

OV and VO in Itelmen: A preliminary study

Jonathan David Bobaljik
Harvard Linguistics

7 Feb 2023
Lang Cog



Observations:

- Flexible word order: $OV \sim VO$

Q&A:

- Q: What conditions choice?
- A: Information Structure (given/new O)
- A': OV is basic, VO derived
(old information “moves”)

An alternative?

- Gibson et al. 2013: VO to disambiguate
GF of sole argument in “noisy” channel



Itelmen [itl]

- Chukotko-Kamchatkan, Kamchatka peninsula, Russia



By Marmelad - Made from Image:Map of Russian subjects, 2008-03-01.svg, CC BY-SA 2.5,
<https://commons.wikimedia.org/w/index.php?curid=3464501>

Itelmen [itl]

- Chukotko-Kamchatkan, Kamchatka peninsula, Russia
- 1990s ~ 80 speakers
- now handful of varying proficiency:
 - 1 completely fluid, native
- all speakers bilingual, Russian-dominant



Acknowledgments



T. N. Bragina



L. E. Pravdoshina

Collaborator:
Susi Wurmbrand

Stats help: Kate Davidson,
Steven Worthington
(Institute for Quantitative
Social Science, Harvard)



E. E. Silina



A. D. Ivashova

OV~VO:

(1) kma mił knin i? t'-il-ał-čen. OV
1SG all 2SG.POSS water 1SG-drink-FUT-1>3SG

'I will drink up all your water.' [AS (37)]

(2) k'-il-ʔin=(n)en mił kəłx^w-čax. VO
PRT-drink-TRANS.PRT=3CL all lake-DIM

'She drank up all the little lake.' [AS (38)]



- VO languages common focus-final (Jackendoff 1972 et seq)
- Russian (VO) (Neeleman & Titov, 2009; Bailyn, 2012; Titov, 2012):
 - Right Focus: new information focus occurs clause-finally (IK1)
 - *O_{old}* may but need not front (also contrastive focus)
- Old English (VO) (Struik & van Kemenade 2022)
 - “discourse-given, lexical objects are optionally OV, but that new objects are near-categorically VO”

Loosely: Given precedes new

- OV languages: post-verbal foci rare (Borise 2019)
- Turkish (OV) (Kornfilt, 2005; Şener, 2010):
 - preverbal focus, *O_{old}* scrambled (L or R)
 - postverbal constituents must be background (given, non-topic)

Generalization: “focus” in most deeply embedded position
old information can/must be in “derived” position

Schmerling (1976); Gussenhoven (1983); Selkirk (1984); Arregi (2016)

Turkish: SOV

- (3) I know that my wife asked Pelin, Pinar, and Can each to read a famous book by Orhan Pamuk titled “White Castle” in her class last week. I ask her today:

A: What happened with that reading assignment of yours from last week?

- (4) a. # Yalnızca Pelin **kitab-1** **oku-muş**. SOV
only Pelin book-ACC read-PAST
- b. **Kitab-1** yalnızca Pelin **oku-muş**. OSV
book-ACC only Pelin read-PAST
- c. Yalnızca Pelin **oku-muş** **kitab-1**. SVO
only Pelin read-PAST book-ACC

‘Only Pelin read the book.’

OV: [new - V] - background

VO: given - [V - new]

- O in its canonical position (complement of V) as most deeply embedded element:
- convergence of sentential “nuclear” stress and focus prominence Jackendoff (1972); Arregi (2016)
- *O_{old}* is in a derived / high position (whether L or R)

Text Study

Title	Stroyteller	Genre
1. Wingless Gosling	Tatiana N. Bragina	Myth/Tale
2. Old Man	Lyudmila E. Pravdoshina	Event
3. Kutxh and the Mice	Ekaterina E. Silina	Myth/Tale
4. Tilval	Agrafena D. Ivashova	Legend

- ~900 clauses
- Sedanka-Tigil (Northern) dialect
- recorded 1994, transcr. & edited with speakers
- Speaker 1 *1906, Speakers 2-4 *1934-1942.

