Book of Abstracts

Workshop

OV to VO, VO to OV: Word order change from an areal perspective

23-24 June 2017
Bielefeld University

Content

Wilbert Heeringa:
Dialect change in the Dutch dialect area
Fryske Akademy

In the period 2008–2011 we compiled a large database of dialect recordings for 86 local dialects of Dutch spoken in the Netherlands and the northern part of Belgium ('Flanders') and of Standard Netherlandic Dutch and Standard Belgian Dutch. In each of the 86 locations two older male speakers and two younger female speakers were recorded, representing conservative dialect speakers and innovative dialect speakers respectively. Using these data we analyze and visualize dialect change and the influence of standard Dutch on apparent time changes in these dialects.

Using this data and focussing on the lexicon, the morphology and the sound components, we will answer the following questions:

1. Do dialects change? Do the linguistic levels correlate to each other? Which level is affected most strongly?

2. Do dialects converge to standard Dutch?

3. What are the explanatory factors of dialect change?

4. What are the explanatory factors of convergence to/divergencence from standard Dutch?

5. Do dialects converge to each other?

6. Do dialect groups fuse or split or do their mutual relationships change otherwise?

As to the last question we will use a new flavour of bootstrap clustering which generates areas, similar to classical dialect maps. In our approach 1) we consider dialect groups as continua, i.e. each local dialect is not necessarily strongly related to any other local dialect in the same group; the local dialects in a group rather constitute a 'network' and 2) we take into account that not every local dialect can be classified with statistical confidence.

Geoffrey Haig:
The interaction of OV/VO with adpositional type in the Mesopotamian transition zone
Bamberg University

In this paper, I investigate adpositional order and the OV/VO distinction, focussing on the languages of what is provisionally termed the Mesopotamian region (Haig 2017), centred on today's North Iraq and extending into the neighbouring regions of Iran, Turkey and Syria. From a macro-areal perspective, the region represents the transitional belt between an OV/postpositional block of Central Asia and the Indian Subcontinent, and the VO/prepositional block of Afro-Asiatic and the languages of western Europe. Unsurprisingly, the area is a global hotspot for languages with mixed word-order typologies (Stilo 2003). For example, in a worldwide sample of more than 500 languages with OV word order (Dryer 2013a, 2013b), just 14 are found with prepositions. Of these 14, three are west Iranian languages (Persian, Tajik, and Central Kurdish), while the fourth is a Semitic language (Jewish Neo-Aramaic of Arbel) spoken in a predominantly Kurdish-speaking region.

Given the reasonably robust correlation between verb/object ordering, and adpositional type, it is worth considering whether the attested shifts in word order in this region have proceeded independently of adpositional type, or whether they occur as sub-changes within a cluster of related parameter shifts. The latter would be predicted by classical parametric approaches to word-order, including more recent contributions (Haider & Szucsich 2012), as well as the processing-based approach of Hawkins (1994), or formal accounts in terms of the Final-over-Final constraint (Sheehan 2012). I will present data from two Iranic languages (Central and Northern Kurdish) and a Semitic language (Jewish Neo-Aramaic of Urmi, Khan 2008), which suggests that the languages of the region in fact crystallize into what appears to be a stable compromise word order, involving OV in combination with prepositions, and accompanied by a post-verbal positioning of other object types (in particular verb-goal order). This raises more fundamental questions regarding the validity of mainstream parametric approaches to word-order typology: the regularly repeated correlations are based mainly on majority patterns attested in the world's languages, but these could be the epiphenomenal outcome of

contingent historical spread of a small number of initial types. In fact, the typological space for word order variation may be much less constrained than the majority patterns would lead us to believe.

Ralf Vogel:
OV and VO in German VP syntax
Bielefeld University

The classification of German as an OV language is common wisdom among linguists today. Only a few decades ago, this was less clear. For instance, Greenberg (1966), in his pioneering work on linguistic universals, classified German as an SVO language. This is due not to some rapid changes in the grammar of German, but rather to different interpretations of the same grammatical facts and different weighting of seemingly contradictory evidence.

