

Atelier international / International Workshop

## OV-IS 2019

Corrélats de l'ordre OV et structure de l'information

OV basic word order correlates and information structure

Mercredi 20 et jeudi 21 novembre 2019

Salle 4.24

### Organisation

SeDyL (Structure et Dynamique des Langues, UMR 8202 Inalco-CNRS-IRD)  
MII (Mondes Iranien et Indien, FRE 2018-Sorbonne Nouvelle-EPHE-Inalco-CNRS) Labex EFL (Empirical Foundations of Linguistics, USPC)

### Contact

Anaïd Donabedian (Inalco-SeDyL) [adonabedian@inalco.fr](mailto:adonabedian@inalco.fr)  
Pollet Samvelian (P3-MII) [pollet.samvelian@sorbonne-nouvelle.fr](mailto:pollet.samvelian@sorbonne-nouvelle.fr)  
Claudine Chamoreau (CNRS-CEMCA/SeDyL) [claudine.chamoreau@cnrs.fr](mailto:claudine.chamoreau@cnrs.fr)

**Inalco**  
65 rue des  
Grands Moulins  
75013 Paris

# Program/programme

Wednesday, November 20<sup>th</sup> / Mercredi 20 novembre - Salle 4.24

**14h30-15h30**

**Judith Aissen**

*Word order and information structure in a so-called 'VOS' language*

**15h30-16h15**

**Tatiana Nikitina**

*Argument linearization and deviations from canonical OV word order: A contrastive analysis of Bashkir and Wan*

**16h15-16h30**

Coffee break / Pause café

**16h30-17h30**

**Elisabeth Verhoeven**

*Word order and optional ergative marking in Cabécar (Chibcha)*

**17h30-18h15**

**Roberto Herrera Miranda**

*The morphosyntactic realization of foci in Maleku*

**19h30**

Dinner / Diner

# Program/programme

Thursday, November 21<sup>th</sup> / Jeudi 21 novembre - Salle 4.24

**10h00-11h00**

**Geoffrey Haig**

*Between OV and VO: exploring word-order variation and change across the Western Asian Transition Zone (WATZ)*

**11h00-11h45**

**Stavros Skopeteas**

*From OV to VO and information structure: the syntax of Caucasian Urum*

**11h45-12h30**

**Pegah Faghiri, Victoria Khurshudian and Pollet Samvelian**

*Modern Eastern Armenian: OV or VO?*

**12h30 – 14h30**

Lunch / Déjeuner

**14h30-15h15**

**Metin Bağrıaçık**

*Word order in Phrasiot Greek: What is from Turkish and what is not?*

**15h15-16h00**

**Masud Mohamadirad**

*Post-predicate oblique arguments across Kurdish*

**16h00-16h30**

Closing coffee break

# Résumés des présentations/ Abstracts

Judith Aissen (University of California, Santa Cruz)

## *Word order and information structure in a so-called 'VOS' language*

This talk considers aspects of the word order of Tsotsil, a Mayan language spoken in Chiapas, Mexico. The first part of the presentation will consider the functions of preverbal positions, as these are linked to the central relations of information structure: Topic, Focus, and Given. The second part of the presentation focuses on the canonical word order of the clause, VOS. I will argue that this order is determined not by grammatical function (contrary to what is suggested by the label 'VOS'), but by definiteness: in postverbal position, definites follow non-definites. This combined with the fact that transitive S is almost always definite, while O need not be, yields a canonical order which conforms to the description 'VOS'. A structural analysis of Tsotsil word order will be suggested which provides a position for definite noun phrases on the right edge of the clause.

Metin Bağrıaçık (Ghent University)

## *Word order in Phasiot Greek: What is from Turkish and what is not?*

Phasiot Greek (PhG) is an Asia Minor Greek dialect that was spoken before 1923 by 2600 people in six villages in what is today the Kayseri province in Turkey. Between 1923 and 1925 all its speakers were relocated in Greece following the population exchange between Greece and Turkey. It is currently a moribund variety, spoken as a heritage variety by about 25 second-generation refugees in northern Greece.

