

Minimal vowel systems in cross-linguistic perspective

This paper presents the results of a comprehensive cross-linguistic survey of minimal vowel systems and discusses a number of theoretical issues which arise from them. Although the existence of minimal vowel systems has been acknowledged since the early days of structuralist linguistics (e.g. Jakobson 1923), even today they are often neglected or misunderstood in discussions around phonological typology, and vowel system typology in particular. Minimal vowel systems pose difficulties for many mainstream phonological theories and for vowel system modelling and for this reason an exhaustive survey of such systems is a desideratum; one that is addressed here.

A 'minimal' vowel system is considered to be one in which the fewest phonological contrasts can serve to define the members of the set of vowels. Typically, this means that only first formant, and not second formant, distinctions are necessary to define the set of vowels. While such systems have been described as 'vertical' or 'linear' by previous scholars, these terms are avoided here in order to include analyses with one vowel (e.g. Barreteau 1988) or no vowels whatsoever (e.g. Kuipers 1960).

The survey was carried out by conducting a review of the secondary literature on vowel systems to identify relevant languages, followed by a detailed analysis of their phonological descriptions. In some instances, minimal vowel systems have been described widely across phylogenetic or areal groups of languages, such as in Northwest Caucasian (Hewitt 2005), the Arandic languages of Australia (Breen 2001), the Sepik-Ramu languages of Papua New Guinea (Foley 1986; Comrie 1991), the Goidelic languages (Ó Siadhail and Wigger 1975; McConville 2013), and the Central Chadic languages (Barreteau 1987). In other cases, individual languages have been described as having minimal vowel systems, such as Chinese (Chao 1968), Marshallese (Bender 1968), the Caddoan language Wichita (Rood 1976), the Salishan language Nuxálk (Nater 1984), and the Arnhem language Anindilyakwa (Leeding 1989).

The great geographical and phylogenetic diversity of these languages offer fertile terrain for scholarly attention in phonological typology, but unfortunately, this has not always been forthcoming. While minimal vowel systems are described in early surveys of vowel systems such as Trubetzkoy (1939) and Hockett (1955), and are discussed in Liljencrants and Lindblom's (1972) work in dispersion theory, in subsequent overviews and attempts at modelling they have often been dismissed (e.g. Crothers 1978), or not discussed at all (Lass 1984; Maddieson 1984; Schwartz et al. 1997).

The survey shows that the majority of minimal vowel systems are two-member or three-member vertical systems, with a smaller number of descriptions with only one vowel (e.g. Smith 1999) and isolated instances without contrastive vowels (Kuipers 1960; Pulleyblank 1983). Often, but not always, the vowels in minimal systems exhibit extensive allophony conditioned by surrounding consonants or by prosodies applying over larger phonological domains. It is argued that these phenomena are best captured by an abstract model of phonology which does not reify the segment. In such a model, a number of further vowel systems, in which only second formant, but not first formant distinctions, are necessary to define the set of vowels, can be understood in the same terms as the minimal vowel systems discussed in this paper.

List of works cited

- Barreteau, Daniel (1987). 'Du vocalisme en tchadique' in Barreteau ed. *Langues et cultures dans le bassin du lac Tchad*. Paris: ORSTOM, 161-91.
- Barreteau, Daniel (1988). *Description du Mofu-Gudur: Langue de la famille tchadique parlée au Cameroun. Livre 1: Phonologie*. Paris: Orstrom.
- Bender, Byron W. (1968). 'Marshallese phonology' in *Oceanic Linguistics* 7, 16-35.
- Breen, Gavan (2001). 'The Wonders of Arandic Phonology' in Simpson et al. eds., 45-69.
- Chao, Yuen Ren (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Comrie, Bernard (1991). 'On Haruai vowels' in Pawley ed. *Man and a Half: Essays in Pacific Anthropology in Honour of Ralph Bulmer*. Auckland: The Polynesian Society, 393-97.
- Crothers, John (1978). 'Typology and universals of vowel systems' in Greenberg et al. eds., 93-152.
- Foley, William A. (1986). *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Halle, Morris (1970). 'Is Kabardian a vowel-less language?' in *Foundations of Language* 10.6, 95-103.
- Hewitt, George (2005). 'North West Caucasian' in *Lingua* 115, 91-145.
- Hockett, Charles F. (1955). *A Manual of Phonology*. Baltimore: Waverley.
- Jakovlev, Nikolai F. (1923). Таблицы фонетики кабардинского языка [Tables for the phonetics of the Kabardian language]. Moscow.
- Kuipers, Aert H. (1960). *Phoneme and Morpheme in Kabardian (Eastern Adyghe)*. The Hague: Mouton.
- Lass, Roger (1984). 'Vowel System Universals and Typology: Prologue to Theory ' in

- Phonology Yearbook* 1, 75-111.
- Leeding, Velma Joan (1989). *Anindilyakwa Phonology and Morphology*. PhD dissertation, University of Sydney.
- Liljencrants, Johan and Lindblom, Björn (1972). 'Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast ' in *Language* 48.4, 839-62.
- Maddieson, Ian. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- McConville, Mark (2013). 'Reverse-engineering the morphophonology of Gaelic vowels from orthography: Language planning to linguistic theory, and back again' in Cruickshank and McColl Millar eds. *After the Storm: Papers from the Forum for Research on the Languages of Scotland and Ulster triennial meeting, Aberdeen 2012*. Aberdeen: Forum for Research on the Languages of Scotland and Ireland, 61-81.
- Nater, Hendricus Franciscus (1984). *The Bella Coola Language*. Ottawa: National Museums of Canada.
- Ó Siadhail, Micheál and Wigger, Arndt (1975). *Córas Fuaimeanna na Gaeilge*. Baile Átha Cliath: Institiúid Ard-Léinn.
- Pulleyblank, Edwin G. (1983). *Middle Chinese: A Study in Historical Phonology*. Vancouver: University of British Columbia Press.
- Rood, David S. (1976). *Wichita Grammar*. New York: Garland.
- Schwartz, Jean-Luc; Boë, Louis-Jean; Vallé, Nathalie ; Abry, Christian (1997). 'Major trends in vowel system inventories' in *Journal of Phonetics* 25, 233-53.
- Smith, Tony (1999). *Muyang Phonology*. Yaoundé: SIL.
- Trubetzkoy, Nikolai S. (1939). *Grundzüge der Phonologie*. Prague: Travaux du Cercle Linguistique de Prague 7. Pagination here from the English translation by Baltaxe, Christiane A. M. [1969]. *Principles of Phonology*. Berkeley: California.

Cormac Anderson

On developments in the vowel systems of two Even dialects

Natalia Aralova, Max Planck Institute for Evolutionary Anthropology

My research aims to present independent developments in the vowel systems of two Even dialects (Northern Tungusic) which are spoken in remote areas of Eastern Siberia and are undergoing contact influence from different neighboring languages. Until recently, the only dialect whose phonetics and phonology have been described in detail was that of Ola (Magadan region). This dialect also serves as the the base for the Standard Even. It has been described as having two vowel sets opposed by pharyngealization: /iʰ aʰ oʰ uʰ iaʰ/ vs. /i e o u ie/. The opposition is realized as root-controlled vowel harmony.

Some peripheral Even dialects do not have the same symmetrical vowel systems as in Ola and at the same time appear to be quite different from each other. I have specifically studied the westernmost one, which is called Lamunkhin and is spoken in Sebian-Küöl in the Republic of Sakha (Yakutia), and one of the Eastern dialects, that of the Bystraia district of Kamchatka. The results of acoustic and perception studies carried out for both of these dialects are shown in Table 1. The upper part of this table shows statistically significant acoustic differences; the lower part corresponds to the ability of Even speakers to recognize quasi-minimal pairs containing vowels of different harmonic sets. One of the interesting features of the Sebian-Küöl dialect is the merger of the harmonic opposition of /i/ vs. /i/ supported by both acoustic and perception data. Another striking feature is the difference between the upper and lower parts of Table 1 demonstrating the reduction of the opposition between harmonic sets in the perceptual data. In the Bystraia dialect all harmonic oppositions are absent except for /a/ vs. /e/, in Sebian-Küöl the oppositions /a/ vs. /e/ and /o/ vs. /ɔ/ are kept.

In my talk, I will enter into more detail about the acoustic parameters investigated and discuss the design of my perception study. The present study is not the first to have observed discrepancies between acoustic and perceptual measurements. I will propose an explanation for the results along the lines of the ‘near mergers’ of Labov (1994). Moreover, I will view the developments in the vowels systems in the light of the contact situations in which the dialects are involved. The clear distinction between /e/ and /o/ (based on frontness) in Sebian-Küöl could be due to Sakha influence. The ongoing simplification of the vowel system together with the increasing role of consonantal cues and confusion of diphthongs in the Bystraia dialect is likely due to Russian influence.

Table 1. Systems of monophthongs in the Bystraia and Sebian-Küöl dialects (on the basis of the acoustic and perception studies)

		Bystraia dialect			Sebian-Küöl dialect		
		front	mid	back	front	mid	back
acoustic study	high	i i:		u u:	i i:		u u:
		ɪ ɪ:		ʊ ʊ:			ʊ ʊ:
	mid	e e:		o o:	e e:	ə ə:	o o:
				ɔ ɔ:			
low		a a:			a a:		
perception study	high	i i:		u u:	i i:		u u:
	mid	e e:		(o o:)	e e:	ə ə:	o o:
	low		a a:			a a:	

Peter Arkadiev (Institute of Slavic Studies, Russian Academy of Sciences, Moscow)
Andrey Shluinsky (Institute of Linguistics, Russian Academy of Sciences, Moscow)

Towards a typology of derivational viewpoint aspect systems

Starting from Dahl's (1985) notion of 'Slavic-style aspect' it is generally recognized that the 'basic' aspectual opposition of perfective vs. imperfective / progressive may be organized in two ways in the languages of the world. Some languages have an inflectional aspect system where perfectivity resp. imperfectivity is expressed by obligatory TAM markers and is often dependent on tense (e.g. the "aorist" vs. "imperfect" distinction in many languages). Other languages have a derivational aspect system where perfectivity resp. imperfectivity is a lexical feature of a verb, and a verb expressing the opposite viewpoint may be derived via imperfectivizing resp. perfectivizing derivational markers, which are non-obligatory, independent of other TAM values and often lexicalized. For all that we know, derivational aspectual systems are a cross-linguistic rarity; to date, they have been discussed almost exclusively based on the data of Slavic and neighbouring languages, see e.g. Breu 1992, Dickey 2000, Tomelleri 2010. As a result, the peculiarities of Slavic aspect are implicitly attributed to derivational aspect systems in general, without recognizing them as a broader cross-linguistic type where Slavic systems form just one — and rather idiosyncratic — subclass.

Still, derivational systems radically different from the better known Slavic type are attested in a number of languages of diverse areas and genetic groupings. First, these are the non-Slavic languages of Eastern Europe and the Caucasus, which are known to be to some extent similar to Slavic languages in the morphology of aspect: Baltic languages [< Indo-European], Ossetic [< Iranian < Indo-European], Yiddish [< Germanic < Indo-European], Istro-Romanian [< Romance < Indo-European], Livonian [< Finnic < Uralic], Ugric languages [< Uralic], Kartvelian languages. Second, these are some languages of Siberia, such as Samoyedic [< Uralic] and Evenki [< Tungusic < Altaic]. Third, these are some (yet) sporadic examples of languages from other parts of the world, such as Chadic languages (Margi), some Oceanic languages (Mokilese and Kusaiean), some languages of South America (Quechuan, Aymaran and Mapuche [< Araucanian]), some languages of North America (Pomoan languages and West Greenlandic [< Eskimo-Aleut]).

We present the results of a pilot cross-linguistic study of such systems based on available descriptions and our own fieldwork data; our sample is neither exhaustive nor balanced, however, we believe that it allows us to make some preliminary generalizations.

We distinguish three subtypes of derivational aspect systems.

1. Perfectivizing derivational aspect systems in which most of the underived verbs are imperfective. In such systems perfectivizing markers are at least partially based on locative expressions (cf. the notion of 'bounder-based perfectives' from Bybee & Dahl 1989), and therefore individual languages usually have many perfectivizers. Imperfectivizing markers may be absent. Perfectivizing markers expressing the culmination of a telic process ('eat smth.' > 'eat smth. up') are always present in such systems, while inchoative or delimitative perfectivizing markers for states or atelic processes may be absent. Such systems are found in Slavic and neighbouring languages, as well as in Margi and Oceanic languages.

2. Imperfectivizing derivational aspect systems in which most of the underived verbs are perfective. In such systems there are always more than two imperfectivizing markers, while perfectivizing markers may be absent. Those systems which have perfectivizing markers always have them for perfectivization of stative or activity verbs, while perfectivization of telic processes may be absent. Such systems are characteristic of Samoyedic languages, Evenki and Mapuche.

3. Derivational aspect systems with no acute trend to perfectivizing or imperfectivizing in which the ratio of perfective and imperfective underived verbs is not abrupt and both perfectivizing and imperfectivizing derivations are productive. Such systems are found in Quechua, West Greenlandic and Pomoan languages.

References

- Breu, Walter (1992). Zur Rolle der Präfigierung bei der Entstehung von Aspektsystemen. In: M. Guiraud-Weber, Ch. Zaremba (éd.), *Linguistique et slavistique. Melanges offerts à Paul Garde*, t.1. Paris, Aix-en-Provence: Presses universitaires de Provence, 119–135.
- Bybee, Joan L., & Östen Dahl (1989). The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13.1, 51–103.
- Dahl, Östen (1985). *Tense and Aspect Systems*. Oxford: Blackwell.
- Dickey, Stephen M. (2000). *Parameters of Slavic Aspect: A Cognitive Approach*. Stanford: CSLI.
- Tomelleri, Vittorio S. (2010). Slavic-style aspect in the Caucasus. *Suvremena lingvistika* 69, 65–97.

In borrowing, anything goes, but not really

Dik Bakker, Amsterdam

Ewald Hekking, Querétaro

Abstract

It has been suggested that there would, in fact, be no constraints on what can be borrowed by one language from another, the only real explanatory factor being cultural, i.e. whatever speakers of the Target language deem interesting in the Source language, with which they are in contact (cf. Thomason 2001). Undoubtedly, all kinds of pragmatic factors may play an important role in what is borrowed and what is not. However, in our research over the last two decades into borrowing from Spanish by several typologically very different Amerindian languages (cf. Bakker et al. 2008 - 2012), we think we have found evidence for the assumption that structural similarities and dissimilarities between Source and Target language might play a supporting or a constraining role, at least in the earlier stages of the loan process. We derive our observations from three corpora that we collected - for Otomi (Otomanguean, Mexico), Quichua (Quechua, Ecuador), and Guarani (Tupi, Paraguay) – each containing excerpts of spoken texts by between 38 and 59 native speakers. For all three languages, we selected two dialects between which no systematic contact had been attested. In all three cases, we established a great deal of lexical borrowing from Spanish, with an average over all informants of between 14.1% (Otomi) and 18.9% (Quichua) of the tokens.

The paper discusses several aspects of the three languages concerned from the perspective of borrowing. We will look at the lexicon - both the major parts of speech and grammatical items -, morphology, and syntax. Illustrative for our approach is the borrowing of adpositions. Otomi, a VOS language, has no adpositions at all, but it has relational particles in pronominal position. Quichua has basic order SOV, and possesses a host of enclitical case markers. Guarani is a SVO language, and has many postpositions. In the Otomi corpus we found an instance of

virtually all current Spanish prepositions, spread over all informants, together totalling to some 20% of all the loanword tokens found in that language. For the other two languages, the number of attested Spanish prepositions was less than 1% of that total, used only by a minority of the speakers, sometimes as a postposition. In all cases, these figures were consistent for the two dialects. This is suggestive of the fact that lexical or grammatical aspects play a role in the borrowing process, at least for these items. Similar cases can be built for the borrowing of adjectives, of derivational and inflectional morphology, of coordination, and subordination, among others.

