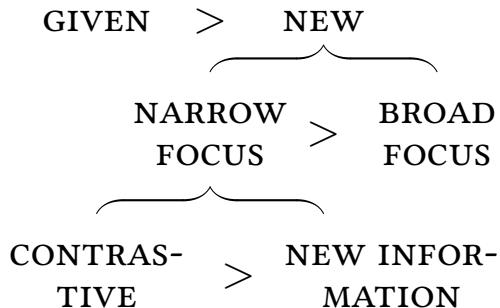
²University of Potsdam, SFB 1287

³Pázmány Péter Catholic University

ICHL25, University of Oxford, 1–5 August 2022

- Information-structural (IS) factors have been shown to play an important role in determining the **availability of postverbal constituents (PVCs)** in languages that undergo a change from more to less rigid V-finality.
- Specifically, **given/backgrounded PVCs** have been noted to appear first, later followed by **new/focused PVCs** (Surányi & Tánzos 2011; É. Kiss 2014; Asztalos et al. 2017).

- We provide extensive evidence for the GIVEN > NEW generalization regarding the development of PVCs and further show that:
 - (i) different languages may distinguish **different subtypes** of GIVEN and NEW (i.e., non-given), which may make the observed progression of events look quite different in individual languages.
 - (ii) among NEW PVCs, there is a strong preference for **NARROWLY** focused ones to become available **earlier** than those that are part of **BROAD FOCUS**.
 - (iii) among **postverbal** **NARROW FOCI**, **CONTRASTIVE FOCI** appear earlier than **NEW-INFORMATION FOCI**.



This succession may seem odd:

- First, **deaccented, prosodically non-prominent** GIVEN material is allowed as a PVC.
- Next, **prosodically prominent** (NARROWLY FOCUSED, especially CONTRASTIVELY FOCUSED) material is allowed as a PVC.
- Finally, material that is part of BROAD FOCUS and bears **default prominence** is allowed as a PVC.

- We account for this seeming paradox and derive the generalizations above from the **prosodic properties of the postverbal domain in V-final languages**.
- We support this analysis by evidence from multiple V-final languages, focusing on **Uralic languages** (but also bringing in evidence from other language families).

Components of the analysis

Prosodic structure:

- **Phonological Phrases** (ϕ s) consist of one or more prosodic words, and are roughly equivalent to **syntactic phrases (XPs)**.
- **Intonational Phrases** (ι s) consist of one or more Phonological Phrases and are roughly equivalent to **syntactic clauses** (Selkirk 2011, Ito & Mester 2013, Elfner 2015, a.o.).

Components of the analysis

- **Nuclear stress (NS)** in V-final languages, as a default, targets the **rightmost VP-internal argument**.
 - or, in its absence, the verb (Truckenbrodt 2006, Büring 2012)
- One way of formalizing NS placement is via alignment of NS with an **edge of a prosodic constituent** (Féry 2013):

- (1) **H- ι -R:**
Align the right boundary of every ι with its head (i.e., the most prominent ϕ).

Components of the analysis

- The clause-final **verb is ‘skipped’** for the purposes of NS placement.
- This may be achieved e.g. by a high-ranked constraint STRESS-XP, which ensures that within a YP properly containing Y^0 and XPs, **XPs but not Y^0 will carry stress** (Truckenbrodt 2006; see also the Non-head stress principle, Duanmu 2004).

Components of the analysis

- Our starting assumption is that information structure (IS) is **not** directly represented in the syntactic structure. Instead, **syntactic structure** interacts with **prosodic structure** (and IS-categories may have prosodic requirements).
- With regard to syntax, we assume for concreteness that VPs in the V-final languages discussed here are **head-final** at basic structure (i.e., we do **not** assume that we are dealing with the change of the direction of headedness within the VP at this point).

Components of the analysis

- The PVCs we examine are licensed even if we put aside any examples that might plausibly involve the **raising of the verb to the left** (i.e, contexts involving raised narrow foci, wh-questions, and possible V2 constructions).
- We take the PVCs discussed here to be derived via **phrasal displacement to the right** (whether by movement or base-generation).
- This approach is supported by the fact that in each of our languages **external arguments/subjects** may appear as PVCs, which would not be easily derivable via the change in headedness or (short) verb movement.