Coding (by hand)

Constituent Order: relative order of S, V, O
Information Structure status of O
Animacy of O

Table 1: Constituent Order (First Pass)

Intransitive	V	SV	VS	Total
	289	234	34	557
Transitive				308

Four Texts

Table 2: Transitive Constituent Order (raw, $n=308$)

O_{\emptyset}	V	SV	VS		
	55	15	6		76
S_{\emptyset}		OV	VO		
		102	60		162
S, O	SOV (33) OSV (6) OVS (4)	43	27	SVO (21) VSO (5) VOS (1)	70
Totals		145	87		232

S&O overt $\sim 25\%$ S omitted $\sim 60\%$ O omitted $\sim 25\%$

OV:VO = 1.7:1

Coding for IS

- O_{new} referent not previously introduced in discourse
- O_{contr} explicitly contrastive ($n=6$), merged with O_{new} cf. *focus*
- O_{old} referent previously mentioned in discourse (=given)
1st and 2nd person always given
- dialogues within a story are new discourses
 - Could drill down much deeper:
 - finer-grained: salience, accessibility, etc
 - part-whole: X's hand, door of house, etc.
 - common ground via world knowledge (chapel in village, etc)

First results: OV/VO × old/new

Table 3: Aggregated (raw)
($\chi^2=18.54$, $p=.000017$)

	OV	VO	Total
O_{new}	89	28	117
O_{old}	56	59	115
Totals	145	87	232

Table 4: Text 1 (raw) (Fisher's Exact
 $p=0.0092$)

	OV	VO	Total
O_{new}	12	4	16
O_{old}	3	10	13
Totals	15	14	29

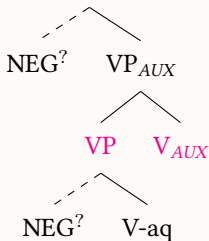
- O_{new} pre-verbal 3:1
- O_{old} evenly split
prefers post-verbal in older text

Itelmen Evidence for head-final V/vP (I)

Negation: NEG ... V-(k)aq ... AUX

- (5) qaʔm ʎəm-aq t'-iʎ-čen
 NEG kill-NEG 1SG.-AUX-1>3SG
 'I didn't kill it.' [S3:9]

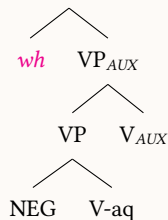
- (6) a. * t'-iʎ-čen qaʔm ʎəm-aq
 1SG-AUX-1>3SG NEG kill-NEG
- b. * qaʔm t'-iʎ-čen ʎəm-aq
 NEG 1SG-AUX-1>3SG kill-NEG
 'I didn't kill it.' [S3:9]



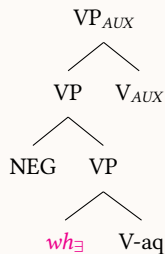
see also Abramovitz (2019) on Koryak

(aside) *wh-in-situ*, but ...

- (7) a. **k'e** qaʔm k'oɬ-kaq ɬ-in
 who NEG come-NEG AUX-3SG
 Who didn't come? [S3:23]



- b. qaʔm **k'e** k'oɬ-kaq ɬ-in
 NEG who come-NEG AUX-3SG
 Noone came. [S3:23]



Itelmen Evidence for head-final V/vP (II)

Aux as Light-verbs:

root[?] + ... AUX

- (8) a. **xaq** tχi-ʔn k-**təl**-knen
know 3PL-PL PRT-AUX-PRT
'They recognized (lit: knew) him.' [TN (63)]
- b. jurte-čχ **ənluʃi-ʃ** k'-**ite**-ʔin
yurt-DIM enflame-INST PRT-AUX-PRT
'(They) burned the little yurt.' [KL (118)']
- c. noz-əʔn knank **napravit** t'-**iʃ**-če-ʔn
dried.fish-PL 2SG.DAT prepare(<Russ.) 1SG-AUX-1>3-PL
'I've prepared jukola (dried fish) for you.' [KL:15]

Two analytical ambiguities

- Perception verbs: VO or V-CP ?
- N-construction: are these transitive?

(9) k'-əłčku-in=(n)in tsxal-astas k-k'ot-knen
 PRT-see-TRNS.PRT=3CL fox-AUGM PRT-come-PRT

'He saw a fox coming.' [AS (28)]

- a. [saw fox_i] [(it_i) came] NP-complement
 b. [saw [fox come]] clausal-complement*

Perception (and similar) verbs:

- (9) k'-əłčku-in=(n)in tsxal-astas k-k'ot-knen
 PRT-see-TRNS.PRT=3CL fox-AUGM PRT-come-PRT

'He saw a fox coming.' [AS (28)]

- a. [saw fox_i] [(it_i) came] NP-complement
 b. [saw [fox come]] clausal-complement*