The recent years have shown remarkable progress in our understanding of the historical development of the OV property of German, which includes, most importantly, an understanding of the factors that lead to the mixed picture of OV- and VO-properties we have been observing from early stages until today.

In my presentation, I will review the evidence and insights from current literature, leading to a rough historical sketch of the core developments. Two factors can be identified that continue to play an important role in the phenomena and developments at issue in this discussion:
i) prosody and the syntax-prosody interface
ii) the constant historical trend towards analytic inflection (with its own prosodic implications)

Aria Adli:
Topic chains in dialogues
University of Cologne

In this talk I deal with a subset of referential expressions, namely topic chains in spoken Spanish. While referential chains describe a property of text structure, topics describe a property of common ground (CG) management and bring in the speaker/hearer perspective. More specifically, I ask whether the different status of discourse participants on the one hand and all other speaker/hearer-external referents on the other is reflected in CG management by means of two different type of chains that span over the sentence topics of a discourse or whether there is a single chain regardless grammatical person.

I discuss the adequacy of a model in which two parallel chains are assumed, making use of the correlation between overt vs. null subject pronouns and shift vs. familiar topics. This correlation between referential form and topic continuity exists in many languages, at least with regard to topics that are the subject of the sentence.

Under a double-chain model each sentence topic is specified for three chain-related features: [±shift] indicating whether it is the start or the continuation of a topic chain, [±local] indicating to which of the two chains it belongs, and [±switch] indicating whether the present topic belongs to another chain than the previous topic referent (which would be [+switch]) or to the same chain (which would be [-switch]). Therefore, the double-chain model distinguishes 8 different feature combinations, 4 regarding the local chain and 4 regarding the nonlocal chain.

[+local]				[-local]			
[+switch]		[-switch]		[+switch]		[-switch]	
[+shift]	[-shift]	[+shift]	[-shift]	[+shift]	[-shift]	[+shift]	[-shift]

The double chain model is assessed using the Spanish data from the sgs corpus (www.sgscorpus.com), consisting of spoken dialogue data that I have collected during fieldwork in Barcelona in the year 2008 with Spanish native speakers. The data consists of transcribed and annotated recordings with 54 persons. The quantitative analyses build on a pool of

3,737 sentences that are then split according to the features [±shift], [±switch], and [±local].

The results show that chain-switching is an important factor. It lowers the effect of topic shift on pronoun rate. Based on a closer examination of discourse examples, two factors are identified: First, *ease of disambiguation* between different referents of the nonlocal chain as the relevant factor. Second, local referents are generally situationally evoked, they generally show a high level of activation. I come to the conclusion that speaker and hearers operate in their common ground management with two parallel or superposing prominent referential structures, one linked to the speakers themselves and one to the entities they talk about. This view also suggests that the interaction between the two information-structural notions topic and givenness deserve further attention, because the entities of the local chain have a special status with regard to givenness.

Michael Rießler:
NLP in endangered language documentation: Building and investigating corpora for under-resourced languages of the Barents Sea region
Freiburg University

Despite the digital nature of current methodology, documentary linguistics has rarely considered applying computational methods in building and analyzing endangered language corpora more efficiently. In my talk, I will provide a brief overview of work-in-progress in a project which is one of the very first attempts to work in the paradigm of endangered language documentation and description while systematically applying methods from Natural Language Processing (NLP).

The languages I work most intensively on at present are endangered and highly endangered Uralic language from the Permic, Saamic and Samoyedic branches of Uralic. All languages are spoken in Barents Sea region of northeastern Europe.

I will illustrate the workflows in my projects – as well as the potential of computational linguistic approaches to endangered language documentation – based on a quantitative small-scale case study on constituent order variation across Kildin Saami and Komi spoken and written varieties.