Stage 1: **No PVCs are allowed.**

- The position of nuclear stress is **aligned with the right edge** of an Intonational Phrase (ι), with the verb ‘skipped’ for the purposes of NS assignment.
- No clause-internal PVCs are allowed, because they would separate the NS-marked constituent from the right ι -edge, and would render the prosodic structure non-optimal.

$$\{ \dots \text{XP} \mathbf{V} \}_\iota$$

- The only kind of PVC allowed is a prosodically and syntactically non-integrated **afterthought**.
- Prosodically non-integrated afterthoughts are available even in strictly V-final languages (Simon 1989, Kim & Shin 1992).
- We take afterthoughts to form **their own** ι s, and do not take the availability of afterthoughts to be a step towards less rigid V-finality.

{ ... XP **V** } $_{\iota}$ {**Afterthought**} $_{\iota}$

Stage 2: GIVEN (**deaccented**) PVCs are allowed.

- GIVEN material is obligatorily deaccented – therefore, in the presence of a GIVEN PVC, **NS is still in its default position**, on the (preverbal) rightmost VP-internal argument, and the overall prosodic make-up of the clause is unchanged.
- We propose that GIVEN PVC material is **outside of the core** ι and forms its own lower-order prosodic constituent, a Phonological Phrase (ϕ).

$$\{ \dots \text{XP } \mathbf{V} \}_{\iota} (\text{PVC})_{\phi}$$

Stage 3: NEW (**accented**) PVCs are allowed.

- PVCs being possible leads to the emergence of **nested** ι s, in which the PVC is part of the ‘maximal’ but not the ‘core’ ι .
- Being higher on the prosodic hierarchy, the ‘**maximal**’ ι attracts the most prominent, accented prosodic constituent to its edge – as opposed to it aligning with the edge of the ‘core’ ι .
- The **marked status** of a prosodic structure with nested ι s may facilitate its use in a marked context: one containing a NEW/ NARROWLY FOCUSED constituent.

$\{\{ \dots \text{XP } \mathbf{V} \}_{\iota} (\mathbf{PVC})_{\phi} \}_{\iota}$

- ⇒ Furthermore, because CONTRASTIVE FOCUS is **more marked** as an information-structural category, and, accordingly, may be more marked in its grammatical form than NEW-INFORMATION FOCUS (Repp 2016, Cruschina 2021), its emergence as a PVC may occur first.

Stage 4: **PVCs that are part of BROAD FOCUS (carrying default prominence) are allowed.**

- The right edge of the ‘maximal’ ι is reinterpreted as that of the ‘core’ ι , and the head aligned with it becomes the only available one.
- NS on the PVC becomes an **unmarked option**, and NEW PVCs that are part of BROAD FOCUS become allowed.

{ ... XP **V** **PVC** } $_{\iota}$

- We observe that many languages that are on the trajectory from more to less rigid V-finality, including **Uralic languages**, conform to the sequence of events proposed here and exemplify **different points** of the GIVEN>NEW trajectory.

Cross-linguistic evidence: Tundra Nenets

- **Tundra Nenets** represents the most rigidly V-final end of the spectrum: it is described as strictly V-final (Tereščenko 1973, Nikolaeva 2014).

(2) a. **xasawa tolábi padar-m?*
man read.3SG book-ACC
(‘A/The man reads a/the book.’) (elicited data)

b. **Irina-m? mǎneʔŋa-ś Pavel.*
Irina-ACC see-3SG.PST Pavel
(‘Pavel saw Irina.’) (elicited data)

{ ... XP **V** }_l

Cross-linguistic evidence: Tundra Nenets

- Tundra Nenets, as we show, only allows for **clause-external, afterthought-like prosodically separate PVCs** that are associated with a clause-internal correlate. The correlate may be **overt or covert**.