We think that studying language contact, borrowing, and language change from this angle may provide us with yet another perspective on understanding the diversity among the languages of the world.

References

- Bakker, D., J. Gómez-Rendón & E. Hekking (2008). 'Spanish meets Guaraní, Otomí and Quichua: a multilingual confrontation'. In Th. Stolz, D. Bakker & R. Palomo (eds) *Aspects of Language Contact*. Mouton de Gruyter, 165-238.
- Hekking, E. & D. Bakker (2009). 'Loanwords in Otomi, an Otomanguan language of Mexico'. In M. Haspelmath & U. Tadmor (eds) *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Mouton de Gruyter, 897-917.
- Bakker, D. & E. Hekking (2010). 'Is Otomí creating a new lexical class for the modification of noun phrases as a result of the contact with Spanish?'. *STUF* 63, 5-19.
- Bakker, D. & E. Hekking (2012). 'Constraints on Morphological Borrowing: Evidence from Latin America'. In: Lars Johanson & Martine Robbeets (eds) *Copies versus Cognates in Bound Morphology*. Leiden: Brill, 187-220.
- Bakker, D. & E. Hekking (2012). 'Clause Combining in Otomi Before and After Contact with Spanish'. *Linguistic Discovery* 10.1, 42-61.
- Thomason, S. G. (2001). *Language contact. An introduction*. Edinburgh: Edinburgh University Press.

On rare typological features of the Zamucoan languages, in the framework of the Chaco linguistic area

Pier Marco Bertinetto & Luca Ciucci (Scuola Normale Superiore, Pisa)

The Zamucoan family only includes two surviving endangered languages: Ayoreo (AY) and Chamacoco (CH), spoken in northern Chaco between Bolivia and Paraguay by approximately 4500 and 2000 people, respectively. The Zamucoan family also includes an extinct language, Ancient Zamuco (AZ), described in the 18th century by the Jesuit Father Ignace Chomé. AZ is very close to AY from the lexical point of view, but shows striking morphosyntactic correspondences with CH; this allows robust diachronic insights (Ciucci 2013; Ciucci & Bertinetto, submitted).

This talk will address typological features which turn out to be rarely attested in the world's languages (cf. the Konstanz *Raritätenkabinett*). Some of these rarities can be observed in both AY and CH and probably stem from Proto-Zamuco. Among these one can mention: (1) the tripartite distinction within nominal morphology, crucially involving one form characteristically devoted to nominal predication plus two forms for nominals in argumental position with specific vs unspecific reference; (2) the radical lack of tense and aspect in verb morphology (Bertinetto 2014); (3) the presence of both gender and classifiers, or more specifically of gender and number in possessive classifiers (Bertinetto & Ciucci, in preparation); (4) the so-called para-hypotactic strategy, involving the simultaneous presence of both coordinating and subordinating connectives in sentences including a proleptic dependent clause (Bertinetto & Ciucci 2012). Other typological rarities, although equally inherited from Proto-Zamuco, are only found in AZ and occasionally in AY: (5) the presence of nominal tense (AZ) and of nominal aspect (AZ, AY) – the latter being an absolute typological rarity – and (6) the traces of a conjunct/disjunct system (AZ). The diachronic reconstruction proposed by Ciucci & Bertinetto (submitted) reveals other rarities stemming from relatively recent developments in the verb morphology of individual languages, such as: (7) the first person zero marking in the AY *realis* mood; (8) the unexpected affix order in the CH third person plural (Bertinetto 2012); (9) the combination of greater plural and clusivity in CH (Ciucci 2013).

According to some scholars, Gran Chaco constitutes a linguistic area (Comrie *et al.* 2010). Indeed, some of the above-mentioned features may also be found in other surrounding languages: in particular (4) and (8), and to some extent (3). As a matter of fact, the Chaco populations, although traditionally in mutual conflict, were in narrow contact for centuries, so that lexical and morphological borrowings could occur. In at least one case (8), one can prove that the Zamucoan languages were the source of the influence (Ciucci, to appear).

Bibliographical References

- BERTINETTO, Pier Marco 2012. How the Zamuco languages dealt with verb affixes”, *Word Structure* 4,2: 215-230.
- BERTINETTO, Pier Marco 2014. Tenselessness in South American indigenous languages with focus on Ayoreo (Zamuco). *LIAMES* 14. 149-171.
- BERTINETTO, Pier Marco & Luca Ciucci 2012. Parataxis, hypotaxis and para-hypotaxis in the Zamucoan Languages. *Linguistic Discovery* 10.1. 89-111.
- BERTINETTO, Pier Marco & Luca Ciucci, in preparation. *Possessive classifiers in gender-marking Ayoreo and Chamacoco*. Paper presented at the workshop *Gender and classifiers: cross-linguistic perspectives*. Guildford, University of Surrey. January 14, 2014.
- CIUCCI, Luca 2013. *Inflectional morphology in the Zamucoan languages*. Ph.D. thesis. Pisa: Scuola Normale Superiore.
- CIUCCI, Luca (to appear). Tracce di contatto tra la famiglia zamuco (ayoreo, chamacoco) e altre lingue del Chaco: prime prospezioni. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa*.
- CIUCCI, Luca & Pier Marco BERTINETTO (submitted). *A diachronic view of the verb system in the Zamucoan languages*.
- COMRIE, Bernard, Lucía A. GOLLUSCIO, Hebe GONZÁLES & Alejandra VIDAL 2010. El Chaco como área lingüística. In: Zarina ESTRADA FÉRNANDEZ & Ramón ARZÁPALO MARÍN (eds.). *Estudios de lenguas amerindias 2: contribuciones al estudio de las lenguas originarias de América*. Hermosillo, Sonora (Mexico): Editorial Unison.

Syntactic mixing across generations in a Chintang/Nepali bilingual environment

Stoll, S., Zakharko, T., Moran, S., Schikowski, R. and Bickel, B.

University of Zurich

One of the key effects of sustained language contact is code-switching, i.e. the switching from one language to another either within or across utterances. Over time this can lead to language change, sometimes resulting in a mixed code with new structures. Here, we explore changes in code-switching during language acquisition and more broadly across several generations in a population with community-wide bilingualism. This kind of bilingualism is the most common one in the world, partly as the result of national languages and *linguae francae* encroaching on vernaculars worldwide. Despite this, most research on acquisition and trans-generational patterns in bilingual situations has so far focused not on community-wide but on individual bilingualism, where a single child, or a small group of children, is raised bilingually in an otherwise monolingual community.

A quantitative analysis of a trans-generational, conversational corpus of Chintang (Tibeto-Burman) speakers with community-wide bilingualism in Nepali (Indo-European) reveals that children show more code-switching into Nepali than older speakers. This confirms earlier proposals in the literature that code-switching in bilingual children decreases when they gain proficiency in their dominant language, especially in vocabulary. Contradicting expectations from other studies, our corpus data also reveal that for adults, multi-word insertions of Nepali into Chintang are just as likely to undergo full syntactic integration as single-word insertions. Speakers of younger generations show less syntactic integration. We propose that this reflects a change between generations, from strongly asymmetrical, Chintang-dominated bilingualism in older generations to more balanced bilingualism where Chintang and Nepali operate as clearly separate systems in younger generations. This change is likely to have been triggered by the increase of Nepali presence over the past few decades.

Semantic factors governing the order of object and verb in Walman, a language of Papua New Guinea

Lea Brown and Matthew Dryer

In most languages that allow both orders of object and verb, the factors conditioning the choice of order are usually pragmatic, although the length of object noun phrases is sometimes also a factor. It is very rare for the order to be determined lexically or semantically. One of these rare cases is Koyraboro Senni, where some verbs require OV order while other verbs require VO order (Heath 1999). Although the choice might be seen as lexically governed, there is a clear semantic principle governing the two classes of verbs: if the object is a patient or theme, the object precedes the verb but if the object bears some other semantic role, the object follows the verb.

In this talk we describe a similar phenomenon in Walman, a language in the Torricelli family spoken on the north coast of Papua New Guinea. As in Koyraboro Senni, there are two classes of verbs, except that in Walman one of the classes allows either OV or VO order while the other one requires VO order. The semantics associated with the two classes of verbs is similar to that in Koyraboro Senni. As in Koyraboro Senni, with verbs whose object is something other than a patient/theme, the object must follow the verb. With verbs whose object is a patient/theme, Walman is slightly different: in Koyraboro Senni, objects of such verbs must precede the verb, while in Walman they may either precede or follow the verb.

This principle actually applies not only to the order of object and verb, but to the order of any nonsubject and verb, including the complements of the verbs for 'hear' and 'put in a pile', which are morphologically intransitive and do not exhibit object inflection for the complement. In particular, the complement of the verb for 'hear' must follow the verb, since it is not a patient or theme, while the complement of the verb for 'put into a pile' can either precede or follow the verb, since it is a theme. Similarly the recipient noun phrase of the verb for 'give' (which is treated as the object in that there is object inflection for it on the verb) must follow the verb because it is neither patient nor theme, while the theme noun phrase (which is not treated as an object by object inflection on verbs) can either precede or follow the verb.

Discussions of semantic roles in the recent literature question whether semantic roles like patient and theme actually play a role in grammar. But the Walman evidence suggests they do. In addition, discussions of semantic roles in the semantic literature do not always agree on what counts as a patient or theme. For example, while linguists often assume that the object of verbs meaning 'hit' are patients, Jackendoff (1972) argues that the object of *hit* in English is a goal, since the act of hitting involves moving some object towards the thing that gets hit. In Walman, at least, word order implies that the object of the verb for 'hit' is treated as a patient/theme since it can precede the verb, whereas if it were a goal it would have to follow the verb.

The Typology of Voice in Malayic: The development of agent-demoting passives

A key typological characteristic of traditional Austronesian languages such as Seediq, Malagasy, Tagalog etc. is that these languages display a so-called 'Philippine' type voice system. While in languages with a European-type voice system, the agent in passive constructions is 'demoted' from (core) argument to adjunct (as evidenced by syntactic behavior, e.g. optionality of agent, binding and extraction patterns typical of adjuncts rather than arguments etc.), in languages with a 'Philippine' type voice system, agent arguments are retained as core arguments in transitive constructions with non-agent subjects.

A large proportion of the Austronesian family has completely lost the 'Philippine' type voice system, but Indonesian-type languages have been claimed to exhibit a 'mixed' voice system. The term 'mixed' refers to the fact that, on the one hand, these languages retain a 'Philippine' type voice construction (often termed 'object voice'), while on the other hand, they also appear to exhibit a European-type passive in which the agent argument is 'demoted' to adjunct. This passive construction is marked with the prefix *di-* in Standard Indonesian.

Although the etymological origin of the passive prefix *di-* has been widely discussed (e.g. by Wolff (2001), Adelaar (2005), van den Berg (2004) and others), little has been said about the stages through which Indonesian-type languages developed a European-type passive construction. In this paper, we present data from 'missing link' varieties: i.e. Malayic varieties in which the European-type passive construction shows the remnants of the earlier, Philippine-type voice system, illustrating potential intermediate steps in the change.

These varieties, which include some dialects of Kerinci, have already developed a European-type passive, yet they exhibit unusual properties which are best understood as the remnants of a Philippine type passive. The 'demoted' agent in the *di-*passive construction exhibits behavior/properties typical crosslinguistically of arguments rather than adjuncts. For example, in the *di-*passive construction, the agent argument is syntactically obligatory, it must appear adjacent to the verb, it cannot appear in a by-phrase, and, as in the object voice construction, the verb exhibits selectional restrictions with regard to the agent.

Through comparison of the syntactic characteristics of these 'missing link' varieties, this paper develops an account of the historical changes which lead to the development of the mixed voice system in Indonesian-type languages.

Bibliography:

- Arka, I. W., & Manning, C. (1998). *Voice and grammatical relations in Indonesian: A new perspective*. CSLI.
- Adelaar, A. (2005). Much ado about di. *Bijdragen tot de Taal-, Land-en Volkenkunde*, 127-142.
- Chung, S. (1976). On the subject of two passives in Indonesian. *Subject and topic*, 57, 99.
- McKinnon, T., Cole, P., & Hermon, G. (2011). Object agreement and 'pro-drop' in Kerinci Malay. *Language*, 87(4), 715-750.
- Van den Berg, R. (2004). Some notes on the origin of Malay di. *Bijdragen tot de Taal-, Land-en Volkenkunde*, 532-554.
- Wolff, John U. (2001) 'The origin of the Malay affixation di-nya'. Paper, Fifth International Symposium on Malay/Indonesian Linguistics, Leipzig, 16-17 June.

Antipassives in Tsezic and beyond

Bernard Comrie*, Diana Forker, Madzhid Khalilov, Zaira Khalilova & Helma van den Berg†
(*corresponding author <comrie@eva.mpg.de>)

So far antipassive constructions have been less well described than other valency-changing constructions, and a number of questions remain unanswered. The Nakh-Daghestanian language family is one of the genealogical groups with a relatively dense concentration of these constructions, and thus represents a promising source of information for their investigation. Nakh-Daghestanian languages overall favor transitivity processes for valency change. Still, a number of detransitivizing processes are attested, and antipassivization is one of them.

We follow basically the definition of antipassives given by Polinsky in WALS: “An antipassive construction is a derived detransitivized construction with a two-place predicate, related to a corresponding transitive construction whose predicate is the same lexical item. In the basic transitive construction, the patient-like argument is realized as a direct object; in the antipassive construction, that argument is either suppressed (left implicit) or realized as an oblique complement.”

In this sense, antipassives are attested for the following Daghestanian languages:

- Godoberi (Andic subgroup);
- Avar;
- Tsezic subgroup: Hunzib, Bezhta, Tsez, and Hinuq;
- Dargi

Nakh-Daghestanian antipassives differ considerably from canonical antipassive constructions. Only Dargi (obligatorily) and Bezhta (optionally) allow the patient to be expressed, in the ergative or the instrumental (less frequently, inter-essive), respectively. Dargi does not make use of overt derivational morphology. In all languages other than Dargi, the derivation is also available for intransitive verbs. The antipassive is not very productive and generally only available with a limited number of verbs, which is at least partially due to its semantics. It commonly has a durative, iterative, or habitual meaning. Therefore its use is (almost) entirely determined by its semantic load. Accessibility to major syntactic processes in Daghestanian languages is not generally constrained by grammatical relations, so the antipassive does not serve to feed syntactic processes as in, for instance, Dyirbal.

The goal of this presentation is to give an overview of antipassive constructions in Nakh-Daghestanian languages, especially within the Tsezic and Dargi sub-branches, with particular regard to:

- productivity
- semantics
- syntactic properties

Examples from Bezhta

- (a) *oždi* *bäbä* *m-üq-čä*
boy.OBL(ERG) bread(III) III-eat-PRS
‘The boy eats the bread.’
- (b) *ožö* *bäbälä-d* *Ø-üⁿq-dä-š*
boy(I) bread.OBL-INS I-eat-ANTIP-PRS
‘The boy is busy eating the bread.’