Diana Forker:
The variablity of constituent order in Sanzhi Dargwa
Universität Bamberg & Universität Jena

Sanzhi Dargwa is a Nakh-Daghestanian language spoken by around 250 people living in Daghestan (Russian Federation, Caucasus). As all Nakh-Daghestanian languages, Sanzhi Dargwa has (predominantly) head-final constituent order. Noun phrases are head-final and thus modifiers generally precede the head. However, all types of modifiers except for demonstrative pronouns can also follow their head, which in some cases requires additional marking. Similarly, the constituent order at the clause level is most commonly SOV, but all logically possible alternative orders are found in texts and can easily be elicited. Among them, SVO is the preferred alternative to SOV. Only subordinate clauses show a clear and sometimes rigid preference for the final position of the verb.

In this talk, I will analyze the constituent order of Sanzhi Dargwa and provide an account of the contexts that allow for heads to be placed in positions other than the final one. It will be shown that information structure plays a major role in distribution of the different constituent order possibilities. I will also discuss the possibility of Russian influence on the word order at the clausal level.

The data analyzed in this talk originate from extensive fieldwork in Daghestan since 2012.

Johanna Lorenz:
Complement clauses in Caucasian Urum
Bielefeld University

The talk deals with the complement clause system in Caucasian Urum. Urum is a variety of Anatolian Turkish spoken by ethnic Greeks in the Small Caucasus in Georgia, the speech community is characterised by bi- and multilingualism, including both genealogically and typologically distinct languages. Almost every speaker is fluent in Russian and most are competent in Georgian as well.

Several studies report changes of subordination patterns in Turkish or Turkic languages in settings with dominant Indo-European languages. They generally claim that a growing realisation of finite complement clauses introduced by subordinating conjunctions substitutes the primarily synthetic morphological marking in left-branching Turkic complement clauses. Similar phenomena are observable in Caucasian Urum.

The study attempts to compare the licensing conditions which determine the distribution of complement clause types in Turkish, Russian and Urum by focusing on constraints in structural variations and their relations to different types of complement-taking predicates. Furthermore, deviations between complement clause types in Caucasian Urum and Turkish will be analysed by looking at their properties as syntactic constructions, e.g. at functional categories within a minimalistic approach.

Stefanie Böhm:
Word order change and information structure in Caucasian Urum
Bielefeld University

A current development which can be observed all of over the world is the emergence of multilingual contact situations. A natural consequence of language contact is the development of bilingual communities. This talk deals with the effect of language contact on the information structure in Caucasian Urum. Caucasian Urum (henceforth: Urum) is an Anatolian Variety of Turkish, which is spoken by a small minority of ethnic Greeks in the Small Caucasus in Georgia. The ancestors of the Urum speakers came from several cities in North Eastern Anatolia (e.g., Kars, Erzurum, Bayburt) and moved to the Caucasus in the beginning of the 19th century. Since that time Urum speakers were in close contact with the other languages of the Caucasus, predominantly with Russian.

The influence of Russian is particularly visible in the syntax of Urum, i.e., both OV and VO orders occur under similar discourse conditions. Within this talk I will present an overview of empirical studies that investigate the effect of word order and information structure in Caucasian Urum. The results of the studies reveal that the change in the Urum word order from OV to a language with a free position of the verb within the VP lead to an extension of the information structural possibilities of the language. Whereas foci in Turkish are restricted to the preverbal field and may not occur in the postverbal area, foci in Russian and Urum may felicitously occur in any position of the clause, i.e., sentence-initially (SFocOV, SFocVO, OFocSV, OFocVS), postverbally (OVSFoc, SVOFoc) or preverbally (OSFocV, SOFocV).