Since the early 20th century, numerous scholars have argued for massive Turkish influence on PhG at various linguistic levels. On one occasion it was claimed for PhG that "the peculiar Turkish order of words has invaded Greek" (Dawkins 1916:198)—a claim adopted in various subsequent studies as well. In this study, we critically assess this claim focusing on word order (variation) at both clausal and nominal domains.

PhG exhibits all possible word order permutations in declarative main clauses with a mono-transitive verb and an overt subject and a direct object. Drawing on (i) corpus of written texts since late 19th century, (ii) corpus of naturally occurring spoken data collected between 2013-2019 and (iii) speaker judgments, we show that PhG can at best be characterized as a language with VS0/SV0 pragmatically neutral word order, which are however not perfectly interchangeable. We further show that all O-initial word orders, SOV order and a subset of SVO orders are pragmatically marked: they contain one or more constituents whose referents receive topic or focus interpretation. This refutes any claim that there is pervasive influence of Turkish, a neutral-SOV language, on PhG in terms of word order at clausal level.

In the nominal domain the situation is different. We show that modifiers of a noun in PhG (along with another Asia Minor Greek dialect—Cappadocian) come in a strict order, verifying the cross-linguistic observation that word order in the noun phrase tends to be less flexible than word order in the clause. Unlike Modern Greek, or various other non-Asia Minor Modern Greek dialects, all modifiers of a noun, i.e., (universal) quantifiers, demonstratives, restrictive relative clauses, numerals and adjectives are strictly prenominal and come in this strict order (although postnominal restrictives relative clauses have emerged only recently, under the influence of Modern Greek). Unlike the case in clausal order, this peculiar ordering constraint is difficult to explain without evoking an analysis that takes into consideration the influence of Turkish on PhG.

# Résumés des présentations/ Abstracts

Pegah Faghiri, Victoria Khurshudian and Pollet Samvelian  
(Universität Köln, Inalco/SeDyL/Sorbonne-Nouvelle/Monde Iranien)  
*Modern Eastern Armenian: OV or VO?*

The “unmarked” or “canonical” word order in Modern Eastern Armenian (MEA) is controversial. While many typological studies group Armenian with OV or (OV flexible) languages (Minassian, 1980: 263; Kozintseva, 1995:8; Siewierska 1998, Dum-Tragut, 2002, among others), Armenian grammars (Abrahamyan et al., 1975, Abrahamyan, 2004) and some other typological studies (ex. Dum-Tragut, 2009, Dryer 2013, WALS) characterize MEA as either VO or underspecified between OV or VO. Studies identifying OV as the basic word order account for VO ordering in terms of information structure and analyze postverbal constituents as “backgrounded” or “dislocated” (Giorgi, and Haroutyunian, 2016; Tamrazian, 1991, 1994 ). This analysis mainly relies on the fact that the argument/narrow focus is always preverbal in MEA and has led to the assumption that the postverbal domain can only host “given” or “old” information. From a quantitative point of view, it is not self evident to disqualify VO as an unmarked word order in MEA and our preliminary corpus studies do not support this view. Firstly, word order distribution in the UD corpus (570 occurrences, <https://universaldependencies.org/>) is 46% of VO vs 39% of OV. Secondly, in an unbiased sample of 338 instances of periphrastic verbal forms (involving the auxiliary em ‘to be’) and 231 instances of simple verbs, extracted from the EANC open access corpus (<http://eanc.net/>), we observe that VO is by far the most frequent order (78.8%). Importantly, when we study these occurrences more closely, we see that the VO order maps into an “all focus” discourse configuration in a fair share of cases. The goal of our study is to clarify the relation between the tendencies highlighted via corpus studies and the distribution of word order in controlled (experimental) production of speakers that aims at eliciting neutral “all focus” sentences. If VO is the most frequent way of realizing all focus sentences, it can be argued that VO is (pragmatically) the unmarked or canonical word order in MEA (cf. Dryer, 2007). On the other hand, from a distributional point of view (cf. Dryer, 2007), knowing that VO has a more restrictive distribution (indeed OV has more various discourse uses than VO), it can be argued that OV is the unmarked word order.