A typological overview of strategies that aid in reference tracking

Thomas J. Conners, Claudia M. Brugman, Nikki A. Adams, , Amalia E. Gnanadesikan, Brook Hefright
University of Maryland

This general overview lays out the specific linguistic phenomena underpinning an ongoing collaborative research project examining the distribution and interpretation of systems of people referring expressions cross-linguistically. We assert that no theory of reference resolution should, from a typological perspective, have a major or exclusive focus on the distribution and interpretation of pronouns, or to limit its further scope to the distribution or interpretation of anaphoric zeroes with respect to overt pronouns.

In addition to categories of people referring expression with independent denotational value, we consider here both exophora—the use through language to refer to an element in the speech situation—and, more particularly, anaphora—the co-reference of one linguistic element with another. In the languages in our survey (including Chinese, Dhivehi, Indonesian, and Korean), we distinguish two large classes of phenomena to be accounted for as the non-categorical members of the set of potential people referring expressions:

Pronoun “imposters” (viz. Collins & Postal 2012)—elements whose unambiguous reference in the linguistic context does not match its denotational (lexical) features, such as

- the use of lexical nouns (such as kin terms and proper names) with first- or second-person reference, and as used anaphorically with third-person reference
- the use of pronouns whose lexical features of person, number, or gender do not match the corresponding properties of the referent
- the use of indexicals whose lexical animacy features do not match the corresponding properties of the referent

Zeros—phonologically null elements with recoverable reference in the discourse context

Both of these classes of phenomena have been the subject of much scholarship. However, the languages of our survey show a wider range of acceptable linguistic contexts for their appearance than is often recognized in the literature. We give an overview of some of the discourse-pragmatic conditions on their use, as well as interactional pragmatic considerations (such as gender taboo prompting a gender imposter pronoun).

Orthogonal to these strategies are a number of lexicogrammatical systems that may interact with, in particular, zero anaphors and imposters to constrain the interpretation of the referent of the zero or other anaphoric element. Two of these are honorific and speech level systems. In each case, lexical choices are made that reflect the speaker’s relationship to the hearer, the speaker’s relationship to an individual spoken of, or the hearer’s relationship to an individual spoken of. In a language that uses speech levels, the lexical choices are not necessarily involved in reference to individuals; they may be, e.g., verbs that index the speech level, thereby indexing the social status of an individual spoken of (for instance, the actor of the named action), and thereby narrowing the potential referential space of that individual (for instance, the zero anaphor that expresses the subject of the verb).

Complementary to the introductory framing of the pragmatic aspects of reference resolution, this overview will be heavy on descriptive data, with a view toward broadening the problem space for reference resolution to give greater attention both to the phenomena implicated in this communicative need, and the sociopragmatic values that may be reflected in language-particular ways.

Collins, Chris and Paul Postal. 2012. *Imposters: a study of pronominal agreement*. MIT Press: Cambridge.

Two systems or one? A Canonical Typology approach

Greville G. Corbett and Sebastian Fedden
Surrey Morphology Group

Depending on the research tradition, there are types of data which are handled as representing a single (complex) system or elsewhere as two related systems. The particular analysis is often assumed rather than argued for. We need a more principled approach here, since we are engaged in a typological study of gender and classifiers, for which a prerequisite is a means of determining the number of systems involved in a given language.

Let us take an example which is relatively straightforward. We analyse Italian as having gender and number, despite the fact that these features are realized together. This can be seen in the following adjective paradigm.

<i>nuovo</i> 'new'	SINGULAR	PLURAL
MASCULINE	nuov-o	nuov-i
FEMININE	nuov-a	nuov-e

Why does it seem evident that our analysis should indeed have two features, each with two values, rather than a single feature with four values? **First**, gender and number in Italian are separate systems because they are orthogonal to each other. Given a noun which takes masculine agreement we cannot predict its number value. Similarly for a feminine. And equally, given a number value we cannot predict gender. The **second** argument concerns form. While our example shows cumulation of number and gender, there are number forms (of synthetic verbs) which show number without gender.

These observations suggest two parameters for our typology: (i) the degree to which the semantics of the two (potential) systems are orthogonal to each other, and (ii) the degree to which their means of realization are distinct. In canonical terms, the more distinct the semantics and the more distinct the forms, the more clearly different the two systems are. We give a typology of nine possibilities, from a canonical single system to a situation with canonically two systems. We then apply this to a small number of languages (from a larger project set). Each of these is traditionally analysed as having a gender system and at least one system of classifiers. While in some instances we can accept the tradition, in others the degree of semantic and formal overlap between the systems is so extensive that we would want to speak of a single system.

We conclude that the method provides the means to be clear and explicit about analyses into one or more than one system. This is the necessary basis for a comprehensive typology of gender and classifiers, and in particular of languages which have been argued to have both systems concurrently.

The Lenguas de Bolivia project: typological prospects

Mily Crevels, Leiden University
Pieter Muysken, Radboud University

The knowledge available on the languages of Bolivia has been gathered and published in a series of four volumes, *Lenguas de Bolivia*. Most of the languages documented are presented in sketches of 30-50 pages, systematically covering their phonology, morphology, and syntax (31 in total). In addition, ethno-historical and sociolinguistic information is provided, and each chapter concludes with one or more texts. We decided to publish the series in Spanish and in Bolivia to make the data available to the general Bolivian public and we have tried to be consistent in our terminology and glosses in such a way that a maximum number of readers would be able to profit from the books. The fourth volume contains elaborate subject, language, and author indexes and a terminology database with definitions and relevant examples from the different languages in the chapters. Overall, the series consists of ca. 1.800 pages, and testifies to the documentation efforts which have intensified considerably in the last two decades.

However, so far the potential for typological and areal research of our project has been underexplored. In this talk we will illustrate some striking typological features (e.g. possession, classifiers, alignment, switch reference, case marking, evidentials, causatives, reduplication, nasals) and their distribution across the languages of the highlands, the foothills, the Amazonian lowlands, and the Chaco. Some features have spread across different zones, while others are limited to one of the zones involved. In any case, a direct highland-lowland split, as postulated by several earlier authors, is not generally valid for the feature set that we have studied. Interestingly, Bolivia appears to lie at the cross-roads of the major linguistic areas of South America: the Andes, the Amazon proper, and the Chaco-Southern Cone. Its languages thus present a microcosm of South American linguistic diversity.

Synchronic vs. diachronic explanations of typological universals: redefining the role of frequency

Frequency has long been recognized to affect linguistic structure to the extent that high-frequency items undergo phonetic reduction at a faster rate than low frequency ones, and are represented and accessed as single units. These effects are plausibly related to cognitive automatization and consequent reduction of articulatory gestures (Mithun 2003, Bybee 2007).

In typology, frequency has been invoked in a different sense: speakers selectively use overt marking only to encode less frequent meanings, because more frequent meanings are easier to identify and hence less in need to be indicated overtly. This idea underlies, for example, several classical explanations of universals pertaining to number marking, possession marking, and alignment (Greenberg 1966, Comrie 1989, Croft 2003, Haspelmath 2006 and 2008, among others). These explanations are based on the synchronic distribution of zero vs. overt marking within these domains, not how this distribution actually originates in individual languages. The paper examines several types of cross-linguistic diachronic evidence about the origins of zero vs. overt marking for different number values, possession types, and argument roles (as found in different alignment types), and argues that this evidence poses various challenges for frequency explanations.

First, overt markers often originate from the reinterpretation of pre-existing elements (through grammaticalization or other processes of form-meaning redistribution within complex expressions). In such cases, the markers are used in contexts compatible with the properties of the source construction and are not used in contexts incompatible with these properties, regardless of whether the various contexts involve more frequent or less frequent meanings. If all of the relevant meanings were originally zero marked, then those incompatible with the properties of the construction that give rise to the overt marker will remain zero marked. This suggests that the distribution of zero vs. overt marking depends on the properties of the sources of individual overt markers, not the frequency of the meanings encoded by these markers.

Also, when overt markers develop from preexisting elements, this process is plausibly driven by context-induced inferences, so there is no obvious evidence that it is also driven by the frequency of the meanings encoded by the resulting markers. Likewise, sometimes the use of zero and overt marking for different meanings within the same grammatical domain (e.g. different number values) originates from a situation where all of these meanings were overtly marked and regular phonetic changes eliminated some of the markers. These changes too are independent of frequency, and, cross-linguistically, they affect both markers encoding more frequent meanings and ones encoding less frequent meanings.

These facts suggest that the distribution of zero vs. overt marking is a result of several particularized and often context-dependent diachronic processes, rather than the frequency of the meanings encoded by individual markers. In order to obtain a full understanding of this distribution, then, we need qualitative and quantitative data about these processes (e.g. what source constructions can give rise to specific overt markers, in what contexts, and through what mechanisms), rather than data about the resulting patterns in themselves.

References

- Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Comrie, B. (1989). *Language universals and linguistic typology. 2nd edition*. Oxford: Basil Blackwell.
- Croft, W. (2003). *Typology and universals. 2nd edition*. Cambridge: Cambridge University Press.
- Greenberg, J. H. (1966). *Language universals, with particular reference to feature hierarchies*. The Hague: Mouton.
- Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics* 42, 25–70.
- Haspelmath, M. (2008). Creating economical morphosyntactic patterns in language change. In J. Good (Ed.), *Linguistic Universals and Language Change*, pp. 185–214. Oxford: Oxford University Press.
- Mithun, M. (2003). Functional Perspectives on Syntactic Change. In R. D. Janda and B. D. Joseph (Eds.), *The handbook of historical linguistics*, pp. 552–72. Oxford: Blackwell.

Induction to the max: Typology through parallel texts

Michael Cysouw

Universität Marburg

One of the many fortuitous and productive coincidences that resulted from the open and collaborative atmosphere at the linguistics department of the MPI-EVA was that my stint there coincided with the stay of Bernard Wälchli while he was working on verb semantics in bible translations. The possibilities of using parallel texts for the typological comparison of the world's languages has haunted me ever since, resulting recently in a project to finally prepare bible translations in a linguistically and technologically useable form. The current state of the project is available at <http://paralleltxt.info/>. I will present the project and its achievements so far, the computational potential of the data, and, most importantly, the fruitful possibilities to induce typological surveys from this source.

Patterns of Variation in Word Order and Agreement in Modern Western Iranian Languages

Western Iranian languages are generally classified into two groups: Southwestern and Northwestern. In this paper, I will adopt a typological perspective to describe the diversities which I have found in my corpus of a number of Modern Southwestern and Northwestern Iranian languages with respect to word order and agreement parameters. The Southwestern Iranian languages for which I will present data and statistics are Delvari, Larestani, Davani, and Modern Persian. From the Northwestern group, I will describe Hawrami, Central and Southern Kurdish, Laki, Balochi, Tati, Northern Talyshi, Vafsi, Shahmirzadi, and Naini. The patterns of the word order variations which I have discovered are based on the application of the twenty four correlation pairs, discussed in Dryer (1992, 2007, 2011, and 2013), to the mentioned languages. The patterns are distributed along a continuum which begins with Central Kurdish, Laki, and Balochi which contain more VO-type correlation pairs and ends with Shahmirzadi and Northern Talyshi which reveal more OV-type correlation pairs. Interestingly, these languages, which are at the two ends of the mentioned continuum, all belong to the Northwestern group of the Iranian languages. Also noteworthy is the fact that all Modern Iranian languages have the SOV constituent order at clause level. Finally, I will propose that a mixed type in terms of word order parameters may last for centuries. In regard to agreement, Delvari, Larestani, Davani, Laki, (most varieties of) Balochi, Central Kurdish, Naini, Vafsi, and Northern Talyshi have grammaticalized split-agreement to encode S, A, and P whereas Southern Kurdish, Modern Persian, Shahmirzadi, and Sistani Balochi are uniformly Nominative-Accusative. The Iranian split-agreement systems are tense-and-transitivity-sensitive. I will adopt the standard treatments of split-systems which are proposed in Comrie (1978) and Dixon (1994) and will suggest modifications based on my corpus of the Iranian languages. More specifically, I will argue that in the past tense domain of transitive verbs of the Iranian languages with a split-agreement system, we find the following patterns: Ergative-Absolutive, Tripartite, Bipartite (namely Oblique-Oblique) or even a combination of them in one language.

References

- Comrie, B. (1978). Ergativity. In W. Lehmann (Ed.), *Syntactic Typology: Studies in the phenomenology of language* (pp. 329-394). Sussex, The Harvester Press.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Dryer, M. (1992). The Greenbergian Word Order Correlations. *Language*, 68, 81-138.
- Dryer, M. (2007). Word Order. In T. Shopen (Ed.), *Language Typology and Syntactic Description*, (2nd ed.) vol. 1: *Clause Structure*, Cambridge: Cambridge University Press, (pp 61-130).
- Dryer, M. (2011). The Evidence for Word Order Correlations, *Linguistic Typology*, 15, 335-380.
- Dryer, M. (2013). On the Six-Way Word Order Typology, Again. *Studies in Language*, 37(2), 267-301.

How WEIRD are WALS languages?

Östen Dahl

Stockholm University

Henrich et al. (2010, 61) note that behavioural scientists tend to make "broad claims about human psychology and behavior" based on samples from "Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies" at the same time as, in their opinion, these societies "are among the least representative populations one could find for generalizing about humans." Majid and Levinson (2010, 103) say that "WEIRD languages have misled us, too", arguing that linguists have "projected assumptions based on English and familiar languages onto the rest". It is certainly possible to agree with this statement. I think, however, that we may also be led a bit astray by the catchy acronym WEIRD in that the adjectives it encapsulates are not necessarily the most adequate for characterizing the biases that have influenced linguistics. It is true that Western (mainly European) languages have been in the focus for a long time; however, even after the Eurocentric bias has started to lose its grip on the choice of languages to be studied, there remains a bias that can be summed up in the acronym "LOL" for "Literate, Official, and with Lots of users". Even in typological works, the bias is visible. To a certain extent, it is probably unavoidable, given the restricted availability of information on smaller languages. The interesting question is how much the bias in the choice of languages influences the results. It turns out that there is a very restricted set of LOL languages which are overrepresented in almost any sample. It can also be shown that these languages as a group have a distinct typological profile. Recently, it has been argued by a number of scholars that certain typological parameters are correlated to factors such as the size of a language in terms of number of speaker and its degree of contact with other languages. This is most obvious in parameters that have to do with morphological complexity but also in other areas, such as word order and argument alignment. The distinct profile of the LOL languages is partly due to their size but also to the overrepresentation of European languages that it still found among them. Looking at the languages that occur in WALS, it can be noted that the basic 100-sample contains at least 20 languages that must be characterized as LOL languages in all respects. This is partly due to a conscious policy – the editors of WALS decided to add a few more "major languages of Eurasia" than would be motivated from the point of view of a sample that would be maximally free of genealogical and areal bias. On the other hand, most maps in WALS contain large numbers of languages in addition to the basic 100-sample, and they are much less obviously biased towards LOL languages. An analysis of how much the statistics in WALS is influenced by the bias in the basic 100-sample shows that this (perhaps not unexpectedly) depends on the size of the samples used in the individual chapters. Literacy, political status, and number of speakers are factors that are not mentioned in the introduction to WALS; they are rarely seen as relevant for typological sampling. To some extent you get a representative spread with respect to these parameters for free if you try to optimize genealogical and areal diversity, since this will automatically include both large and small language families. But this presupposes that the sample is large enough. For WALS, it can be shown that whereas the largest samples are rather biased towards small languages, some of the results reported in chapters with small samples give misleading impressions of the frequency of certain features which can be blamed on the overrepresentation of LOL languages.