(3) a. *xasawa fuku-mʔ_i tolabi(-da), fuku padar-mʔ_i.*
man this-ACC read.3SG(SG.3SG) this book-ACC
'A/The man reads this, this book.' (elicited data)

b. *xasawa pro_i tolabi-da, fuku padar-mʔ_i.*
man pro read.SG.3SG this book-ACC
'A/The man reads [it], this book.' (elicited data)

{ ... XP V }_l {Afterthought}_t

Cross-linguistic evidence: Turkish

- In **Turkish**, according to the literature, only GIVEN PVCs are allowed, which undergo obligatory **deaccenting** (Özge & Bozsahin 2010; Öztürk 2013; Göksel 2013).

- (4) *Semra-ya ben gönder-miş-ti-m dün hediyе-yi.*
Semra-DAT 1SG.NOM send-PRF-P-1SG yesterday gift-ACC
'I sent the gift to Semra yesterday.' (Göksel 2013)

{ ... XP V }_t (PVC)_φ (PVC)_φ

- Similarly, only backgrounded, GIVEN PVCs are allowed in SOV **Amharic** (Semitic; Kramer & Eilam 2012), **Quechuan** (Sánchez 2010), and **Siouan** (Gordon 2015) languages.

Cross-linguistic evidence: Amharic

Context: Who gave the book to my son yesterday?/Who did the teacher give the book to?

- (5) *astämariw tənant lä-lijē sət'e mätshafun.*
teacher.DEF yesterday to-my.son gave book.DEF.ACC
'The teacher gave the book to my son yesterday.'

(Wakweya Gobena, p.c.)

Context: What did the teacher give to my son yesterday?

- (6) **astämariw tənant lä-lijē sət'e mätshafun.*
teacher.DEF yesterday to-my.son gave book.DEF.ACC
'The teacher gave the book to my son yesterday.'

(Wakweya Gobena, p.c.)

Cross-linguistic evidence: Quechua

- In **Quechua**, SOV, there is **morphological marking** of IS-categories like topic and focus.
- Topics can appear as PVCs, (7a), but foci cannot, (7b).

- (7) a. *Mariya Xwanaman qun libruta -qa.*
Mariya Xwana.DAT give.3s book.ACC -TOP
'As for the book, Mariya gives to Xwana.' [sic]
(Sánchez 2010: 94)
- b. **Mariya Xwanaman qun libruta -n.*
Mariya Xwana.DAT give.3s book.ACC -FOC/EVID
int. 'It is the book that Mariya gives to Xwana'.
(Muysken 1995: 383, as cited in Sánchez 2010)

Cross-linguistic evidence: Siouan

- **Siouan**, SOV, allows for **deaccented, prosodically integrated, non-afterthought PVCs** that are “recoverable” (i.e., given).
- The example below is from Hidatsa (Siouan) (only corpus data; Gordon 2015).

Context: a war story with GIVEN *enemies*, and an utterance-final afterthought *the Snake people* (i.e., the Shoshone people).

- (8) *Hii šee awá ihtúutiru ú’šiak káawarec*
and that ground hill.base.at arrive.ss be.there.PL.NE
maaiháa’š Waapúkšaruxpáaka’š.
enemy.PL.DEF.the Snake.People.PL.DEF.the
‘And the enemy, the Shoshone/the Snake people, were on that
ground, having gotten to the base of the hill.’ (Gordon 2015: 400)

Cross-linguistic evidence: Surgut Khanty

- **Surgut Khanty** is also described as **rather strictly V-final** (Honti 1984; Nikolaeva 1999; Schmidt 2008), but it allows for certain types of **non-afterthought PVCs**.
- We demonstrate that their emergence also adheres to the GIVEN > NEW generalization.

Cross-linguistic evidence: Surgut Khanty

- In the texts collected in 1901, only GIVEN PVCs were allowed (9).

Context: At one place where he found the fox,

(9) *ʃəm utəm wärsəyən wöqinat. Töł mannəs*
good dream do.PST.3DU fox.COM/INSTR from.there go.PST.3SG
wöqi.

fox

‘They said good bye to each other with the fox where he found him. The fox went away.’ (Paasonen & Vértes 2001)

{ ... XP V }_ι (PVC)_φ

Cross-linguistic evidence: Surgut Khanty

- In today's Surgut Khanty, NEW PVCs are allowed too (10).