## Résumés des présentations/ Abstracts

Geoffrey Haig (Bamberg University)

*Between OV and VO: exploring word-order variation and change across the Western Asian Transition Zone (WATZ)*

In this talk I present on-going research on the dynamics of word-order change in the Western Asian Transition Zone (WATZ), an area roughly corresponding to contemporary eastern Turkey, Northern Iraq, Western Iran, Armenia and Azerbaijan. The term 'transition zone' (Haig 2017, Haig & Khan 2018, see also Stilo 2015) refers to the fact that these languages occur at the region of overlap between a central Asian macro-region dominated by OV languages (Turkic, Iranian, Indo-Aryan, Mongolic, Tungusic, Uralic), and a Western European/North-African macro-region, dominated by VO languages (Celtic, Romance, arguably Germanic, Afro-Asiatic). Within WATZ, attested shifts of VO to OV include the comparatively well-documented Western Armenian (Donabedian 2018), but also the Jewish Neo-Aramaic dialect of Urmi (Khan 2008). Historically OV languages switching wholesale to VO have yet to be documented in WATZ (Haig & Khan 2018), but the OV languages of the region nevertheless exhibit several features typical for VO languages, in particular high frequencies of post-verbal obliques (Haig 2014, Stilo 2018, Stilo & Haig in prep.), which are absent in their genetic relatives outside the region. I will utilize a corpus-based approach (Adamou 2016, Schnell et al, under review) to compare data from a range of genetically diverse languages in WATZ, which shed light on one of the most thoroughly researched, yet still poorly-understood, typological shifts. I will also advance some remarks on the importance of transition zones as havens for typologically unusual constellations.

# Résumés des présentations/ Abstracts

Roberto Herrera Miranda (INALCO/SeDyL - Universität Göttingen)

## *The morphosyntactic realization of foci in Maleku*

Maleku, also known as Guatuso (ISO 639-3: gut), is an ergative head-marking Chibchan language which exhibits most of the traits typical of SOV languages with respect to word order, such as postpositions, possessor-possesum and main verb-auxiliary order, etc. (cf. Dryer 2007). Perhaps unusual for a verb-final language is its tendency to realize absolutes postverbally:

(1) Taca na=rrri=quiye ni=ja=to ni Ø=tarinh coja  
CONJ 1ABS=3ERG=say 3SG=LOC=FUT 3SG 3ABS=disappear fever  
'And he told me: the fever will disappear with this.'

(2) Jue i=punhe tan ninhata lecu marama=ti ni Ø=cuye piuju  
MOD 1ABS=EXIST but now person PL=AF 3SG 3ABS=take balsa\_wood  
'There is (royal cedar) but now people use balsa wood.'

This is also characteristic of obliques and adverbs, but not of ergatives. This talk explores the non-canonical realization of agents and patients, specially in focus environments, based on data extracted from a corpus of spontaneous speech. Additional data gained using the fieldwork tool QUIS (Skopeteas et al., 2006) is also discussed.