References

- Henrich, Joseph, Heine, Steven J., & Norenzayan, Ara (2010). The weirdest people in the world? *Behavioral and Brain Sciences* 33, 61-83. doi: 10.1017/S0140525X0999152X
- Majid, Asifa, & Levinson, Stephen C. (2010). WEIRD languages have misled us, too [Comment on Henrich et al.]. *Behavioral and Brain Sciences*, 33(2-3), 103. doi:10.1017/S0140525X1000018X.

Agreement of personal pronouns in Nakh, Dargwa and Archi (East Caucasian)

Michael Daniel

Most East Caucasian languages are rich in **agreement classes**, alias **genders**. Agreement shows different configurations in the singular and in the plural. For example, Archi (Lezgif) and Chechen and Ingush (Nakh) have four formally different agreement patterns in the singular, and Dargwa languages have three. In the plural, however, only two agreement patterns are present in Archi and Dargwa, and three in Nakh. (Note that while this produce only three genders in Dargwa and four genders in Archi, there is a complex pairing between singular and plural agreement patterns in Nakh, so that Nichols 2011 counts up to six genders). The pivotal distinction in the plural is that between human and non-human plurals (HPL vs. NPL), and this is also true of other East Caucasian languages.

On the contrary, apart from some special patterns surveyed in (Sumbatova 2011) and morphologization of personal pronouns in Udi and Tabasaran (Lezgif) and Bats (Nakh), East Caucasian is remarkable in how it **lacks personal agreement**. Against this family background, Archi, Dargwa and Nakh stand out, having a rather peculiar agreement pattern. While other languages of the family show HPL agreement with plural personal pronouns, plural personal pronouns ‘we’ and ‘you.pl’ (and, where available, ‘incl’) control **non-human rather than human** plural agreement markers. I will refer to this as **plural locutive agreement**.

This pattern has been discussed extensively in connection with Archi (Kibrik 1972); (Chumakina, Corbett & Kibort 2007) and (Corbett & Baerman 2013) whose focused is also on Archi, mention comparable data from Nakh and Dargwa, too. This is a rather uncommon strategy of personal agreement. Why should this happen to Archi, Dargwa and Nakh? The three solutions suggested to this problem are as follows:

- A. plural personal pronouns do not control a **special type** of agreement (personal agreement) but constitute a **separate gender** (agreement class) (Kibrik 1972, 1977, 1994)
- B. agreement with plural personal pronouns is **indeed personal agreement** using – somewhat parasitically – gender morphosyntax (Chumakina et. al 2007)
- C. the agreement of plural personal pronouns is **set to default value**; this happens because of the conflict between language specific requirement that pronouns, as all NPs, must control gender agreement, on the one hand, and typological incompatibility of gender with personal pronouns, on the other (Corbett & Baerman 2013)

This paper is an attempt to revise the interpretations suggested in the above discussion. Although I do not fully support the argument of B (based on gender resolution rules), there is empirical data that shows that A indeed can not be valid. My conclusion is thus similar to that of B but made on different grounds. Both A and B are purely structural account, though, while C is a typological claim. It is on C that this paper focuses. Although Corbett and Baerman’s typological argument is rather appealing, there are several conceptual problems. The most important one is as follows: if incompatibility of gender with personal pronouns is strong enough to set the agreement to default in Archi, Dargwa and Nakh, why the same does not also happen in other languages of the family: Lak, Avar, Andic, Tsezic etc. – languages whose plural personal pronouns do control HPL?

Note that the pattern can hardly be explained away as a spread of innovation or inherited retention: the three linguistic areas are not contiguous, and Nakh and Dargwa have never been considered to be especially close in the family tree. As for Archi, it is the only Lezgif language that shows plural locutive agreement, and is not in contact with either Nakh or Dargwa; it is thus plausible to suggest that this may be a language internal innovation.

I believe that any account of the phenomenon of plural locutive agreement as attested in East Caucasian should address the following issue: why should three languages in one family develop, apparently independently, a pattern which is typologically so rare? (Corbett & Baerman 2013 do suggest a typological parallel from Amazonia, but this parallel is, to my eyes, a partial one – at least under certain aspects). I suggest that to solve the paradox is to seek an explanation combining a *universal tendency* attested in various languages of the world with a *structural feature* specific to the languages of the family. If it is shown to be plausible that a cross-linguistic tendency interacting with a family structural feature may give rise to an unusual pattern, this is a viable explanation of a situation like the one described above.

I suggest that, in this case, *universal tendency* is probably negative politeness, while *family structural feature* is absence of personal agreement and argue that plural personal pronouns originally assumed an agreement pattern typical of **impersonal** (non-specific human) reference as a means of indirect reference to the locutor(s).

In Nakh and Dargwa, NPL agreement pattern is one of the means of non-specific human reference; in Archi it is indeed the main one. Indirect reference as a strategy of (negative) politeness is attested cross-linguistically (see Siewerska 2004). In some languages, third person constructions are known to develop into regular first (French *on*) or second (Portuguese *você(s)*) person reference device (in these two cases – notably also in the plural). However, examples usually come from person-prominent languages, where the opposition to be switched is that of person: from first/second to third person. In East Caucasian languages, on the other hand, there is no person agreement, at least not in the usual sense, and indirect reference as a means of politeness is achieved by switching gender agreement to NPL, the pattern shared by non-humans with human non-specific NPs. In this way, negative politeness in combination with human vs. non-human rather than person agreement gave rise to the rare phenomenon of plural locutive agreement.

References:

- Chumakina, Marina, Anna Kibort & Greville Corbett (2007) 'Determining a language's feature inventory: person in Archi' In: Austin, Peter K. & Andrew Simpson (eds) *Endangered Languages*. Special Issue of *Linguistische Berichte* 14. Hamburg: Helmut Buske. 143-172.
- Corbett, Greville & Matthew Baerman. 2013. Person by other means. *Languages across Boundaries: Studies in Memory of Anna Siewierska*, ed. by Dik Bakker and Martin Haspelmath. Berlin: MdG.
- Kibrik, Aleksandr E. 1972. 'O formalnom vydelenii soglasovatelnyx klassov v archinskom jazyke'. *Voprosy jazykoznanija* 1. 124-131.
- Kibrik, Aleksandr E. 1994. 'Archi'. In Smeets, Rieks (ed.), *The Indigenous Languages of the Caucasus* 4, 297-365. Delmar, New York: Caravan Books.
- Nichols, Johanna. *Ingush Grammar*. University of California Press: 2011.
- Siewerska, Anna. *Person*. Cambridge. CUP: 2004.
- Sumbatova, Nina. 'Person hierarchies and the problem of person marker origin in Dargwa: facts and diachronic problems.' - *Tense, aspect, modality and finiteness in East Caucasian languages* / Ed. by G. Authier and T. Maisak. Vol. 30. Bochum : Brockmeyer, 2011. P. 95-130.

Relative clause formation in Lule (Argentine Chaco)

Lule was spoken by a semi-nomadic people who lived in the Gran Chaco between the Pilcomayo river and the Andean foothills of northwestern Argentina. It is an extinct language, which was also spoken by other groups of the present-day Argentine provinces of Tucumán and Santiago del Estero. Virtually all of our documentation of Lule comes from the Sardinian Jesuit missionary Antonio Maccioni (1688-1753) in his *Arte y vocabulario de la lengua lule y tonocoté*, published under the name Machoni in Madrid in 1732. Mentions of Lule disappear from the historical record after the expulsion of the Jesuits in 1776. We may surmise that the Lule language lost its last speakers sometime in the course of the early 19th century. The Lule language is generally considered to be genealogically related to the Vilela language, forming a Lule-Vilela family, but it remains possible that the lexical similarities are due to contact rather than to shared ancestry. The language described in Machoni (1732) provides interesting challenges for the study of linguistic diversity in the Chaco area, because it is typologically rather distinct from the other known Chaco languages, and because we have to rely almost entirely on Maccioni's examples and linguistic abilities in trying to reconstruct its grammar in modern terms. In this paper, we will interpret Maccioni's account of relative clause formation. There are 18 examples of what can be interpreted as relative clauses in Machoni's (1732) grammar section, and 33 examples in his texts, i.e. the doctrine and catechism sections. We suggest that it is possible to draw tentative conclusions from this very small corpus. The most common type of construction involves subject relative clauses with the agent nominalizer *-ton* as in examples (1) and (4), and this is the only type present in the texts. Another type involves no nominalization, but the anaphoric and resumptive pronoun *mima*, as in example (2), and which can be optional, as in example (3). It would appear from examples (2) and (3) that *mima* is only used when the head is an object, but example (4) (the only occurrence in the corpus of a relative clause construction involving both *-ton* and *mima*), shows that this is not the case. We suggest that the presence of *mima* is pragmatically conditioned, and that *mima* is likely to occur only when the subject of the relative clause is not coreferent with the first word of the sentence. We will conclude with some observations regarding the validity of philological interpretations of early missionary writings of extinct languages.

- (1) *pele inle wotikeyu-ton wetsi-p*
 man yesterday work-NMLZ die-3SBJ
 'the man who worked yesterday is dead' (Machoni 1732, grammar section, p. 56)
- (2) *tanta kis kai-ts mima p'o-p*
 bread 1PRO eat-1SBJ ANA.DIST be.white-3SBJ
 'the bread that I am eating, (that one) is white' (Machoni 1732, grammar section, p. 65)
- (3) *pele Dios alap-ip (mima) tso=ta ka-n-t*
 man God choose-3SBJ (ANA.DIST) heaven=LOC go-FUT-3SBJ
 'the man that God chooses, ((that one)) will go to heaven' (Machoni 1732, grammar section, p. 64)
- (4) *Dios amaitsi-ton mima ti-t usa Dios ne-kekes-p*
 God love-NMLZ ANA.DIST do-3SBJ what God INS-command-3SBJ
 'the one who loves God, (that one) observes what God commands' (Machoni 1732, grammar section, p. 64)

(Note: the abbreviations used follow the Leipzig Glossing Rules, with the addition of ANA 'anaphoric', and PRO, 'independent pronoun'.)

Role recovery in subject and object relative clauses

Holger Diessel

Relative clauses have been studied intensively in syntax (e.g. Sag 2010), psycholinguistics (e.g. Gordon and Lowder 2012), and typology (e.g. Comrie 1989, 1998, 2003, 2006; Comrie and Kuteva 2005; Givón 1990; Keenan and Comrie 1977, 1979; Lehmann 1984, 1986). One aspect that has played an important role in this research is the encoding and processing of the relativized role, i.e. the semanto-syntactic role of the ‘head’ in the relative clause. Research in linguistic typology has been mainly concerned with the differential use of pronouns, i.e. relative pronouns and resumptive pronouns, to indicate the relativized role; but pronouns are primarily used to mark some of the lower syntactic roles, e.g. the objects of adpositions and genitives attributes. The core roles, i.e. subject and object (or S, A, P), are only rarely marked by a case-marked pronoun. This raises the interesting question how do languages differentiate between subject and object relative clauses if the relativized role is not indicated by a case-marked pronoun (or pronoun in situ) as in the vast majority of the world’s languages?

In the syntactic and typological literature, it is commonly assumed that in the absence of a pronoun the relativized role is expressed by a ‘gap’; but the so-called ‘gap-strategy’ subsumes a wide range of very different linguistic means to express the relativized role, or as Comrie (1989: 152) put it 25 years ago: “In practice, a number of strategies can be invoked, ranging from strategies based on syntactic properties of the language in question to knowledge of real-world properties.” While some of these properties are well-known from the analysis of relative clauses in particular languages, they have never been systematically investigated from a cross-linguistic perspective (but see Givón 1990).

It is the purpose of this paper to analyze the ‘role recoverability problem’ from a novel perspective. The analysis concentrates on subject and object relative clauses, which often account for more than 90 percent of all relative clauses in spoken and written discourse (e.g. Roland et al. 2007; Wiechmann 2009; Wu et al. 2014). Building on research from psycholinguistics, I argue that the traditional analysis of the ‘gap-strategy’ ought to be replaced by a ‘cue-based approach’ in which the relativized role is invoked by the interaction between different types of cues (or features) that can be described from various perspectives. Cues can be (i) structural or non-structural, (ii) obligatory or optional, (iii) categorical or probabilistic, and (iv) narrow or broad. Since cues are based on specific aspects of relative clauses (e.g. particular morphemes, linear order, or semantic features), it is not unusual for relative clauses to include multiple cues for the relativized role.

Analyzing data from a stratified sample of 100 languages, the paper provides an overview and classification of the various types of cues that contribute to the interpretation of subject and object relative clauses. The proposed analysis of cues involves three important dimensions: First, I distinguish structural cues (e.g. case-marked pronouns) from semantic/pragmatic cues (e.g. the animacy of the head or the topicality of the relative-clause subject) (cf. Mak et al. 2008). Second, I distinguish special relativization cues that are specifically used in relative clauses to mark a particular role (e.g. a relative pronoun) from general structural cues that can help to constrain the interpretation of the relative clause (e.g. regular cross-reference marking on the verb). And finally, I distinguish nominal cues (based on pronouns, determiners, and nouns) from verbal cues (based on verbs and auxiliaries) from other structural cues (e.g. word order, particles, floating case markers).

The primary goal of the paper is to outline a new descriptive framework for the analysis of the role recoverability problem (inspired by research from psycholinguistics); however, the study also presents some preliminary results concerning the frequency and cross-linguistic distribution of particular structural cues and their relationships to other structural aspects of relative clauses (e.g. the order of relative clause and noun and the degree of nominalization or deranking).

Pragmatic typology: unity and diversity in systems of language use

Mark Dingemans^a, Simeon Floyd^a & N. J. Enfield^b

^a Language & Cognition Department, Max Planck Institute for Psycholinguistics

^b Department of Linguistics, University of Sydney

Typology, the comparative study of linguistic systems, has long been concerned with aspects of language structure at sentence level and below: from information structure and morphosyntax to lexicon and phonology. In what is almost a parallel universe, sociologists following in the footsteps of Garfinkel, Goffman and Sacks have developed robust descriptions of structures beyond the sentence: systems of language use, or normative regularities in how conversations are structured. The fields of diversity linguistics (encompassing a wide range of formal and functional approaches to cross-linguistic comparison) and interaction studies (including conversation analysis, discourse analysis and corpus linguistics) have much to offer each other (Ford, Fox, and Thompson 2002; Enfield 2013; Dingemans and Floyd 2014).

In this talk we trace recent developments in this interdisciplinary area and present results from two large-scale collaborative projects in pragmatic typology: the comparative study of systems of language use and the principles that shape them. A first project focuses on *repair*: the techniques that languages offer for the real-time resolution of communicative problems in social interaction. We make a systematic comparison of conversations in 12 languages of 8 distinct families, and find evidence for a universal repair system with detailed common properties, but also language-specific calibration of construction types (Dingemans and Enfield in press). A second project looks at *recruitments*, or requests: the ways in which people get other people to do things for or with them. Here, too, we find a combination of unity and diversity in social action and linguistic formatting (Floyd et al. 2014).

After presenting the findings we describe our methodological approach and the conceptual framework underlying it in order to sketch prospects and challenges for the future of pragmatic typology. A key move is to see formats for implementing actions in conversation as ultimately language-specific. Comparison is best done at the level of their constitutive properties, which can be functional (concerning aspects of linguistic formatting) as well as sequential (concerning aspects of the immediate interactional context). This approach can capture patterns of commonality and diversity in conversational structures within and across languages.

Dingemans, Mark, and N. J. Enfield. in press. "Other-initiated repair across languages: towards a typology of conversational structures." *Open Linguistics*.