Strategy: raw pass — count everything
 refined pass — only unambiguous O

*clausal complements always? post-verbal

Perception (and similar) verbs:

(9) k'-əłčku-in=(n)in tsxal-astas k-k'ot-knen
 PRT-see-TRNS.PRT=3CL fox-AUGM PRT-come-PRT
 'He saw a fox coming.' [AS (28)]

- a. [saw fox_i] [(it_i) came] NP-complement
 b. [saw [fox come]] clausal-complement*

- All perception verb complements post-verbal, regardless of IS
- 3/4 VO_{new} in Text 1 (15/28 in aggregate)

Interim Summary: Text 1

Table 5: Text 1 (raw) (Fisher's Exact $p= 0.001$)

	OV	VO	Total
O_{new}	12	1	13
O_{old}	3	10	13
Totals	15	11	26

- O is nominal objects only, excluding perception/restructuring
 - clausal complements always? post-verbal

N-construction: passive/impersonal

- (10) a. **Mit-enk** n-qzu-z-um
Miti-LOC N-wait-PRES-1S.OBJ
'Miti (is) waiting for me.' [KL (17)']
- b. **tsxal-enk** mił i? n-γil-čen
fox-LOC all water N-drink-3SG.OBJ
'All the water was drunk by the fox.' [AS (47)']

S: oblique (LOC/INSTR), fixed prefix *n-* [=3PL]

O: unmarked, triggers object agreement

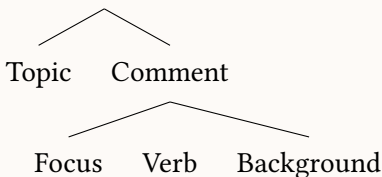
Source of Ergativity in Chukotkan? (Bobaljik, 2019)

Passive/Impersonal

- (11) a. **kma** n-oms-qzu-βum
1SG N-leave-ASP-3>1SG
'I was left behind.' [AS (46)]
- b. sinanɛβt-enk **kma** n-txunt-qzu-βum anɫe
Sinangewt-LOC 1SG N-raise-ASP-1SG.OBJ winter
'I was raised by Sinangewt in the winter.' [AS (85)]

O_{old} preverbal,
but O is **topic** in many such examples.
n=6 (Text1), n \geq 27 (aggregate)

(12)

 $\text{Given}_{\text{Top}} > \text{New} > \text{Given}_{\text{Bkgd}}$

- 2 NPs: S = Topic, O can be O_{new} or O_{old}
- 1 NP: O follows pattern for Topic, not $O_{\text{background}}$

Georg & Volodin (1999) N-construction as “Agensausblendung”
(backgrounding of Agent)

Table 6: Text 1 (final) (Fisher's Exact $p=0.0003$)

	OV	VO	Total
O_{new}	10	1	11
O_{old}	1	10	11
Totals	11	11	22

Table 7: Aggregated OV/VO \times old/new (fin) ($\chi^2=39.9008$, $p<.00001$)

	OV	VO	Total
O_{new}	85	13	98
O_{old}	45	57	102
Totals	130	70	200

Aside: Interacting factors

VO_{new} in Text 1:

- (13) a. k-zun^ɬ-qzu-kne-ʔn **qsə-ʔn**
PRT-live-ASP-PRT-PL goose-PL
‘There lived some geese.’ [AS (1)]
- b. **qsə-ʔn** k'-ənsxt-ʔeʔn **tɣi-ʔn p'e-ʔn**
goose-PL PRT-give.birth-TRNS.PRT.PL 3PL-PL child-PL
‘The geese, they gave birth to children.’ [AS (2)]

Narrative Inversion

(Once upon a time, ЖИЛ БЫЛ)

- (14) it'e k-zun^{jɬ}-qu-knen li staroj tnaqol.
when PRT-live-ASP-PRT very old old.man
'Once there lived a very old man.' [TN (2)]
- (15) it-qat k-zun^{jɬ}-qu-kne-ʔn Kusxɬnequ i Miti,
when-already PRT-live-ASP-PRT-PL Kutkh and Miti
'Once there lived Kutkh and Miti,' [KL (1)]
- (16) it-qata zin-k k-zun^{jɬ}-qzu-kne-ʔn č'amza-ʔl.
when-already woods-LOC PRT-live-ASP-PRT-PL person-PL
'Once in the woods there lived some people.' [TL]

Narrative Inversion

k'-ənsxt-ʔeʔn

tʃi-ʔn p'e-ʔn

PRT-give.birth-TRNS.PRT.PL 3PL-PL child-PL

'They gave birth to children.' [AS (2')]

SUBJECT

OBJECT_{NEW}

VERB

SUBJECT

VERB

OBJECT_{OLD}

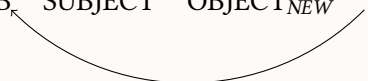
N.I.