While patient focus contexts exhibit a non-canonical word order, such as that induced by the questions in (3-4), agents are morphologically marked and remain in-situ:

(3) I=cuquicha ni Ø=rrri=joye unhe quirri  
3ABS=hand 3SG 3ABS=3ERG=make PROG light  
'He (his hand) is hitting the CANDLE.' (answer to a patient correction question)

(4) Coracuinh lacachi i=rrri=alu~aluye to orachuma-o  
branch one 3ABS=3ERG=drag~RED go child-DIM  
'A BRANCH a child is dragging along.' (answer to a patient completion question)

In transitive clauses, the subject often exhibits dependent marking, in a construction which precludes ergative verbal agreement (see Constenla 1998). This agent-marking construction is triggered by different factors, including focusing of the agent (5), but also changes in word order (6), among others (Herrera 2019). This is reminiscent of the conditions under which agents are flagged in other Chibchan languages (cf. Quesada & Skopeteas 2010 for Teribe, Bajorat 2014 for Ika):

(5) Curijuri=ti arapcha ora ye unhe i=cuquicha=c  
woman=AF child DIM lift PROG 3POSS=hand=LOC  
'A WOMAN.' (answer to an agent-completion question)

(6) Taca epeme ni lica tui=ti Ø=chie  
CONJ NEG DET liquid 1PL=AF 3ABS=drink  
'And the (plant's) SAP we do not drink.' (only the flower)

### References

- Bajorat, J. 2014. Ergativity in ika. function and use of the marker -se'. Humboldt-Universität zu Berlin. MA thesis.  
Constenla, Adolfo. 1998. Gramática de lengua guatusa. Heredia, Costa Rica: EUNA 1st edn.  
Dryer, Matthew S. 2007. Word order. In Timothy Shopen (ed.), Clause structure, vol. 1 Language Typology and Syntactic Description, 61-131. Cambridge University Press 2nd edn.  
Herrera, Roberto. 2019. Agent focus in maleku. Presented at the ACALING Fifth International Congress. UNA: Heredia, Costa Rica  
Quesada, J Diego & Stavros Skopeteas. 2010. The discourse function of inverse voice: An experimental study in teribe (chibchan). Journal of Pragmatics 42 (9). 2579-2600.  
Skopeteas, S, I Fiedler & et al. Hellmuth, S. 2006. Questionnaire on information structure: Reference manual Vol 4. Interdisciplinary Studies on Information Structure. Potsdam: University of Potsdam.



# Résumés des présentations/ Abstracts

Tatiana Nikitina (CNRS/LLACAN)

## *Argument linearization and deviations from canonical OV word order: A contrastive analysis of Bashkir and Wan*

This study is an attempt to compare the interaction between word order and information structure in two genetically and culturally unrelated SOV languages: Bashkir (Turkic, Russia), and Wan (Mande, Côte d'Ivoire). I focus on two aspects of cross-linguistic variation: factors responsible for non-canonical post-verbal placement of objects, and the interaction of arguments in pre-verbal positions. I show that while Bashkir and Wan behave in strikingly different ways with respect to the former, they display some non-trivial similarities with respect to the latter.

Bashkir is a typical verb-final language with a pre-verbal contrastive focus position (cf. Johanson 1998 on Turkic). In spontaneous data, however, arguments often follow the verb when associated with certain discourse-structuring functions. I present evidence for two distinct post-verbal positions in which arguments can appear, corresponding to different sets of functions. Afterthought-like constituents are commonly used to provide additional information on arguments introduced in their canonical pre-verbal positions; they are separated from the verb by an intonational break, suggesting very high attachment in the clause structure. A different position is used for constituents introducing new arguments; this position is integrated intonationally with the verb. Bashkir also shows evidence for the fronting of topic constituents, as well as for a clause-initial predicate focus position. Overall, word order in Bashkir is to a large extent influenced by information structure, in ways that suggest Russian influence.

Wan is an OVX language which places objects before the verbs but adjoins all oblique arguments and adjuncts in an unusually high, clause-level position (Nikitina 2009, 2019). Unlike in Bashkir, its word order is extremely rigid: there is no focus position, and all sorts of topicalized and afterthought-like arguments must be referred to by a pronominal element in the canonical position. A major prosodic break separates the verb from anything that follows. Overall, word order in Wan shows no effects of information structure, and neither is information structure reflected in prosody.