Dingemans, Mark, and Simeon Floyd. 2014. Conversation across cultures. In *Cambridge Handbook of Linguistic Anthropology*, ed by. N. J. Enfield, Paul Kockelman, and Jack Sidnell, 434–464. Cambridge: Cambridge University Press.

Enfield, N. J. 2013. *Relationship Thinking: Agency, Enchrony, and Human Sociality*. Oxford: Oxford University Press.

Floyd, Simeon, Giovanni Rossi, N. J. Enfield, Julija Baranova, Joe Blythe, Mark Dingemans, Kobin Kendrick, and Jörg Zinken. 2014. Recruitments across languages: A systematic comparison. Talk presented at the 4th International Conference on Conversation Analysis, University of California at Los Angeles, CA.

Ford, Cecilia E., Barbara A. Fox, and Sandra A. Thompson. 2002. *The language of turn and sequence*. New York: Oxford University Press.

Nina Dobrushina
Higher School of Economics
nina.dobrushina@gmail.com

Negation in the complement clauses of fear-verbs

According to Michael Noonan [Noonan 1985: 119], ‘fear’ complements are peculiar in that “the languages differ in the assignment of negation to such complements”.

Indeed, some languages use positive polarity in the complement clause of the verb ‘to fear’:

Archi (Nakh-Daghestanian)

Zon l'inč'a-r, un ha tər-čaj də-χəl-kut.
I fear-IPFV you river-ERG 2-bring.away-APPR
'I am afraid you may be taken away by the stream.'

In the same type of complements, some other languages use negative polarity without apparent semantic motivation:

Latin (Sobolevskij 1998)

Time-o, ne pater veni-a-t
fear-1SG NEG.CONJ father come-CONJ-3SG
'I am afraid that father may come.'

Some languages alternate between constructions with and without negation: *Je crain-s qu'elle (ne) vienne...*

In this paper, I consider several cases of positive and negative marking of ‘fear’-complements and suggest a functional explanation of this phenomenon. I argue that negative polarity is typical of those complement clauses where the predicate of the complement is a volitional form, patterning with ‘X fears; may he not P’, such as prohibitive in Hunzib (van der Berg 1995: 131), and optative in Kumyk:

Hunzib

Kid hič'e ož-di-l žu hehe-dor
girl fear.PRES boy-OBL-ERG self hit-PROH
'The girl fears that the boy should see her.'

Kumyk

Ol bir zat ojlaš-yan, tek qorq-a
he/she one thing think-PST but be.afraid-PRS
bašxa zat gošul-up qal-ma-yađ edi
other thing interfere-CVB remain-NEG-OPT AUX
'S/he's got a plan but s/he's afraid lest something will interfere with fulfilling it.'

Positive polarity is associated with subordinate verb forms expressing possibility:

Armenian (p.c. Victoria Khurshudian)

Na čoč-v-um ēr ev es vaχen-um
s/he stagger-MED-IPFV.CVB be.AUX.PST.3SG and I fear-IPFV.CVB

ēi or inz ēl karov ē khch-el
be.AUX.PST.1SG that I.DAT also might be.AUX.PRS.3SG make.fall-INF
'He was staggering so much, I was afraid that he may make me fall'.

TAM categories and information structure in Komnzo, a language of Papua New Guinea

CHRISTIAN DÖHLER - ANU, CANBERRA

This paper looks at the TAM categories from the perspective of information structure. I will present text excerpts from my documentation project of Komnzo, a member of the Yam language family, spoken in the Morehead district in Southern New Guinea.

The method of text analysis allows us to connect the complex TAM system to pragmatic notions of backgrounding/foregrounding, event-sequencing, tail-head linkage. Based on the observations from texts, I will construct a hierarchy of most-foregrounding to most-backgrounding TAM category in Komnzo.

Languages, ‘Languoids’, ISO-codes and the Glottolog: Creating Reference Systems for Language Diversity and Variation

Sebastian Drude, CLARIN ERIC & Max-Planck-Institute for Psycholinguistics

There is a growing need to uniquely and unequivocally refer to languages and similar entities – not only by linguists concerned with language diversity (descriptive, theoretical, historical linguists and of course typologists), but also by institutions such as repositories offering language materials and by institutions and sites such as OLAC, LinguistList, WALs, Wikipedia, UNICODE and others. Indeed, the identification of the language some text or other resource is using, such as localization of user interfaces and language-specific services is a necessary requirement for many software and hardware manufacturers and the world’s communication services.

So far, ISO 639 has been answering this need in many domains – due to its wide adoption in any computer, mobile device, XML or HTML-document, even in washing machines with multi-language-display, this may very well be the most used ISO standard of all. Some 10 years ago, as the need also to be able to identify smaller and lesser widespread languages grow within and outside academia, ISO 639-3 was established based on SIL’s Ethnologue (Lewis et al. 2014), and despite several inaccuracies and problems (cf. Hammarstöm in press) it was widely taken up not only in linguistics, in lack of a better alternative. It is foreseeable that in the future, when we will more and more often interact orally with technology, there will also be a growing need for identification of language varieties, and that holds also for linguistics where linguistic diversity and variation below the language level is becoming more prominent.

But is ISO 639 the right answer to these needs? Recent discussions have brought up doubts and criticisms, and it is obvious that the first proposal for ISO 639-6, covering linguistic varieties by a flat hierarchical system of four-letter-codes, is clearly inadequate. There are several alternatives to the Ethnologue to catalogue the world’s linguistic diversity, for instance the UNESCO atlas of endangered languages, specialized and general encyclopedias, endangeredlanguages.com, and in

particular the Glottolog, a child of the MPI-EVA, which also offers and uses the new terminological conception around the term ‘languoid’.

Based on the author’s background in language documentation, as former head of one of the world’s leading language archives as well as on his participation in ISO TC 37 SC 2 WG 1 (responsible for ISO 639), this contribution offers an insider’s view on the discussion about ISO 639 and related work dealing with linguistic diversity, focusing on the following points:

- Many of the criticisms in Morey et.al’s paper and similar recent notes are based on misconceptions of the role of ISO and Ethnologue.
- Even if arguably best combined (or used alongside) with other catalogues such as Glottolog, ISO 639 is too important for diversity linguists not to care about and not to get more involved.
- The term ‘languoid’ may well be a useful cover term for languages, language varieties and perhaps language families, but the theoretical conception offered by Cysow and Good (2013) to underpin this term (and the related notions ‘doculect’ and ‘glossonym’) is inadequate as a conception to formalize the notion of ‘language’.
- An appropriate way of cataloguing and coding languoids cannot be done without an appropriate answer to the question what names like ‘English’ or codes like ‘eng’ actually refer to. Here, a conception of languages and similar entities as sets, allowing for fuzzy and overlapping borders, will be offered.
- The topology of ‘languoids’ is empirically more complex than what can be covered by the traditional triad ‘family’ – ‘language’ – ‘dialect’ (or even more general: ‘variety’), but that means we should expand and refine our model, not abandon these terms.
- A more appropriate framework for identifying and coding linguistic variation than the previous ISO 639-6 proposal will be outlined.

References

Cysouw, Michael & Jeff Good (2013). Languoid, Doculect, Glossonym: Formalizing the notion "language". *Language Documentation and Conservation* 7. 331-359. <http://hdl.handle.net/10125/4606>

See also a Blog contribution:

Cysouw, Michael (2014-01-04). Languoid, Doculect and Glossonym: Formalizing the Notion 'Language'. *Diversity Linguistics Comments*. <http://dlc.hypotheses.org/623>

Hammarström, Harald (in press). Review article: The *Ethnologue* 16th & 17th eds. *Language: Journal of the LSA*.

Haspelmath, Martin (2013-12-04). Can language identity be standardized? On Morey et al.'s critique of ISO 639-3. *Diversity Linguistics Comments*. <http://dlc.hypotheses.org/610>

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the World*, Seventeenth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.

Morey, Stephen & Post, Mark W. & Friedman, Victor A. (2013). The language codes of ISO 639: a premature, ultimately unobtainable, and possibly damaging standardization. Handout of a talk given at the PARADISEC RRR Conference, December 2013.

http://www.academia.edu/4725536/The_language_codes_of_ISO_639_A_premature_ultimately_unobtainable_and_possibly_damaging_standardization

Version for publication is in preparation.

Nordhoff, Sebastian & Hammarström, Harald & Forkel, Robert & Haspelmath, Martin (eds.) (2013). *Glottolog* 2.2. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://glottolog.org> Accessed 2014-11-25.

Affix as a Comparative Concept

Matthew Dryer

Haspelmath (2011) argues that the widespread assumption that there is a clearly defined crosslinguistic concept of word does not hold up to scrutiny, that there are no good criteria for distinguish words from affixes and that claims that make reference to a distinction between words and affixes are suspect. He suggests that the widespread assumption that there is a crosslinguistic category of words may be an artifact of the writing systems used by European languages. The worry he raises is that the decisions that linguists describing languages make in terms of what they represent as words may at best be based on inconsistent criteria and at worst reflect unconscious influence of how similar constructions are represented in European languages. It is not obvious, for example, that the suffixing preference (Greenberg 1957) actually represents a typological pattern as opposed to a tendency to represent grammatical morphemes that follow stems or content words as suffixes but to represent grammatical morphemes that precede stems or content words as separate words simply because suffixes are more common than prefixes in the orthographies used by European languages.

Haspelmath argues that the problems that he discusses for word (and thus affix) as a crosslinguistic category are equally well problems for word and affix as comparative concepts. In this paper, I propose a comparative concept for affixes that avoids the problems that Haspelmath discusses. Haspelmath argues that there are serious problems with the notion of grammatical or morphosyntactic word; I believe his arguments are quite convincing. In fact, I would go even further and claim that there is no reason to posit a notion of morphosyntactic word in analysing any language. However, Haspelmath assumes that any notion of affix must be defined relative to morphosyntactic words. I argue that this is not in practice what people describing languages usually mean when they say that a morpheme is an affix. Rather, I argue that most linguists describing languages presuppose a notion of affix defined in terms of phonological words, not morphosyntactic words. I thus define the comparative concept affix (as a first approximation) as a grammatical morpheme that is part of a phonological word. There are various additional parts of the definition that I will discuss, but the most important one is that of phonological word: I use the maximal domain of segmental rules as the basis of a definition of phonological word as a comparative concept.

While I argue that I have provided a viable definition of a comparative concept affix, there remains the problem that it is not obvious that descriptions of languages always assume such a notion of affix in deciding whether to treat something as an affix or a separate word, so the possibility that the suffixing preference is an artifact of orthographic conventions cannot be ruled out. However, I will report on a study in progress of the phonological properties of tense-aspect affixes in a sample of over 400 languages. Since grammatical descriptions rarely treat morphemes containing only consonants or morphemes that exhibit phonologically-conditioned allomorphy as separate words, the hypothesis that the higher frequency of suffixes is an artifact of conventions in European languages must be based on morphemes that contain vowels and that do not exhibit allomorphy. But preliminary results provide evidence that the suffixing preference for tense-aspect affixes is just as strong if we restrict attention to morphemes that contain only consonants or that exhibit phonologically-conditioned allomorphy.

Abstract

Causatives and causation in Benabena

Carola Emkow

For the past few decades or so, the causative construction has been a recurrent research topic in linguistics (Comrie 1981, Givón , Dixon 2000, Shibatani & Pardeshi 2002)

A causative is a grammatical or lexical indication of the causal role of a referent in relation to an event or state expressed by the verb. The *causer* causes someone or something (the *causee*) to do or to be something, or causes a change of state in a non-volitional event.

In a causative event a new participant is introduced. A causative event thus involves ‘two events’, the causing event and the caused event (cf. Shibatani 1976: 1). The caused event may have only one participant, the intransitive subject (S), or two, the transitive subject (A), and the transitive object and (O).

An underlying intransitive clause with an intransitive subject (S) becomes a transitive object (O) in the causative, and a transitive subject (A) is introduced as a new argument.

I	made	him	leave.
A	Caus	O	leave
Causing event		Caused event	
A causes O to leave.			

An underlying transitive clause already has two arguments, in A and O functions. Provisions have to be made to accommodate all three arguments; two arguments in A-function, the newly introduced *causer*, and the transitive subject of the underlying transitive clause.

All languages have ways to express causation, but differ in the means. A causative construction may be marked by

- a morphological process applied to the verb in the clause (i.e. affixation, stem alternation, tone alternation, reduplication)
- two verbs in one predicate (i.e. compound predicate, auxiliary verb construction, serial verb construction)
- periphrastic construction (two clauses)
- lexically (different verb forms, one verb for both)

The fact that many languages have more than one causative construction has given rise to meaning-form correlations in causatives. Comrie (1981 164-167; 1989: 171-174, Haiman 1983: 783 – 785; Givón 1990: 556) suggested that the continuum of formal expressions of causatives from analytic/periphrastic to morphological to lexical, correlates with a meaning that is from less direct to more direct causation, and also from high control to low control on the part of the *causee*. Dixon (2000: 74-77) proposes a number of correlations between the various semantic parameters relating to the verb, the *causer* and the *causee* correlating with ‘compactness’ of expression. Direct causation, for instance tends to be expressed by a more compact mechanism than indirect causation. i.e., with a morphological process rather than an analytic causative if a language has both.

Another formal parameter in the semantics of causatives is their productivity. Shibatani & Pardeshi (2002: 111-112) claim that the less productive and more restrictive the causative, the more likely it is to express direct rather than indirect causation.

This talk will investigate causatives in Benabena, a Papuan language of the Gorokan family, spoken in the Eastern highlands of Papua New Guinea.

The language is a case-marking language, accusative-nominative in its pronominal system, and agent-patient marking in its nominal system. It is a predominantly suffixing agglutinating language. The predicate of a clause is often a light verb construction. The language has serial verb constructions and medial verb constructions.

There are three grammatical means to express a causative event

- a causative compound verb construction involving *le'mo* 'carry'
- a causative compound verb construction involving *li* 'take'
- a causative medial verb construction

“In the compound type, the causer’s action is expressed by a separate verb instead of a morphological element, but that verb must appear next to a basic verb so that other elements are not able to intervene between the two.” (WALS, chapter 9, non-periphrastic causatives).” (WALS, chapter 9).

(1) *Ai le'mo giya hu'ehive.*
 ai Ø-le'mo giya hu-ehive
 3Sg 3Sg.O-Caus laugh.LV.Pst.mf-Pst.3Sg
 He/She made me laugh.

(2) *Ana bo le'mo he'milave.*
 ana bo le'mo he'mi-lave
 that man Caus leave.Fut-Fut.3Pl.Dcl
 'They will get rid of that man.' (lit.: They will make that man leave.)

(3) *Nani li gala fuka i'ohuve.*
 nani li gala fuka i-ohuve
 1Sg Caus spill.water.Pst-Pst.1Sg.Dcl
 'I spilled the water.'

(4) *Gu'mo li huto hu-ohuve.*
 gu'-mo li hu-to hu-ohuve
 bilum-the Caus make/be.MV.Pst.mf make.Pst.mf-Pst.1Sg.Dcl
 'I made the bilum.' (I am the causer.)

In a causative medial verb construction the causing event and the caused event are linked by means of a non-finite clause containing the medial verb expressing the causing event, and a finite clause whose verb expresses the caused event.

(5) *Nani hugo buto ho'ehive.*
 nani hu-go bu-to ho-ehive
 1Sg make.MV.mf.Pst go.MV.mf kill.Pst.mf-Pst.3Sg.Dcl
 'I made him go and kill it.'

The talk will analyze all three means in terms of their formal properties and suggested meaning-form correlations.