Context: I live with my father and my elder sister, and with my brother-in-law.

- (10) *tut owtinə os mən quŋjiwnə wałł ma mańitəm.*
that surface.LOC also 1PL next.to.us live.3SG 1SG brother.PL.1SG
'Besides, my younger brothers also live with us.' (Csepregi
1998: 56)

$\{\{ \dots \text{XP } \mathbf{V} \}_\iota (\mathbf{PVC})_\phi \}_\iota$

Cross-linguistic evidence: Udmurt

- **Udmurt** is described as **flexibly V-final** and is further along the GIVEN > NEW trajectory.
- Already in the texts collected in 1885 and 1891-92 (Munkácsi 1887; Wichmann 1901), both GIVEN (11), and NEW PVCs were allowed.

Cross-linguistic evidence: Udmurt

Context: A fox has stolen a chicken, a goose, a ram and an ox.

- (11) *vańze* *ik,* *kuregze* *no* *żażegze* *no*
all.3SG.ACC PTCL chicken.3SG.ACC also goose.3SG.ACC also
takaze *no* *ošse* *no* *vandem* *žícy.*
ram.3SG.ACC also ox.3SG.ACC also slaughter.EV.3SG fox
‘The fox slaughtered all of them: the hen, the goose, the ram
and the ox.’ (Wichmann 1901, text. 24)

{ ... XP V }_ℓ (PVC)_φ

Cross-linguistic evidence: Udmurt

- In these older sources, CONTRASTIVE FOCI, a subtype of NEW constituents, can appear as PVCs, (12).

Context: A man and a bear negotiated planting beets together.

- (12) *Mon bašto vyžyze, ton, gondyr, jylze.*
1SG take.FUT.1SG root.3SG.ACC 2SG bear top.3SG.ACC
'I'll take its root, you, bear, (take) its top.'
(Wichmann 1901, text. 23)

$\{\{ \dots \text{XP } \mathbf{V} \}_\iota (\mathbf{PVC})_\phi \}_\iota$

Cross-linguistic evidence: Udmurt

- Finally, constituents that are subparts of BROAD FOCUS are allowed as PVCs, (13), which means that Udmurt is the **furthest** along the GIVEN>NEW trajectory.

Context: A poor Udmurt went to hire himself out as a farmhand.

- (13) *śures vylyśen pumitam nulesmurtez.*
way from.on meet.EV.3SG forest.spirit.ACC
'On the way he met a forest spirit.' (Wichmann 1901, text. 30)

{ ... XP V PVC }_l

- In today's Udmurt, **all types of PVCs are allowed**, and the frequency of NEW PVCs is higher than in the older sources (Asztalos et al. 2017: 54–56).

- Numerous V-final languages that are undergoing a shift from more to less rigid V-finality follow the GIVEN > NEW, NARROW > BROAD, and CONTRASTIVE FOCUS > NEW-INFORMATION FOCUS trajectories in allowing for PVCs.
- We propose that the factor that governs this process is the **gradual loosening of restrictions on prosodically prominent post-V material**.

- We argue that this process involves a stepwise shift from prosodically non-integrated PVCs to the introduction of a **nested** ι -structure, and then to the **re-analysis** of the ‘major’ ι as the ‘core’ ι , thus getting rid of recursive ι s again.
- **Main takeaway:** Both in Uralic languages and beyond, the trajectory of diachronic syntactic changes, including the development from V-finality to more flexible word order, is shaped in crucial ways by principles of the prosodic interface.

Thank you for your attention!

This research has been supported by the research projects NKFIH K 135958 and NKFIH KKP 129921 of the National Research, Development and Innovation Office of Hungary.



Asztalos, E., K. Gugán & N. Mus. 2017. Uráli VX szórend. Nyenyec, hanti és udmurt mondat szerkezeti változatok [VX order in Uralic. Syntactic variability in Nenets, Khanty, and Udmurt]. In K. É. Kiss, A. Hegedűs & L. Pintér (eds.), *Nyelvelmélet és diakrónia* 3, 30–62. Budapest: Pázmány Péter Catholic University BTK.