Despite being clearly OV languages, Bashkir and Wan are strikingly different with respect to their word order, and in particular with respect to the extent of interaction between word order (as well as prosody) and information structure. Yet they show subtle and non-trivial similarities in linearization effects in pre-verbal positions.

In Bashkir, the ordering of the subject and the object is sensitive to definiteness: definite objects are commonly placed before the subject, while indefinite objects tend to follow it. This is consistent with the overall prominent role of information structure in Bashkir constituent structure. In Wan, definiteness has no effect on word order, but one important aspect of clausal structure is sensitive to the relative definiteness of the subject and object. Subjects and objects are separated in Mande by an auxiliary position, which may be filled by TAM markers. Strikingly, this position is sometimes filled in wan by a dummy auxiliary that carries no TAM meaning. The dummy auxiliary – also known as the bidirectional case marker (Heath 2007) – shows the tendency to appear in cases of subject-object prominence reversals, e.g. when the object outranks the subject in definiteness or information status (Aissen 1999, 2003). It is used, effectively, to avoid SOV structures with linearly adjacent “marked” combinations of subject and object. I draw parallels between this phenomenon and the tendency to front definite objects in Bashkir, suggesting that both derive from the same kind of interaction between word order and argument properties, manifested in SOV languages and structures.

In sum, the contrastive analysis of Bashkir and Wan reveals striking differences in some and subtle similarities in other aspects related to word order. The languages are completely different with respect to factors driving deviations from the canonical OV word order, in particular with respect to the licensing of post-verbal arguments. On the other hand, they both show effects of argument prominence on argument linearization in pre-verbal positions. I interpret this result as indicative of universal constraints on the encoding of different argument combinations within the same clause.

### References

- Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *NLLT* 17: 673–711.  
Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *NLLT* 21: 435–483.  
Heath, Jeffrey. 2007. Bidirectional case-marking and linear adjacency. *NLLT* 25: 83–101.  
Johanson, Lars. 1998. The structure of Turkic. L. Johanson & É. Á. Csató (eds.). *The Turkic Languages*. London: Routledge, 30–66.  
Nikitina, T. 2009. The syntax of PPs in wan, an “SOVX” language. *Studies in Language* 33(4): 907–30.  
Nikitina, Tatiana. 2019. Verb phrase external arguments in Mande: New evidence for obligatory extraposition. *NLLT* 37(2): 693–734

# Résumés des présentations/ Abstracts

Stavros Skopeteas (University of Göttingen)

*From OV to VO and information structure: the syntax of Caucasian Urum*

A change from OV to VO is reported for several languages in contact situations: Quechua in contact with Spanish, Southern Uto-Aztecan languages in contact with Mayan languages, Karaim in contact with Russian and Lituianian, etc. This presentation deals with Caucasian Urum, an Anatolian dialect of Turkish spoken on the Small Caucasus (Georgia). This language differs from Standard Turkish in several respects: postverbal specifics are possible, stress can appear on the postverbal arguments, postverbal arguments may be within the scope of preverbal operators, scope and binding asymmetries are maintained in the postverbal domain. Standard Turkish and Caucasian Urum are two different types of V-final languages: while the right edge of the verb is aligned with the right clause boundary in Standard Turkish, postverbal material is part of the core clause in Caucasian Urum.

The currently spoken varieties of Caucasian Urum display properties of language change. A structural analysis by means of syntactic facts (order of objects, order of adjuncts, behaviour of embedded verb projections) reveals that Young Urum speakers are competent in two registers: beyond the V-final register of their ancestors, they are competent in a register in which the language is re-analyzed as VO. Some parts of this development may have been motivated by their exposure to Russian, but the emerged data pattern differs from Russian in several respects.