**The dynamics of linguistic diversity:
Language contact and language maintenance in Amazonia**

Patience Epps

University of Texas at Austin

While efforts to understand global patterns of linguistic diversity have explored a wide range of nonlinguistic correlates, associations with sociocultural patterns have generally tended to assume a correspondence between the maintenance and/or genesis of distinct languages and a lack of contact among groups. In this talk, I argue that the maintenance of the extensive diversity of languages and families in the Amazon basin has in fact been widely grounded in the dynamics of interaction among groups, as opposed to being simply a factor of isolation. I focus on linguistic evidence for contact, drawing on an extensive survey of lexical and grammatical features across dozens of Amazonian languages. An evaluation of patterns of lexical borrowing, *Wanderwörter*, and grammatical diffusion suggests that multilingual interaction has been widespread in native Amazonia, facilitated by particular activities such as trade, intermarriage, and participation in networks of ritual practice. Interaction has been particularly heavy in particular zones, such as the Upper Rio Negro and the Guianas, but evidence of language contact indicates that these zones were themselves linked within much larger networks of interaction. Even while contact has led to a convergence of linguistic features on some levels, the particular dynamics of interaction across Amazonian regions – themselves arguably a product of diffusion in culture and discourse – have encouraged the maintenance of linguistic distinctions, most notably in the lexicon, thereby fostering the region's impressive levels of linguistic diversity.

A Universal on Generalized Sluicing

Study of syntax within formal traditions has uncovered a vast array of phenomena whose typology is yet to be explored. I propose here an implicational universal regarding the existence of reduced embedded *y/n*-questions, alternative questions, and *wh*-questions in the languages of the world. Starting from Ross (1969) who discovered the phenomenon and coined the term, *sluicing* – ellipsis in embedded *wh*-questions – has been extensively written upon.

(1) *Mary baked something, but I don't know [what ~~Mary baked~~].*

It is the *focus* of a *wh*-question that is retained under sluicing. This allows generalizing the notion of sluicing for other types of embedded sentences, Van Craenenbroeck and Liptak (2006, 2013). In particular, in some languages of the world a similar construction is possible for the focus of a *yes/no*-question or of an alternative question, as illustrated schematically in (2a-b) and by data from Lingala (Bantu C30, Congo) in (2a'-b'):

- (2) a. *y/n*-question
The cat caught something, but I don't know whether ~~it caught~~ a mouse.
- b. alternative question
The cat caught something, but I don't know whether ~~it caught~~ a mouse or a bird.
- a'. *Mary alambaki eloko kasi ngaji najebi te soki loso*
 Mary cook.PST¹ something but I know.PRS NEG whether rice
 'Mary cooked something but I don't know whether (Mary cooked) rice.'
- b'. alternative question
Mary alambaki eloko, (kasi) ngaji nakanisi te Sue
 Mary cook.pst something but I think.PRS that Sue
ajembi soki loso tope madeso
 know.PRS whether rice or beans
 'Mary cooked something and I think that Sue knows whether (Mary cooked) rice or beans.'

I call *y/n*-sluicing the construction exemplified in (2a) and *or*-sluicing the construction in (2b). In many languages, only *or*-sluicing is grammatical along with *wh*-sluicing, as is shown in (3) for Kannada (Dravidian).

- (3) a. *namma bekku yeen-oo tin-tu*
 our cat what-EMP eat-3.SG.N
aadare [[yeen-u anta] nana-ge got-illa]
 but what-NOM COMP I-DAT know-NEG
 'Our cat ate something, but I don't know what.'
- b. (...) *aadare [[hakki-n-oo ili-n-oo anta] nana-ge got-illa]*
 but bird-DISJ mouse-DISJ COMP I-DAT know-NEG
 'Our cat ate something, but I don't know whether (it ate) a bird or a mouse'
- c. (...) *aadare [[hakki-n-aa anta] nana-ge got-illa]*
 but bird-Q COMP I-DAT know-NEG
 'Our cat ate something, but I don't know whether (it ate) a bird.' (intended)

Finally, in some languages, e.g., in Amharic (Semitic, Afroasiatic); Degema (Edoid, Niger-Congo), Kaingang (Gê, Brasil), Khmer (Austroasiatic), and Chechen (East Caucasian) ellipsis of this type is impossible in any kind of embedded question. When grammatical, these varieties of ellipsis satisfy all standard tests for sluicing, as discussed in (Merchant 2001), that is, their syntax is similar to that of *wh*-sluicing. Cross-linguistically, their appearance is subject to an **implicational universal**:

- (a) If a language allows *y/n*-sluicing, it will also allow *or*-sluicing and regular sluicing.
 (b) If a language allows *or*-sluicing, it will also allow regular sluicing.

The universal has been checked on a sample that includes Polish, Russian, Serbian, Lithuanian, Hebrew, Hungarian, Tyvan, Yakut, Turkish, Lingala, Hindi, Persian, Digor and Iron Ossetic, Italian, French, Brazilian Portuguese, German, Slovenian, Albanian, Bulgarian, Georgian, Svan, Basque, Kannada, English, Dutch, Romanian, Finnish, Amharic, Chechen, Lezgian, Degema, Mandarin, Kaingang, Khmer, Adyghe, and Vietnamese.

¹ Glosses: COMP complementizer; DAT dative; EMP emphatic; NEG negation; NOM nominative; PRS present; PST past; Q interrogative.

Clause chaining and nominalization in Tarahumara: Towards a diachronic source of clausal nominalizing suffixes

Zarina Estrada Fernández
Universidad de Sonora
zarinaef@gmail.com

Jesús Villalpando Quiñónez
University of Colorado at Boulder
jesusvillalpandoq@gmail.com

Long sequences of foreground clauses or sequential event clauses are a characteristic of clause chaining according to Dooley (2010:90). In this line of reasoning, the author considers that the most important property of such sequences is the presence of switch-reference markers. On the other hand, Overall (2014) defines clause chaining as a “multi-clausal construction consisting of one or more dependent clauses associated with a single finite clause.” (p. 314). In line with the aforementioned definitions, this paper provides an analysis of clause chaining in Tarahumara, a Uto-Aztecan language from northwestern Mexico.

In this presentation, a diachronic perspective is taken. The analysis argues that the phenomenon of clause chaining in Tarahumara makes use of other operators rather than switch-reference markers, which seem to be lost in this language.

Our analysis shows that clause chaining in Tarahumara is encoded by means of simple and complex conjunctions and subordinators, as it is shown in (1) and (2), as well as by a set of different nominalizing suffixes:

(1) Simple conjunctions

a’rí ko échi kochí wirísi-ri,
CONJ EMPH DEM dog stand_up-PFV

a’rí chóta-ri échi músa neká-mia.
CONJ start-PFV DEM cat bark-IRR

‘And then the dog stood up and started to bark to the cat.’

(2) Complex conjunctions

mapu-a’rí chóta-ri bacháwara ukí wichí-ya
SUB-SIM start-PFV first rain fall-NMLZ

Pedro simí-re iwé-chi
Pedro go-PFV field-LOC

/a’rí karéwi-ri échi iwé kíti ichi-méa sunú.
CONJ weed_out-PFV DEM field PURP sow-IRR corn

‘When the first rains started to fall down, Pedro went to the field / and weed it out in order to sow corn.’

An identified set of nominalizing suffixes *-a* ~ *-ya*, *-ka*, *-cho* ~ *-chi* are also consistently used in Tarahumara clause chaining. This set of suffixes seems to have developed in order to mark specific functional domains: *-a* relevant to complement clauses, *-ka* to some adjectival constructions, and *-cho* to adverbial clauses. Clause chaining in Tarahumara also demonstrates that the boundaries between clauses at discourse level are not always clear and that clause types are fuzzy and may overlap.

REFERENCES

Dooley, Robert A. 2010. *Exploring Clause Chaining*. *SIL Electronic Working Papers in Linguistics*.

http://www.sil.org/system/files/reapdata/13/04/21/130421360311569160306235201738876375986/silewp2010_001.pdf [Accessed on September 30, 2014].

Overall, Simon E. 2014. "Clause chaining, switch reference and nominalisations in Aguaruna (Jivaroan)." In *Information Structure and Reference Tracking in Complex Sentences*. Rik van Gijn, Jeremy Hammond, Dejan Matic, Saskia van Putten and Ana Vilacy Gaucio (eds). Amsterdam/Philadelphia: John Benjamins Publishing Company. pp. 309-340.

A typological analysis of loan translation in contact languages

Loan translation (also called calquing) has not received a lot of attention either in general linguistics or in contact linguistics. In the field of contact linguistics loan translation has either been sidelined (e.g. Mittelsdorf 1978) or has been treated in passing in works that focus on word-formation (e.g. Lefebvre 1998; Farquharson 2012). This is unfortunate given that loan translations are a signal of deep cognitive processes at work in language contact and subsequent language creation. They also provide crucial evidence for testing whether certain lexical concepts are more susceptible to loan translation and what types of concepts are translated in specific contact situations.

The current paper seeks to address the lack of focus by analysing putative loan translations that have been identified for various Atlantic Creoles and Mauritian Creole (which has a Narrow Bantu substrate). The analysis is based on 50 potential African calques drawn from the list provided in Parkvall & Baker (2012). For manageability, it will be restricted to bi-morphemic constructions: phrases (e.g. *get belly* ‘be/become pregnant’) and compounds (e.g. *big-eye* ‘greed/y’).

The paper will tackle several aspects of the typology of these loan translations centering on their etymology, formal structure, and semantics. The issues to be addressed include: (a) whether loan translations tend to follow the morphological/linearisation patterns of their model language (ModL) or those of the language from which they draw their material (MatL); (b) whether loan translations create new sememes and/or lexemes from the perspective of the superstrate language; (c) whether there is a tendency for loan translations to be traceable to one versus multiple substrate languages and what this can tell us about the role of multiple reinforcements in the translation process; (d) the semantic domains (e.g. body parts, human characteristics) in which we find loan translations and how that relates to the nature of contact; (e) whether there is any correlation between the semantic domains in which we find loan translations and the substrate/superstrate languages involved; (f) the typological orientation of loan translations in specific semantic domains, such as zoological vs. anthropocentric orientation in translations of human characteristics and (g) what lexical gaps exist in specific semantic domains and how these gaps can be explained.

The paper hopes to provide the starting point for a broader study of loan translations not only in contact languages but also those that are not known to be the product of contact.

"FORWARD TO THE PAST" - MULTIMODALITY OF TEMPORAL MARKING IN
EASTERN KHANTY.

The fact that languages consistently conceptualize time through space is part of fairly uncontroversial typological universals. Within this metaphorical extension from space to time, languages normally code the speaker as facing the flow of time and approaching events, thus future is conceptualized (universally) as in front (Núñez, Sweetser, 2006). In the majority of the cases, the events are viewed as approaching the speaker-observer, thus future events are «in front», while the past events are «behind».

Co-gesturing is a universal phenomenon, which adheres to the same principles of conceptualization and metaphorization as the spoken language (McNeill, 1992; Sweetser, 1998; Smith 2003; Núñez, 2006). Moreover, it was repeatedly noted that co-gesturing does not merely duplicate, but rather supplements the spoken language, and as such can be a valuable empirical resource for the study of cognitive processing of abstract notions, perhaps, even as leading data register (Kendon 1982; McNeill, 1992; Iverson & Thelen, 1999; Mayberry & Jaques, 2000; Kita & Essegbey, 2001; Núñez & Sweetser, 2001; Goldin-Meadow, 2003). Spatial co-gesturing is directly iconic as it operates in the spatial domain: speaker's body serves as a reference point in space in relation to other objects/bodies; either at the moment of speech, or construing a situation different from the moment of speech (cf. (Haviland, 1993; Levinson, 2003)).

Eastern Khanty consistently demonstrates the metaphoric extension «space» - «time», both in the domain of lexical means (postpositions, adverbs, case system) (cf. examples (1-3)), and in the domain of co-gesturing (cf. picture (1-2) and (3-4)). Following the methodology for the analysis of co-gesturing in Ayamara (Núñez & Sweetser, 2006), video recordings of Eastern Khanty speech events were used in the study. One consultant was selected for the pilot phase, who consistently displayed co-gesturing across genres and registers, and in code-switching. Spontaneous spatial and temporal co-gesturing was registered in the data within the spoken language contexts, and later cross-checked in elicitation, blind to the exact controlled parameters, study objective and hypothesis.

Overall, co-gesturing is quite frequent, with spatial gestures (cf. (Pic.1) and (Pic.2)) dominating over temporal, and when temporal, gestures prevalingly refer to the plain of the Past (cf. (Pic.3) and (Pic.4)). In temporal gesturing, prevailing majority are within the sagittal plain (<front-back>). In sagittal gestures, the majority manifest the type <Past=In front> (cf. (Pic.3-Pic.4) and (Pic.5-Pic.6)), while the minority are of the type <Future=Behind> (cf. (Pic.7)-(Pic.8)). The types <Past=Behind> and <Future=In front> are not registered. Temporal gesturing co-occurring with an adverb /il'-'/ 'fore, front', both for spatial and temporal domains, is signified by hand motion forward from the corpus (also registered in code switching, co-occurring with the Russian lexemes /remäl' 'time' or /vpered/ 'forward' (cf. (Pic.5) and (Pic.6)). Both, the spoken lexical means (adverb /il'') and the gestured 'forward', are used for coding absolute (with a deictic center in "now") and relative (without a deictic center) temporal relations. Similarly to Ayamara observations (Núñez & Sweetser, 2006: 438-439), Eastern Khanty demonstrates the use of a dynamic conceptual metaphor, preferring specifically the construal «Known=Visible» thus «Past=In front» to the construal «Path covered=Known» thus «Past=Behind». The preference is likely associated with the salience within the system of the evidential status of the information. In Eastern Khanty, the key TAM opposition is between Future-Present and Past (Tereskin 1961; Gulya 1966; Filchenko 2007), that is a typical irrealis-realism dichotomy (Filchenko 2007: 254). Availability in the Eastern Khanty system of 4 Past Tense forms combining the temporal and evidential senses (Remote Definite Past, Remote Indefinite Past, Proximal Definite Past, Proximal Indefinite Past) testifies to the importance (manifest in the obligatory grammatical coding) of the temporal-evidential status of information, which maybe seen as an evidence of the visual conceptualization of knowledge (Núñez & Sweetser, 2006: 440-441).

The study is its early stages and awaits further methodological and empirical development. The prospects of the project include a more exhaustive inventory of the spatial and temporal coding means, both lexical and gestural, as well as the aspects of their individual and coordinated distribution. Naturally, extension of the empirical data, both genetically and areally is anticipated.

1) Khanty, Vas. (Filchenko 2007: 103):

- | | |
|---|---|
| a) <i>ämp qat-nə aməs-wəl</i>
dog house-LOC sit-PRS.3SG
'dog sits in the/a house' | b) <i>toyj-nə... men-s-əw</i>
spring-LOC go-PST2-1PL
'Once in spring. . ., in spring we went . . .' |
|---|---|

2) Khanty, Vas.:

- | | |
|---|---|
| a) <i>köt-äm pirt-a nirimtä-s-im</i>
hand-1SG back-ILL pull-PST2-1SG/SG
'I pulled my hand back' (Filchenko 2007: 204) | b) <i>jöy mən-m-äl pir-nə ...</i>
3SG go-PP-3SG back-LOC
'After he went away...' (Gulya 1966) |
|---|---|

3) Khanty, Vas.: (Filchenko 2007: 179):

- | | |
|--|--|
| a) <i>il'-pä äjri-nä jaxəntə-s-əm</i>
fore-ALL canoe-COM paddle-1PST2-SG
'I am paddling forward in my canoe' | b) <i>köröy il'-ən jöy-nämə ran'it' werä-γən</i>
eagle fore-LOC 3SG-RFL wound do-PST0.3SG
'Earlier, the eagle got himself wounded' |
|--|--|

Pic 1.



il'-əpä
fore-EP-ILL
'forward'
'I am moving on forward in my canoe'

Pic 2.



il'-əpä
fore-EP-ILL
'forward'

Pic 3.



il'-län-nə
fore-time-LOC
'long time ago'
'not long ago there were more people in the village'
'awhile ago there was a school in the village'

Pic 4.