Büiring, D. 2012. Predicate Integration – Phrase Structure or Argument Structure? In I. Kučerová & A. Neeleman (eds.), *Contrasts and Positions in Information Structure*, 27–47. Cambridge; New York: Cambridge University Press.



Csepregi, M. 1998. *Szurguti osztják chrestomathia*. Studia uralo-altaica.













Duanmu, S. 2004. Left-headed feet and phrasal stress in Chinese. *Cahiers de Linguistique Asie Orientale* 33(1). Publisher: Brill, 65–103.



É. Kiss, K. 2014. From Proto-Hungarian SOV to Old Hungarian Top Foc V X. *Diachronica* 30(2). 202–231.



Elfner, E. 2015. Recursion in prosodic phrasing: Evidence from Connemara Irish. *Natural Language & Linguistic Theory* 33(4). 1169–1208.

-  Féry, C. 2013. Focus as prosodic alignment. *Natural Language & Linguistic Theory* 31(3). Publisher: Springer, 683–734.
-  Göksel, A. 2013. Flexible word order and anchors of the clause. *SOAS Studies in Linguistics* 16. Publisher: Citeseer, 2–25.
-  Gordon, B. J. 2016. Information-structural variations in Siouan languages. In C. Rudin & B. J. Gordon (eds.), *Advances in the study of Siouan languages and linguistics*, 393–423. Berlin: Language Science Press.
-  Honti, L. 1984. *Chrestomathia ostiatica*. Budapest Tankönyvkiadó.
-  Ito, J. & A. Mester. 2013. Prosodic subcategories in Japanese. en. *Lingua* 124. 20–40.
-  Kramer, R. & A. Eilam. 2012. Verb-medial word orders in Amharic. *Journal of Afroasiatic Languages* 5. 75–104.
-  Nikolaeva, I. 1999a. *Ostyak*. München: LINCOM.
-  Nikolaeva, I. 1999b. *Ostyak*. Lincom Europa.
-  Nikolaeva, I. 2014. *A grammar of Tundra Nenets*. Berlin: de Gruyter.
-  Özge, U. & C. Bozsahin. 2010. Intonation in the grammar of Turkish. *Lingua* 120(1). 132–175.

-  Öztürk, B. 2013. Postverbal constituents in sov languages. *Theoretical approach to disharmonic word orders*. 270–305.
-  Paasonen, H. & E. Vértes. 2001. *H. paasonens surgutostjakische textsammlung am jugan*. Vol. 240. Suomalais-ugrilainen seura.
-  Sánchez, L. 2010. *The morphology and syntax of topic and focus. Minimalist inquiries in the Quechua periphery*. Amsterdam: John Benjamins.
-  Schmidt, Éva. 2008. Északi osztják nyelvtani jegyzet (serkáli nyelvjárás). *Fejes László szerk.: Serkáli osztják chrestomathia= Schmidt Éva Könyvtár* 3. 13–75.
-  Selkirk, E. 2011. The syntax-phonology interface. In J. A. Goldsmith, J. Riggle & A. C. L. Yu (eds.), *The Handbook of Phonological Theory*, vol. 2, 435–483. Hoboken, NJ: Wiley Blackwell.
-  Tereščenko, N. M. 1973. *Sintaksis samodijskih jazykov*. Leningrad: Nauka.
-  Truckenbrodt, H. 2006. Phrasal stress. In K. Brown (ed.), *Encyclopedia of Language and Linguistics*, 572–579. Amsterdam: Elsevier.