**INFORMATION STRUCTURE.** Both varieties share in common that the focus can be postverbal (which applies to VO languages and to OV languages with V-fronting, e.g., in Georgian) and may appear immediately left adjacent to the verb (which applies to OV languages such as Basque and Turkish as well as to VO languages with a left peripheral focus position). The crucial developments are: (a) preverbal material can be focused in the canonical order, i.e., focused subjects may appear in SFOV (while OFSV is excluded) only in Old Urum, but not so in Young Urum; (b) prefocal topics are interpreted as contrastive topics in Young Urum, but not so in Old Urum. While the old variety is a type of V-final language, Young Urum changes to a VO language, with a left peripheral position for topics preceding foci.

# Résumés des présentations/ Abstracts

Elisabeth Verhoeven (Humboldt-University Berlin)

## *Word order and optional ergative marking in Cabécar (Chibcha)*

Chibchan languages are robustly head-final, with basic OV order and postpositions. A subset of the Chibchan languages, among them Cabécar, have been analyzed as ergative on the basis of the morphological properties of the verbal arguments (cf. Quesada 1999, Margery Peña 2003).

Furthermore, the Chibchan languages are reported to allow for case omission under certain conditions related to both word order and information structure (e.g., Chamoreau 2017 on Pesh, Frank 1990 on Ika; Quesada (1999), based on anecdotal evidence from texts, characterizes Cabécar as an optional ergative marking language. However, the specific triggers of ergative case drop and its relation to word order and discourse factors are still largely unexplored.

This presentation provides a detailed account of the syntactic and discourse properties of ergativity in Cabécar, focussing on its intricate interrelation with word order variation. First, it establishes the syntactic facts with elicited data presenting the crucial constructions that provide evidence for

subject/object asymmetries in this language. Second, it examines the properties of ergative marking in a large corpus of narratives (96 narratives produced by 24 native speakers). The data were collected in Costa Rica, Ujarrás, Province of Puntarenas, and Grano de Oro, Chirripó, both traditional settlements of Cabécar people, in which the language is still used in everyday life.

The Cabécar facts are particularly challenging because they show the full-range of properties that are expected for a language with syntactic ergativity (see Dixon 1994, Polinsky 2017). In particular, ergative case is assigned to the agents of transitive verbs (and only these) with a postpositional marker (see Margery Peña 2003:xii). Ergativity is reflected in word order: the object of transitives and the single argument of intransitives must be left adjacent to the verb, while the agent of transitives can either precede or follow the Absolutive-V complex, a property that is otherwise typical for adjuncts.

Furthermore, quantified absolutes are preferably discontinuous with the quantifier appearing in the postverbal domain; this syntactic property does not hold for ergatives. The account is corroborated by evidence from tests of syntactic ergativity.

The second part of this talk will present the corpus facts on the distribution/overt marking of the ergative argument. In particular, we will investigate the impact of givenness and word order on overt ergative marking. By means of Bayesian Network Modelling we will test several dependency scenarios for the explanation of the corpus data distribution. The analysis shows that ergative drop is less pervasive than originally assumed; it occurs in 10.8% of the cases. Ergative marking is primarily determined by word order, i.e. it does not occur in the postverbal domain. Givenness has only an indirect effect on morphological ergative marking, which is clearly mediated through word order.

### References

- Chamoreau, C. 2017. A preliminary grammatical sketch. ELDP Report. August 2017.  
Dixon, RMLJ, 1994. Ergativity. Cambridge: CUP.  
Margery Peña, E. 2003, Diccionario cabécar - español, español - cabécar. San José C.R.: Editorial de la Universidad de Costa Rica.  
Polinsky, M. 2017. Syntactic ergativity. In Everaert & Riemsdijk (eds.) The Wiley Blackwell Companion to Syntax. John Wiley & Sons.  
Quesada, D., 1999, Ergativity in Chibchan. Sprachtypologie und Universalienforschung STUF 52(1), 22-51.

יִשְׂרָאֵל    национален    שפה    文化    شرقية  
i n a l c o

Institut national  
des langues  
et civilisations orientales

Inalco  
65 rue des  
Grands Moulins  
75013 Paris