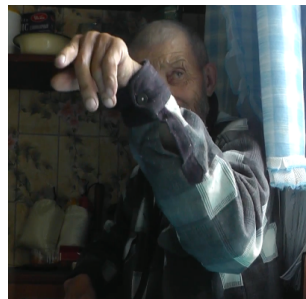
il'-län-nə
fore-time-LOC
'in the old days, before'

Pic 5.



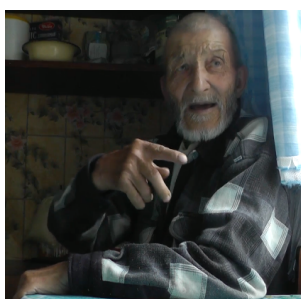
il'-remä-nə
fore-«time»-LOC
'long time ago, before'
'long time ago there lived warriors here'

Pic 6.



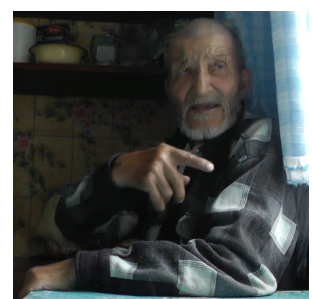
(Russ.) *раньше*
'earlier'

Pic 7.



pir-nə
back-LOC
'after that'
'smth will happen afterwards, later'

Pic 8.



pir-nə sayi
back-LOC along
'after that, later'

TOMSK EASTERN KHANTY AND SOUTHERN SELKUP DOCUMENTATION PROGRAM: INTERIM REPORT

The presentation concerns the ongoing ELDP sponsored project “Comprehensive Documentation and Comparative Analysis of the endangered Eastern Khanty and Southern Selkup: legacy data archival and fieldwork with the last speakers”.

The cooperative project between the University of Zurich and Tomsk State Pedagogical University builds on the applicants’ extended research experience with the respective languages, pursuing documentation and analysis of two endangered languages of Western Siberia. Both languages of the project (i) are severely endangered, each numbers under 10 last proficient speakers; (ii) have existing legacy data in Tomsk with limited access, which is to be improved as a result of the project; (iii) are characterized by areal contact contiguity; (iv) are ethnographically consistent.

The project contains a strong interdisciplinary research component and builds on previous cooperation with the members of academic and indigenous communities. The analytical component of the project encompasses a range of questions and approaches within the wide but interconnected disciplinary field of linguistic anthropology and typology of western Siberia. Special attention is paid to the issues of local contact, variation and change in describing the specifics of the organization of the language systems.

The central intellectual merit of the project is that it surveys a large and linguistically diverse area of Western Siberia, with the languages and cultures on the verge of extinction. It also maintains a robust research group, who have worked in the area and have knowledge of the local communities. The project also builds on a long research tradition, going back to the 1940s, of fieldwork and analysis of indigenous languages of Siberia, preserved in the department's legacy archive. The project utilizes expertise in language archiving technology and best practices of language documentation accumulated in the previous projects when aiming to produce analytical results in comparative and typological linguistics based on conventional functional theoretical and methodological basis, contributing to a number of ongoing debates in linguistics and broader humanities.

The broader impacts of the project are expected to be felt in the local communities of Western Siberia, receiving materials on their language and culture, in durable and easily accessible formats. Linguists and anthropologists throughout will benefit from a new archive of language materials and the theoretical investigations made possible by the project.

New methodological infrastructure developed by the project makes language documentation work more efficient and creates new possibilities for semantic annotation of language data.

The presentation outlines the results of the first two years of the project.

Approaching valency patterns in Spanish Sign Language (LSE)

José M. García-Miguel & Carmen Cabeza
(Universidade de Vigo)

Abstract

We will present a preliminary approach to valency and predication in Spanish Sign Language, aiming to provide a few data that may contribute to the development of intermodal typology. The Valency Patterns Leipzig database [ValPaL] (Hartman, Haspelmath, and Taylor 2013) is an important for study of valency patterns in a sample of 35 languages from around the world. However, this sample does not include any sign language. Other relevant typological databases, such as WALS (Dryer and Haspelmath 2013) and APiCs (Michaelis et al. 2013) also make reference to sign languages only sporadically or marginally.

One may even wonder is the same principles and tools that have been applied to oral languages could also be applied to sign languages. Typologically, sign languages seem to be relatively similar to creoles and relatively close to what Gil calls IMA languages (i.e., isolating, monocategorical, associational), which represent a minimum degree of grammaticalization. Like many creoles, sign languages lack almost entirely core-argument flagging (Gil 2014). However, they make use of other strategies such as indexation and incorporation, that have already been used as the basis of verb classes.

We will present an analysis of Spanish Sign Language based on the elicitation of examples corresponding to 80 core verb meanings of ValPaL. By using the same set of basic verb meanings as ValPaL, we expect to make the comparison with oral languages easier. The goal is to describe and evaluate the resources and strategies used in LSE for the lexicalization of predicative meanings, for the expression of event participants, and for the distinguishing of participant roles.

References

- Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/>.
- Gil, David. 2014. Sign languages, creoles, and the development of predication. In Frederick J. Newmeyer & Laurel Preston (eds.), *Measuring grammatical complexity*, 37–64. New York: Oxford University Press.
- Hartman, Iren, Martin Haspelmath & Bradley Taylor (eds.). 2013. *Valency Patterns Leipzig (ValPaL)*. Pre-release version. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://valpal.info>.
- Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). 2013. *APiCS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/>.

Noun phrase construction in Remo and Gutob

Arun Ghosh

Georg-August University Goettingen and University of Burdwan

Noun phrase (NP) as a rule can be used as subject, object or compliment in a sentence. Remo and Gutob, the two south Munda languages spoken in India, form noun phrases in different ways. In formulaic shape it may be presented like the following:

- NP -----→ N
- NP -----→ ADJ. + N
- NP -----→ N + ADJ.
- NP -----→ N + QUAN.
- NP -----→ QUAN. + N
- NP -----→ NUM + ADJ. + N
- NP -----→ N + ADPOS. + N
- NP -----→ N + ADPOS. + QUAN. + N
- NP -----→ QUAN. + N + CONJ. + N
- NP -----→ NUM + INTENS. + ADJ. + N
- NP -----→ N + V + PTCPL. + DEM. + V
- NP -----→ V + PTCPL. + N
- NP -----→ N + ADPOS. + V + PTCPL. + N
- NP -----→ N + ADPOS. + V + PTCPL. + ADJ. + N
- NP -----→ N + ADPOS. + V + PTCPL. + CONJ. + N + ADPOS. + V + PTCPL. + ADJ. + ADJ. + N
- NP -----→ V + PTCPL. + N + CONJ. + N + ADPOS. + V + CONV. + PTCPL. + ADJ. + ADJ. + N
- NP -----→ N + ADPOS. + V + PTCPL. + V + ADPOS. + N
- NP -----→ N + ADPOS. + V + PTCPL. + NUM + ADJ. + ADJ. + N
- NP -----→ N + ADPOS. + V + PTCPL. + CONJ. + N + V + PTCPL. + NUM + ADJ. + CONJ. + + ADJ. + N
- NP -----→ N + V + PTCPL + N + ADPOS. + V + PTCPL. + ADJ. + NUM + ADJ. + N

Sometimes whole sentence takes the form of a phrase in some clause types. In the presentation different types of NP constructions of the languages concerned will be looked into along with parallels in other Munda languages.

Typology Reflects Phylogeny

David Gil

Ontogeny Recapitulates Phylogeny is an often-repeated (albeit also much-disputed) slogan expressing the hypothesis that stages in the development of an animal from embryo through infancy to adulthood mirror stages in the evolution of the animal's own ancestors. Among others, this hypothesis has also been applied in attempts to shed light on the evolution of language, with claims that early child language may provide a model for the evolution of early human or pre-human language.

This paper presents a companion hypothesis encapsulated in the slogan *Typology Reflects Phylogeny*, arguing that cross-linguistic variation may provide an alternative and in some respects better model for the evolution of language. The substitution of "reflects" for "recapitulates" acknowledges that typology differs in one crucial respect from both ontogeny and phylogeny, namely, it is non-directional. Whereas ontogeny and phylogeny are inherently bound to the arrow of time, linguistic variation has no intrinsic "earlier" and "later". For example, if Czech and Polish differ with respect to the relative order of noun and adjective, then there is no a priori reason to suppose that either of the two orders is more representative of a prior stage in the evolution of language.

However, one basic measure provides a means for characterizing some feature values as likely to be evolutionarily prior to others: *complexity*. If one particular feature value is simpler than its more complex alternative, then the simpler feature is likely to represent an earlier stage in the evolution of language with respect to the feature in question. Underpinning this inference is the fact that evolution typically proceeds from simpler to more complex. In other sciences, similar inferences are commonplace. For example, life today consists of simpler prokaryotes, more complex unicellular eukaryotes and even more complex multicellular eukaryotes; such progression in contemporary complexity is then taken to represent the evolutionary path of life. By the same token, numerous aspects of cross-linguistic variation involve features with respect to which one particular value may be said to be more complex than another. Most straightforwardly, languages differ with respect to inventory sizes of various items such as phonemes, affixes, parts of speech, speech registers, words, and so forth; for each of these, fewer is clearly simpler than more, and hence fewer may in principle be hypothesized to represent an evolutionarily prior stage in the evolution of that particular linguistic feature.

This paper provides support for the methodology of using cross-linguistic variation as a window into phylogeny, focusing on the domain of compositional semantics. When two or more expressions are brought together, the meaning of the juxtaposition must be associated in some way with the meaning of its constituent parts; this is the most basic principle underlying compositionality. However, languages vary considerably with respect to the extent to which the meaning of such a juxtaposition may be further delimited by various grammatical rules making reference to morphosyntactic strategies such as agreement, flagging, linear order and others. Whereas in some languages, of *high articulation*, there is an abundance of such rules, substantially limiting the range of possible interpretations, in other languages, of *lower articulation*, there are fewer such rules, and the range of possible interpretations is thus broader. In terms of their compositional semantics, low articulation languages are thus less complex than their high articulation counterparts, and may accordingly be hypothesized to represent an earlier stage in the evolution of compositional semantics.

This paper presents the results of an ongoing cross-linguistic experiment measuring the degree of articulation of compositional semantics in the domain of thematic role assignment; to date, close to 50 languages have been examined. The main factor

governing the degree of articulation is shown to be the *polity type* of a language, ranked in accordance with the following four-valued scale of increasing socio-political complexity: (1) regional language until recently not part of a larger polity (e.g. Ju|'hoan, Mursi, Mentawai, Marind); (2) regional language that has been part of a larger polity for a long period (e.g. Galician, Yoruba, Shan, Minangkabau); (3) regional variant of a standardized national language (e.g. Neapolitan, Cantonese, Osaka Japanese, Riau Indonesian); and (4) standardized national language (e.g. Italian, Mandarin, Japanese, Indonesian). The experimental results show that languages of greater socio-political complexity tend to exhibit higher articulation. Thus, within in the domain of compositional semantics, grammatical complexity correlates positively with complexity of polity type. However, with respect to polity type, there is abundant extra-linguistic evidence, from history, archaeology and anthropology, that polity types evolved from simple to complex. Thus, the correlation between polity type and degree of articulation of compositional semantics provides further extra-linguistic support for the claim that low-articulation languages represent a prior stage in the evolution of compositional semantics. In so doing, the results of the experiment provide independent support for the use of cross-linguistic variation with respect to complexity as a window into the evolution of language.

Vilela

- (4) hate *wah-el dupmawe jasi-e*
man eat-CONV tasty.food sit.down-3SBJ
'the man who ate tasty food sat down' (Golluscio 2014, forthcoming)

Ayacucho Quechua

- (5) *Tusu-q warmi=qa kuya-wa-n*
dance-NMLZ woman=TOP love-1OBJ-3SBJ
'The woman who is dancing loves me.' (Zariquiey and Córdova 2008: 88)
- (6) *warmi hamu-chka-n chay, pay=mi rima-n-qa*
woman come-PROG-3SBJ DEM 3SG.PRO=EVID speak-3SBJ-FUT
'the woman who is coming, (that one) will speak' (Cerrón-Palomino 2008:240)

Aymara

- (7) *kapurala thuqhu-ri jaqi=xa tiyu-ja=wa*
caporal dance-NMLZ person=TOP uncle-1POS=EVID
'the person that dances the caporal is my uncle' (Cerrón Palomino 2008: 230-231)
- (8) *warmi juta-sk-i uka=xa, jupa=wa parla-ni*
woman come-PROG-3SBJ DEM=TOP 3SG.PRO=EVID speak-3SBJ.FUT
'the woman who is coming, (that one) will speak' (Cerrón-Palomino 2008:240)

Chipaya

- (9) *walja lik-ñi soñi=ki sel-at-ki=ca*
a.lot drink-NMLZ man=TOP be-IMP-REP=DEC
'there was a man who drank much (lit. who was a drinker)' (Peña 2008:10)
- (10) *tii mizi thaj=tra, nii=ki we-t-ta=tra*
DEM cat sleep=EVID DEM=TOP 1SG.PRO-COP-GEN=EVID
'that cat that is sleeping, that one is mine' (Cerrón-Palomino 2006:263)

Lule

- (11) *pele inle wotikeyu-ton wetsi-p*
man yesterday work-NMLZ die-3SBJ
'the man who worked yesterday is dead' (Machoni 1732, grammar section, p. 56)
- (12) *tanta kis kai-ts mima p'o-p*
bread 1PRO eat-1SBJ ANA.DIST be.white-3SBJ
'the bread that I am eating, (that one) is white' (Machoni 1732, grammar section, p. 65)

Mapudungun

- (13) *Kütran-küle-y ti wentru kupa-ke-lu*
be.sick-PROG-IND.[3SG] DET man come-HAB-CONV
'the man who always comes is sick' (HM, own data)