Erika Asztalos:
From SXV to SVX in Udmurt: a Russian-induced ongoing change
Hungarian Academy of Sciences

Claims: Contrary to the commonly assumed view that Udmurt (Uralic; Russian Federation) is a relatively consistent SOV language (cf. Vilkuna 1998, Winkler 2011), I argue that Udmurt is undergoing a shift from the SXV to the SVX type under the influence of Russian, but this change is proceeding more slowly in the areas where Tatar is also spoken. The change proceeds via a gradual broadening of the range of information structure roles associable to the postverbal constituents. The *head + complement* directionality change manifests itself on different syntactic levels (VP, CP, NP, AdjP), but the single phrase types are not equally prone to change their internal order.

Methods: The research was conducted by means of textual analysis and a questionnaire. Textual analysis included the comparison of the proportion of head-initial and head-final constituents in old and contemporary texts, as well as the information structural analysis of sentences containing head-initial phrases. The questionnaire (which was filled out by 90 native speakers) tested, using discourse-neutral sentences, whether the acceptability and the usage of the head-initial variants of Dryer's (1992) *correlation pairs* is influenced by generational and/or by areal factors.

Results:
Textual analysis: The proportion of the head-initial constituents resulted to be higher in the contemporary texts than in the old ones. Textual analysis also reveal that the information structural distribution of the SVX sentences is rather wide: SVX sentences can contain informationally old and informationally new postverbal elements (i. e., postverbal topics and foci), and they can be discourse-neutral, as well. These findings support the claim concerning the ongoing typological change of Udmurt. I suggest that the change is occurring via a gradual broadening of the range of the possible information structural roles of the postverbal elements (discourse-old → discourse-new → discourse-neutral).

Questionnaire: Head-initial order variants were produced and judged grammatical with almost all of the examined constituent types.

The percentage of the informants who produced the head-initial orders and judged them grammatical was consistently higher among the young speakers compared to the older ones. In an *apparent time hypothesis-* approach, such a difference might indicate that Udmurt is undergoing a change from the SXV to the SVX type. A plausible reason for that can be the fact that while the older speakers are Udmurt-dominant bilinguals, the younger ones are Russian-dominant bilinguals (Salánki 2007: 59).

Areal factors also play a role. The informants from Tatarstan had a stronger preference for the head-final orders than the ones living in Udmurtia. This was especially striking in the case of the older informants, but at several test questions the young speakers from Tatarstan used head-final orders more consistently than those from Udmurtia. I suggest that the influence of (head-final) Tatar makes Udmurt more "resistent" to the influence of (typologically distinct) Russian.

Inclination of the examined phrase types to word order change: The examined constituent types did not show an equal inclination to change their internal order. The most prone to the change resulted to be V + S (in existential and possessive sentences) and N + Rel, whereas the least prone proved to be N + Gen, Adj + Stand and Cop + Pred. This suggests that the V + O pair of elements does not play a particular role in the directionality change of Udmurt.

References:
Dryer, Matthew S. 1992. The Greenbergian Word Order Correlations. Language 68: 81–138.
Salánki Zsuzsanna 2007. Az udmurt nyelv mai helyzete. Doctoral dissertation. Budapest, Eötvös Loránd Tudományegyetem.
Vilkuna, Maria. 1998. Word order in European Uralic. In: Anna Siewierska (ed.), Constituent order in the languages of Europe. Berlin–New York: Mouton de Gruyter, 173–235.
Winkler, Eberhard. 2011. Udmurtische Grammatik.

Stavros Skopeteas:
From OV to VO in Caucasian Urum
Bielefeld University

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 Lituanian, etc. Generalizations on the OV-to-VO change are mostly based on linearization preferences, which is a reasonable observable but not necessarily the transferred entity at issue. The crucial question is whether the observed preferences result from a change in the structure of the involved syntactic projections.

This presentation deals with Caucasian Urum, an Anatolian dialect of Turkish spoken on the Small Caucasus (Georgia). An examination of the syntax of Urum reveals that this language does not display the set of properties that appear with right-dislocated constituents in Turkish: postverbal specifics are possible, stress can appear on the postverbal arguments, postverbal arguments can be under the scope of *only*, binding asymmetries are maintained in the postverbal domain.