VERB

SUBJECT

OBJECT_{NEW}

OBJECT_{OLD}



A “Noisy Channel” Alternative?

Gibson et al. (2013):

dog chased

Assumptions:

- No case marking
semantically reversible (animate), 1 NP “missing”
- S-initial (Universal strong trend); (S > O)
- SOV, SVO, VSO

Proposal use of SVO order disambiguates

Prediction preference for VO order for **animate O**

(NB: Gibson et al. (2013) consider only gesture, no actual language data)

A Noisy Channel?

Itelmen:

- No case marking S/O
- usually only one overt NP
- OV, VO available

V	55	
V,NP	183	(S: 21, O:162)
V,NP,NP	70	
<hr/>		
Total	308	(raw)

A Noisy Channel?

- binary logistic regression model

glmer(*ORDER* ~ *INFSTR* + *ANIMACY* + (1/*TEXT*))

(results same if *TEXT* treated as fixed)

- Both factors significant, but
- holding other predictor variables constant:
 - odds ratio $VO_{anim} : VO_{inan} = 2.2$ (95% CI [1.1, 4.8]).
 - odds ratio $VO_{old} : VO_{new} = 7.7$ (95% CI [3.7, 16.9]).

A Noisy Channel?

- Often just one NP (S: 21, O:162).
- NP_{new} is never transitive S, always O
- If there is an issue of disambiguation, it's with O_{old}
- Interaction of animacy and IS?
*glmer(ORDER ~ INFSTR * ANIMACY + (1|TEXT))*
- none found: only INFSTR as a significant factor ($p < .001$)

A different animacy factor?

Small effect of animacy independent of IS, but $O_{inan}V$ or VO_{anim} ?

- (17) esx-ank-əŋ li plex-aʔn txuʔ-iʔn_{new} q-la-qzu-čx-eʔn
 father-DAT-DAT very big-PL greeting-PL IMP-tell-ASP-II-PL
 ‘Give (tell) your father big greetings!’ [TN (110)]
- (18) muza-ʔn miʔ esx-aʔn-k-əŋ txuʔ-iʔn_{old} ntʼ-la-aʔ-xŋ-eʔn
 1PL-PL all father-PL-DT-DT greeting-PL 1PL-tell-FUT-II-PL
 ‘We will all give (tell) our fathers greetings!’ [TN (112)]

Non-referential O? (Pseudo-incorporation?)

Default OV, typically inanimate?

Noisy Channel?

- Main driver of $OV \sim VO$ is IS, not animacy
- small effect of animacy, various potential sources
- since $S \rightleftharpoons$ topic, IS alone enough to resolve GF
- missing/dropped NP most likely topic
- S can be post-verbal
- Noisy channel effect doesn't seem to contribute to explanation of alternation

Conclusions

Questions:

- Descriptive: what conditions $OV \sim VO$?
 - IS: $O_{New} V O_{Old}$
 - obscured by: perception complements, N-construction
- Analytical: Mixed word order?
 - Itelmen is an OV language, like other OV langs
 - optional extraposition of O_{old}
- Theoretical: Role of Communicative Efficiency?
 - Recognizing role of IS and context reduces the putative ambiguity

FOFC and extraposition

Biberauer et al. (2014), citing Koptjevskaja-Tamm:
Even in OV languages, obligatory extraposition of C-initial CP:

- (19) a. * ...dass sie [_{CP} dass es regnet] behaupten.
...that they that it rains claim
'...that they claim that it is raining.'
- b. ...dass sie behaupten, [_{CP} dass es regnet].
...that they claim that it rains
'...that they claim that it is raining.'

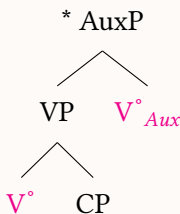
German: V-final in embedded clauses

An indirect argument for high VO_{Old}

Biberauer et al.: Extraposition (CP is not in VP)

Headedness parameters (hypothetical):

VP-AUX	head-final AuxP
DP-V	head-final VP
V-CP	but CP complement to right



Predicts [[V O_{CP}]_{VP} AUX]

False! Unattested.