Appendix: Diachronic development of PVCs in Uralic languages

● Tundra Nenets

- ‘old’ texts (1911–1912): 24/1.079 (2,22%)
- ‘new’ written texts: 2/543 (0,37%)
- ‘new’ spoken texts: clause-external, afterthought-like prosodically separate PVCs that are associated with a clause-internal correlate

● Surgut Khanty

- ‘old’ texts (1901): 74/824 (9%)
 - ⇒ old/known information: 39 (c.93%)
 - ⇒ new information: 3 (c.7%)
- ‘new’ texts: 39/499 (7,82%)
 - ⇒ old/known information: 2 (c.20%)
 - ⇒ new information: 8 (c.80%)

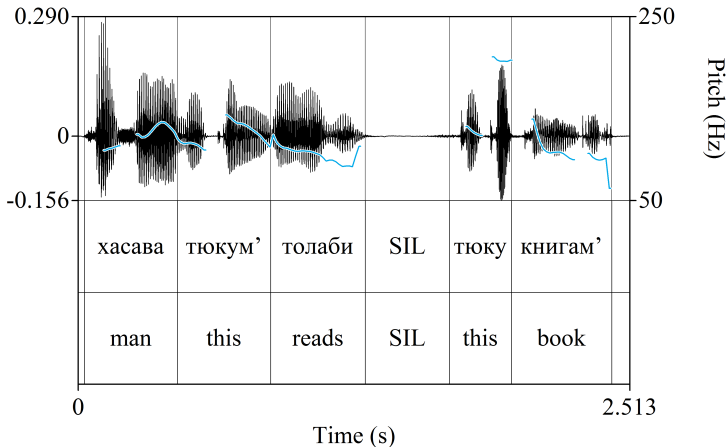
● Udmurt

- ‘old’ texts (1885, 1891–1892): 191/549 (35%)
 - ⇒ old/known information: c.67%
 - ⇒ new information: c.25%
- ‘new’ texts: 260/591 (44%)
 - ⇒ old/known information: c.36%
 - ⇒ new information: c.50%

Appendix: Prosody of AT in Tundra Nenets

- The post-verbal element is prosodically not integrated into the host sentence, but it involves special intonation: there is always a prosodic break, marked by a pause (represented by a comma or dash in the written sources), between the verb and the phrase that follows it, and the PVC is realized with a falling contour.
- The VX construction is not considered to be acceptable without this prosodic break.

Appendix: Prosody of AT in Tundra Nenets



Appendix: Multiple PVCs in Udmurt


	old texts	new texts
frequency (in non-verb-final clauses)	23.5%	10.5%
number of PVCs if > 1	2	mostly 2, sometimes 3
relative ordering	identical IS roles (GIVEN + GIVEN OR NEW + NEW)	neutral ordering
	different IS roles	<i>old-before-new</i> (14)

Context [text piece from a personal blog]: “Yesterday I told that the world is in chaos. And today I will tell...”

- (14) *mar malpa-lo so xaos śaryś šor jozo pokoleñi-yś adami-os*
what think-3PL that chaos about middle of.age generation-ELA man-PL
'(...) what the middle-aged generation thinks about that chaos.'

(MZ 30.11.2015)

Appendix: Russian contact

- **Extremist** view: word order changes are above all due to language contact (Smith 1981: 52, quoted by Harris and Campbell 1995: 137).
- A more **moderate** view: borrowing needs to be recognized as an independent mechanism of syntactic change (Harris and Campbell 1995: 150).
- Clausal constituent order is particularly prone to change in contact situations (Aikhenvald 2006).
- Diffusion of new patterns is facilitated by:
 - pragmatic salience (“The order of clausal constituents also typically correlates with discourse functions of arguments, and is highly diffusible” (p. 27));
 - tendency to achieve word-for-word and morpheme-per-morpheme intertranslatability. “Languages in contact, especially those with a high degree of bilingualism, will often come to have matching discourse patterns, and intonation unit contours” (p. 28).
- ‘Old’ Khanty texts: representing a period with presumably few Khanty-Russian bilingual speakers (“The informants’ command of Russian is truly deficient in places (for instance, the current Tremyugan informant of K[arjalainen] [...] speaks Russian downright frightfully, although he is the best speaker of Russian in this area, and around the Kazym [...] there isn’t a single Ostyak whom one could communicate in Russian with” (Paasonen’s letter to Otto Donner, quoted by Csepregi 2021)

- ‘New’ texts from an era when bilingualism is general. ⇒ Reasonable to assume that new word order patterns emerge / rarer patterns become more frequent owing to interference.