This presentation reports the results of an acceptability study on the order of adjuncts and the order of verb clusters (conducted with 24 speakers of Urum, 12 speakers of Turkish, and 12 speakers of Tbilisi Russian). The findings of this study show an interesting difference between Age Groups of Urum: in contrast to Turkish, the linearization of the VP subconstituents follows general principles of V-final languages and is independent of the position of the matrix verb. Young Urum speakers show a complex data pattern that contains two components: (a) V-final linearizations display the properties of the variety of their ancestors; (b) when the V is non-final, their linearization preferences reveal a pattern that corresponds to head-initial VPs.

Nicolaos Neocleous:
VO and OV alternation in Romeyka: Word order variation in a minimalist system
University of Cambridge

The goal of this paper is two-fold: (i) descriptively, to investigate VO and OV alternation in Romeyka, and (ii) theoretically, to account for such word order variation in a minimalist system. Romeyka is the only variety of Asia Minor Greek that is still spoken in the area historically known as Asia Minor (modern-day, Anatolia, Turkey). Crucially, all Romeyka speakers are bilingual in both Romeyka and Turkish (*see* Sitaridou (2013)). As such, Romeyka shares linguistic features with Greek and Turkish. Interestingly, while the canonical order in Romeyka is VO, focus always appears to the left of the verb, resulting OV orders. In order to account for such word order variation in Romeyka, I aim at pursuing of a third-factor (principled) explanation for word order variation, within the minimalist program.

Given Richards' (2007) assumption that the system that is in conformance with a maximally empty UG is one in which phases are pairs of phase heads and nonphase heads (hence the core sequence C–T–v*–V), the possible expansions of the core functional sequence into more richly articulated hierarchies have to be constrained. Therefore, in a minimalist system discourse-related features cannot project into the narrow syntax.

If Richards' (2007) assumption that C-T-v*-V is the backbone of a clause is right, two questions arise: how can we account for such a theory if argument structure information projects into syntax, and if specifiers have a different syntactic role from complements? In line with Lohndal (2012), I assume that specifiers have the same syntactic role as the complements, and that argument structure information does not project into syntax.

Further, I argue for a linear syntax. In particular, there is good reason to assume that verb raises to T^0 in Romeyka, and subject is left dislocated; hence in an SOV order, in which the subject is in [Spec, CP], and the verb raises to T^0, the (focused) object must appear in [Spec, TP]. In order to account for this, I propose that along uninterpretable features

(*u*Fs) and Edge features (EFs), there must be another kind of formal features, i.e. the linearisation feature ^, in the sense of Biberauer et al. (2014); when ^ is associated with the EF of v^0, it triggers movement of the complement VP to [Spec, vP]. The ^ feature can be spread upwards, triggering further movement of the vP to [Spec, TP]. Said that, it is not the object that raises to [Spec, TP], but rather the vP. The ^ feature can account for efficient computation into the interface conditions of UG, i.e. it has a crucial role in mapping syntax into PHON and SEM; phases are mapped into the interfaces as ordered sets. This is crucial to account for intervention effects (in the sense of Beck (2006)).

References

Beck, S. (2006). Intervention Effects Follow from Focus Interpretation. *Natural Language Semantic , 14* (1), 1-56.

Biberauer, T., Holmberg, A., & Roberts, I. (2014). A syntactic universal and its consequences. *Linguistic Inquiry* (45.2), 169-225.

Lohndal, T. (2012). *Without Specifiers: Phrase Structure and Events.* University of Maryland: Unpublished PhD dissertation.

Richards, M. (2007). On feature-inheritance: an argument from the Phase Impenetrability Condition. *Linguistic Inquiry , 38* (3), 563-572.

Sitaridou, I. (2013). Greek-speaking enclaves in Pontus today: The documentation and revitalization of Romeyka. In M. C. Jones, & S. Ogilvie, *Keeping languages alive. Language endangerment: documentation, pedagogy and revitalization.* (pp. 98-112). Cambridge: Cambridge University Press.