∴ [V CP] order involves extraposition
not complement on right

An indirect argument for high VO_{Old}

Biberauer et al.: Extraposition (CP is not in VP)

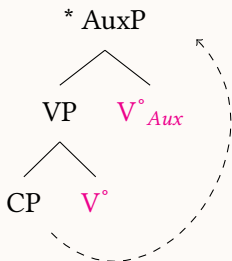
Headedness parameters (hypothetical):

VP-AUX	head-final AuxP
XP-V	head-final VP
V-CP	but CP complement to right

Predicts [[V O_{CP}]_{VP} AUX]

False! Unattested.

∴ [V CP] order involves extraposition
not complement on right



An indirect argument for high VO_{Old}

Elicitation:

(20) βeqaʔnʲɩ ‘bear’:

<1> qaʔm <2> ɬəm-aq <3> tʻ-iɬ-čen <4>
 NEG kill-NEG 1SG.-AUX-1>3SG

‘I didn’t kill (the/a) bear.’ [S3:9-10]

<2> preferred

<1,4> acceptable

<3> very hesitantly acquiesced but in ‘repeating’ switched to <2>

An indirect argument for high VO_{Old}

Texts:


- one occurrence of [V O Aux] in recordings (Lt. Vb)
- but auxiliary omitted in edited version V O

Tentative conclusion: * V O Aux

Unexpected, if O in [VO] within VP

prima facie case that VO is (high) extraposition of O.

- Itelmen is an OV language, VO is derived

- Abramovitz, Rafael (2019) Word Order and Negation in Koryak, handout of a paper presented at MIT.
- Arregi, Karlos (2016) Focus projection theories. In *The Oxford Handbook of Information Structure*, Caroline Féry & Shinichiro Ishihara, eds., Oxford: Oxford University Press, 185–202.
- Bailyn, John Frederick (2012) *The Syntax of Russian*. Cambridge: Cambridge University Press.
- Biberauer, Theresa, Anders Holmberg, & Ian Roberts (2014) A syntactic universal and its consequences. *Linguistic Inquiry* 45(2): 169–225.
- Bobaljik, Jonathan David (2019) The Chukotkan “inverse” from an Itelmen perspective. In *Лингвистика и другие важные вещи: Памяти Александра Павловича Володина [Linguistics and other important things]*, E. V. Golovko, A. M. Pevnov, M. Yu. Pupynina, A. A. Syuryun, & A. Yu. Urmanchieva, eds., Sankt-Peterburg: Institute of Linguistics, Russian Academy of Sciences, 415–435.
- Georg, Stefan & Aleksandr P. Volodin (1999) *Die itelmenische Sprache*. Wiesbaden: Harrassowitz Verlag.
- Gibson, Edward, Steven T. Piantadosi, Kimberly Brink, Leon Bergen, Eunice Lim, & Rebeccas Saxe (2013) A noisy-channel account of crosslinguistic word order variation. *Psychological Science* 24(7): 1079–1088. 

- Gussenhoven, Carlos (1983) Focus, Mode and the Nucleus. *Journal of Linguistics* 19(2): 377–417.
- Jackendoff, Ray (1972) *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Kornfilt, Jaklin (2005) Asymmetries between pre-verbal and post-verbal scrambling in Turkish. In *The free word order phenomenon : its syntactic sources and diversity*, Joachim Sabel & Mamoru Saito, eds., Berlin: Mouton de Gruyter, 163–181.
- Kural, Murat (1997) Postverbal constituents in Turkish and the Linear Correspondence Axiom. *Linguistic Inquiry* 28: 498–519.
- Neeleman, Ad & Elena Titov (2009) Focus, Contrasts, and Stress in Russian. *Linguistic Inquiry* 40(3): 514–524.
- Schmerling, Susan F. (1976) *Aspects of English Sentence Stress*. Austin: University of Texas Press.
- Selkirk, Elisabeth (1984) *Phonology and Syntax*. Cambridge, MA: MIT Press.
- Şener, Serkan (2010) *(Non-)Peripheral Matters in Turkish Syntax*. Ph.D. thesis, University of Connecticut, Storrs, CT.
- Titov, Elena (2012) *Information Structure of Argument Order Alternations*. Phd dissertation, University College London, London. 