Erika Asztalos, Katalin Gugán and Nikolett Mus:
A path from OV to VO: glimpses from Uralic
Hungarian Academy of Sciences

The order of clausal constituents interacts with their discourse pragmatic function, and constituent order and patterns of discourse-pragmatic organization are both prone to contact- induced change (Aikhenvald 2006). In our talk, we survey patterns of (X)VO, or more precisely, (X)VX orders in three genetically related (S)OV languages, i.e., in Udmurt (Permic), Surgut Khanty (Ob-Ugric) and Tundra Nenets (Samoyedic), all of which are under heavy Russian, i.e., SVO influence. On the basis of a predominantly corpus-based research involving the comparison of two periods, we will show that the frequently hypothesized XV > VX change involves, self evidently, an increase in the proportion of non-verb-final clauses. However, the analysis of the postverbal constituents with respect to their syntactic and their discourse-pragmatic function shows that the first steps of such a change can be described as the broadening of the potential discourse-pragmatic function of the postverbal elements, which may ultimately lead to the reanalysis of the basic word order. The comparison of the individual languages also opens the door to outline the phases of such a change, the units of which are discourse-pragmatic categories.

Our research is based on corpus analyses and consultations with native speaker informants. The examination covers diachronic data collected during the second half of the 19th century and in the beginning of the 20th century, that is, during the first expeditions that aimed at documenting and describing these languages. These (relatively) old data will be compared to those representing contemporary language use, obtained from an electronic corpus that is under construction at RIL, HAS.

Comparisons of genetically related SOV languages (e.g., Hindi and Bangla, see Simpson & Choudhury 2015, as well as, Khalkha Mongolian and Uyghur Turkic, see Öztürk 2013) have shown that both the potential discourse-pragmatic functions of postverbal constituents and the syntactic derivations accounting for their postverbal position are different in each of these languages.

Our investigation focussed on the discourse-pragmatic function of postverbal constituents and their syntactic function in the clause, in order to see which category – the syntactic or the pragmatic one – allows for a broader generalization. The analysis of the postverbal constituents leads to establish the following types: (i) the postverbal constituent is pragmatically analyzed as a clarifying afterthought, having an (optional) coreferential intra- clausal element; (ii) the postverbal constituent is a topical element of the clause, the postverbal position of which is due to backgrounding; (iii) the postverbal constituent is either an argument focus or part of the predicate focus (in Lambrecht's sense); (iv) postverbal constituents of all-new, i.e., sentence-focus sentences. Whereas types (i) and (ii) are assumed to be derived from an SOV word order by a pragmatically motivated syntactic operation, type (iii) and type (iv) are argued to reflect interference. In the last type, furthermore, the VX order is not motivated by any pragmatical and/or informational structural roles, i.e., it can be analysed as a non-derived SVO order. As opposed to the pragmatic types, the syntactic function of the postverbal constituents varies considerably across the languages and the time stages to be discussed. Besides, the available patterns are different both language-wise and period-wise, lending themselves to postulate certain phases of the ongoing change.

References:
Aikhenvald, Alexandra Y. 2006. Grammars in Contact: A Cross-Linguistic Perspective. In: Aikhenvald, Alexandra Y. - Robert M.D. Dixon (eds), *Grammars in Contact: A Cross- inguistic Typology.* Oxford University Press, Oxford. 1-66.Lambrecht, Knud 1994. *Information structure and sentence form*. Cambridge University Press, Cambridge.
Simpson, Andrew–Choudhury, Arunima 2015. The nonuniform syntax of postverbal elements in SOV languages: Hindi, Bangla, and the Rightward Scrambling debate, *Linguistic Inquiry* 46: 533–551.
Öztürk, Balkız 2013. Postverbal constituents in SOV languages, In: Biberauer, Theresa – Sheehan, Michelle (eds.), *Theoretical approaches to disharmonic word order.* Oxford University Press, Oxford. 270–305.