References

- Bertinetto, P. & Ciucci, L. 2012. Parataxis, Hypotaxis and Para-Hypotaxis in the Zamucoan Languages. *Linguistic Discovery* 10.1: 89-111.
- Campbell, L. & Grondona V. 2012. Languages of the Chaco and Southern Cone. In: L. Campbell, and V. Grondona, (eds.), *The indigenous languages of South America. A comprehensive guide*. Berlin: De Gruyter Mouton, 625-667.
- Cerron-Palomino, R. 2008. *Quechumara. Estructuras paralelas del quechua y el aymara*. La Paz: PROEIB Andes.
- Cerron-Palomino, R. 2006. *El chipaya o la lengua de los hombres del agua*. Lima: Pontificie Universidad Católica del Perú.
- Comrie, B.; Golluscio, L.; González, H. & Vidal, A. 2010. El Chaco como área lingüística. In Z. Estrada, Z. & R. Arzapalo (comps). *Estudios en lenguas amerindias 2. Contribuciones al estudio de las lenguas originarias de América*, 85-132. Hermosillo: Universidad de Sonora.
- Golluscio, L. 2014. Huellas de trayectorias y contactos en el sistema lingüístico: el caso vilela. In B. Comrie & L. Golluscio: *Languagecontact and Documentation/Contacto lingüístico y documentación*, 77-120. Berlín: De Gruyter (forthcoming).
- Golluscio, L.; Hasler, F. & de Reuse, W. 2014a. "Adverbial subordination at the peripheries of the Andean and Chaco linguistic areas". Presented at SSILA/Linguistic Society of America Annual Meeting, Minneapolis, January 2014.
- Golluscio, L.; Hasler, F. & de Reuse, W. 2014b. "El papel de la nominalización en la identificación de áreas lingüísticas en Sudamérica". Presented at the Workshop "Nominalización en las lenguas de América", organized at the Pontificia Universidad Católica del Perú, Lima, Peru, by Matt Shibatani & Roberto Zariquiey.
- Golluscio, L.; Hasler, F. & de Reuse, W. 2014c. "Estrategias de subordinación en dos áreas lingüísticas de América del sur: Chaco y Andes". Presented at the XVI Jornadas de Lengua y Literatura Mapuche and the V Congreso Internacional de Lenguas y Literaturas Indoamericanas. Universidad de La Frontera, Temuco, Chile.
- Gonzalez, H. 2014. El Chaco como área lingüística: una evaluación de los rasgos fonológicos. In B. Comrie & L. Golluscio: *Language contact and Documentation/Contacto lingüístico y documentación*, 201-247. Berlín: De Gruyter (forthcoming).
- Machoni, A. 1732. *Arte y vocabulario de la lengua lule y tonocoté*. Madrid: Herederos de Juan Garcia Infanzon.
- Messineo, C. and Porta, A. 2009. Cláusulas relativas en toba (guaycurú). *International Journal of American Linguistics* 75.2:49-68.
- Muysken, P. 2000. From linguistic areas to areal linguistics: a research proposal. Gilbers et al. (eds.) *Languages in Contact. (Studies in Slavic and General Linguistics, Vol. 28.)* Amsterdam: Rodopi. Pp. 263-75.
- Nercesian, V. 2011. Gramática del wichí, una lengua chaqueña. Interacciones fonología-morfología-sintaxis en el léxico. Doctoral thesis, Universidad de Buenos Aires.
- Peña, J. 2008. "Complex clauses in Chipaya: main strategies for complementation, relativization and adverbial subordination". *Proceedings of the Conference on Indigenous Languages of Latin America III*. University of Texas: Austin.
- Van Gijn, R.; Golluscio, L.; González, H. & Vidal, A. 2013. "Adverbial subordination strategies in the Chaco and beyond". Presented at SSILA/Linguistic Society of America Annual Meeting, Boston, January 2013.
- Zariquiey, R. & Cordova, G. 2008. *Qayna, Kunan, Paqarin. Una introducción práctica al quechua chanca*. Lima: Pontificie Universidad Católica del Perú.

Selection of element sources in the verbal syntagms of some stable mixed languages.

Anthony Grant

Much work on a range of stable mixed languages (SMLs) has sought to examine ways in which elements of differing origins in the language are combined (Bakker 2013, Christiansen-Bolli 2010, Gómez Rendón 2008, Matras 2003, Meakins 2011, Mous 2003), but the picture is complex and the search for a definitive explanation continues as further and more detailed data become available. Of cardinal importance to an understanding of the development of SMLs is the structure of the verb group or syntagm and the linguistically diverse range of entities which can be included within the boundaries of the verb syntagm. Languages which divide their components between Verb and Noun according to the origins of the stems and morphology differ in this respect from those which make a primary division of sources between lexicon and bound inflectional morphology.

The range of compulsory and optional forms and what else can express ideas which may be expressed by bound morphs vary from one SML to another when the source of the verbal morphology is compared with the source of the basic or high-frequency lexicon. Drawing from data from a wide range of SMLs I argue that there is a very high degree of correlation between the sources of the inflection of the minimally complex SML verb (whatever the source of the verb stem may be) and the source of the elements which more complex verb syntagms in the source languages of the SMLs exhibit. These elements (pronominals, negators, etc) may nonetheless be expressed with free rather than bound morphemes in these source languages, and they are often usually so expressed in the SMLs. This correlation is strongest in the case of SMLs which arose in a strongly bilingual situation which was stable for several generations. I further illustrate the fact that this state of affairs can, however, be strongly disrupted in cases where multilingualism has been characteristic of the SML speech community.

- Bakker, Peter. 2013. Michif. In Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *The survey of pidgin and creole languages. Vol. III: Contact languages based on languages from Africa, Asia, Australia, and the Americas*, 158-165. Oxford: Oxford University Press.
- Christiansen-Bolli, Regula. 2010. *A Grammar of Tadaksahak, a Northern Songhay Language of Mali*. Köln: Rüdiger Köppe Verlag.
- Gómez-Rendón, Jorge. 2008. *Mestizaje Lingüístico en los Andes. Génesis y Estructura de una lengua Mixta*. Quito: Abya-Yala.
- Matras, Yaron. 2003. 'Mixed languages: re-examining the structural prototype.' *The mixed language debate*, edited by Peter Bakker and Yaron Matras, 151-176. Amsterdam: Benjamins.
- Meakins, Felicity. 2011. *Case-marking in Contact: The Development and Function of Case Morphology in Gurindji Kriol*. Amsterdam: John Benjamins
- Mous, Maarten. 2003. *The making of a mixed language: the case of Ma'a/Mbugu*. Amsterdam: Benjamins.

The Central Eastern German KL>*TL cluster shift in a perspective of areal phonetic typology

Sven Grawunder, Max Plank Institute for Evolutionary Anthropology, Leipzig, Germany
grawunder@eva.mpg.de

This talk seeks to provide a demonstration of the approach of a phonetic areal typology, as quantitative approach of doing areal typological research. A phonetic typology needs to consider the diversity and variability that is given by the phonetic implementation of a phonemic contrast. In this perspective it makes the most sense to adopt an understanding of phonetics and phonology as inseparable entity like e.g. argued for by Blevins (2006). Such view allows to talk about phonetic implementation of abstract contrasts on different phonetic domains (inter-, supra-, segmental) as well as about diachronic processes and universals at the same time. Here most sound change relevant for phonemic distinctions would be driven by the listener's misinterpretation of articulatory variations, which occur in the speaker. Regular speaker variations can also reflect social group norms and thus social barriers. Other variations can be caused by a number of (external) mostly speaker-related factors, like L2-influence, but also physical variations like denture, adverse morphology, hearing loss etc. The regularity and systematicity within a population, e.g. driven by analogy and universal tendencies, lead then to a homogenous looking areal pattern.

At another place (Grawunder in prep.) I show that the type frequency distribution of segments in the lexicon is related to the actual phonetic expressions (realizations) as found in an areally-structured collection of comparable phonetic data from the Northeast Caucasus. Here I could demonstrate that acoustic analysis can also be applied to a more wide-meshed grid, where individual data points represent a language variety.

In this talk I will present data from an areal pattern of sound change that was described by Blevins & Grawunder 2006. The sound change comprises a fronting of the cluster /kl/ to */tl/ in parts of Central East German (CEG) dialects and colloquial forms. This sub-phonemic shift goes in parallel with other cluster shifts like /gl/ to */dl/, /kn/ to */tn/ and /gn/ to */dn/. The acoustically oriented analysis is based on a set of recordings from 32 localities with 3 speakers per village equally spread across the CEG area, overlapping the area of the cluster shift. By means of an analysis of gradual variance I can e.g. argue for a higher robustness within the cluster-shift area than at its borders. This can in part be shown for within-speaker, between-speaker and between-locality relations.

However, it can also be shown that the cluster shift is a truly areal pattern, specifically by its implementations in the other adjacent Southern dialects of German, but also by its pervasiveness with regard to dialect borders and language borders (Sorbian). Recent research e.g. on the /tl/ and */tl/ cluster perception (Halle & Best 2007) would corroborate the idea that the underspecified stop+lateral cluster allows for a reduced variety of pathways to develop into, especially if we see it in the light of language contact (L2-speaker behavior, incomplete learning). This would then also explain the independence of similar cluster shifts, as we see in some varieties of English (Blevins & Grawunder 2006).

"Grammatical relations in Taa (West !Xoon)"

Tom Güldemann & Christfried Naumann

This paper gives an overview of the grammatical relations in the West !Xoon dialect of Taa (Tuu family, Namibia, Botswana), before the background of an intensive documentation of the entire language complex.

Starting out in particular from Dryer's (1997) idea that grammatical relations are not universal but language-specific, we investigate how the existing formal encoding means reflect different types of confluences of relevant semantic-functional properties of noun phrases, which relate first of all to semantic role, animacy hierarchy, and information structure. Like in other Non-Khoe languages in the Kalahari Basin, syntactic relations in Taa are regulated first of all by a strict template-like clause organization with the basic constituent order S V O OTHER and a rare type of semantically generic prepositional flagging on postverbal participants, the so-called "multi-purpose oblique (MPO)" marker. However, Taa stands out even against this cross-linguistically unusual profile because of its typologically rare agreement pattern and its unusually large number of prepositions beyond the MPO marker.

Dryer, Matthew S. 1997. Are grammatical relations universal? In Bybee, Joan, John Haiman and Sandra A. Thompson (eds.), *Essays on language function and language type: dedicated to T. Givon*. Amsterdam: John Benjamins, 115-143.

ILLUSTRATIONS AND VIOLATIONS OF THE SYNTAX-PHONOLOGY INTERFACE: DATA AND THEORETICAL INTERPRETATION

Claude Hagège, Collège de France, Paris

The syntax-phonology interface is illustrated, among others, by some languages, in which syntactic structures are directly accessible to phonological phenomena, including tones. This is what we find in the noun-phrases of several dialects of Dogon (Heath & McPherson 2013), where inherent lexical tones are overridden by syntactically conditioned replacive tones. Thus, the inherent tonal contours H(igh)L(ow) of *kúrù* "stone" and LHL of *ìsê*: "village" remain unchanged with an "all" quantifier, as in examples (1)a.1 and (1)b.1. But when the noun is followed by a deictic, as in (1)a.2, or preceded by a possessor, as in (1)b.2, these contours are overridden, respectively, by an LL and a LH overlays, these types depending on the position of the controller; this is also true of adjectives and internally headed relative clauses. All of these, therefore, are reference restrictors, whose syntactic-semantic status is thus pointed out in Dogon by a dedicated (tonal) mark. This is not the case, in Dogon languages, of such elements as topic markers meaning "only" and "even", as well as nonsingular numerals and the independent plural particle. All of these, just like "all" quantifiers, do not impose any replacive tone to the adjacent element in the sentence.

Kpelle (Mande, Manessy 1964) exhibits a (fairly rare) alternation between two consonant sets, one of which has a preceding low tone. There are, among others, five alternation types: 1) $p\sim \times b$, 2) $X\sim m \times$, 3) $s\sim \times z$, 4) $k\sim \times g$ and 5) $f\sim \times v$. In (2)a.2, we see that the second type L(ow) T(one) C(onsonant) is used as a mark of the definite article (a deictic particle *-i* being also suffixed to the noun). The first type LTC has the same usage, as seen in (2)a.1), but it also appears as a mark of a predicative adjective, as in (2)c.2 vs (2)c.1). The third type LTC, when affecting a quantity word, refers to the totality from which this quantity is extracted, as shown by (2)b.2 vs (2)b.1). The fourth and fifth types LTCs mark, but only in the 3SG, the subject of the well-known four predicative auxiliaries of the Mande languages, two of which appear here, i.e., when not submitted to the LTC alternation, positive progressive *ká* in (2)d.1 vs (2).d.2, and negative perfective *fé* in (2)d'1 vs 2d'.2. Thus, in Kpelle, interestingly, the same tonological process applies to two distinct syntactic structures: predication and determination. The fact that at least one language presents, through this phonology-syntax interface, an identical treatment shows that determination and predication are two aspects of one and the same syntactic-semantic phenomenon.

In Kiyaka (Bantoid, Van den Eynde & Kyota 1984: 68-70), the basic forms for "father" and "stood up" are, respectively, *ba-taáta* (mid-tone unmarked) and *vuumbukídí*. But examples (3)a.1-(3)a.2 show that when the noun receives a modifier, like demonstrative *bà-na*, and, in a parallel way, when the verb receives a complement, like *ha-tháangi* "in bed", an identical tone-lowering morpheme marks both of these phenomena. And (3)b.1-(3)b.2 show that the noun when it is itself a modifier, as well as the verb when it appears in a modifying relative clause, are both affected by an identical tone-lifting morpheme. Therefore, in Kiyaka, a modified noun and a modified verb are treated by the same tonal morpheme, and this is also true of a modifying noun as well as of a modifying verb (within a relative clause). Thus verbs and nouns are morphologically treated in the same way (here by tones) in terms of dependency, i.e. when both function as centers, as well as when both function as dependents. This means that phonology can reflect and stress essential syntactic homologies.

However, besides these cases of strict tonology-syntax correspondence, there is another case in which phonological phenomena are totally independent of syntax, namely clitics. In Mukri (Central Kurdish, Öpengin 2013), example (4)a shows that clitic placement does not take into account at all the link between agent pronoun and verb, even less so when the same clitic, in addition to the agent, also marks, by syncretism, the possessor, as in (4)d.; nor does clitic placement take into account the link between a preposition and its complement,, as shown by (4) b.-c..

In conclusion, the correspondence between phonology and syntax (itself semantics-driven) is illustrated by striking tonal phenomena such as those exhibited by certain African languages. But such phenomena are relatively rare among languages. Moreover, clitics constitute a remarkable counterexample to this correspondence. This suggests that meaning in languages tends to be dominated by form. Human languages transmit meanings through forms, but when there is a conflict between meaning and form, it is the latter which overrides the former.

Examples

(1) Ben Tey (Dogon (Niger-Congo), Mali) :

a.1. *kùrù dàwòy* (stone all) “all (the) stones”; a.2. *kùrù* □TM*gú* (stone. PROX.INAN) “this stone”;

b.1. *isé: dàwòy* (village all) “all (the) villages”; b.2. *yǎ-m isè* (woman-AN.SG village) “a woman’s village” (Heath & McPherson 2013: 269-275).

(2) Kpelle (Mande (Niger-Congo), Liberia):

a.1. *pK∇rK* « house »; *×bK∇rKi* “the house”; a.2. *Xála* « sheep »; *m×alai* “the sheep”;

b.1. *nūu sàaXa* (man three) « three men »; b.2. *×zàaXa* “three of them”;

c.1. *nūu pΔ=Δ* man old “an old man”; c.2. *×bΔ=Δi* “(s)he is old”;

d.1. *dí ká pár* (3PL POS.PRG come) “they are coming”; d.2. *×gá pár* 3SG.SBJ.POS.PRG come “(s)he is coming”;

d'.1. *dí fé pání* (3PL.NEG.PRF come) “they didn’t come”; d'.2. *×vé pání* (3SG.SUBJ.NEG.PRF) come “(s)he didn’t come” (Manessy 1964 : 120-121).

(3) Kiyaka (Bantoid (Niger-Congo), Congo Democratic Republic):

a.1 *bà-tàatà bà-na bà-vuumbuk-ídí* (1.CL.PL-father 1.CL.PL-DEM 1.CL.PL-stand.up-REC.PST) “these fathers stood up”;

a.2 *tù-à-vùumbùk-rdr ha-tháangi* (1PL-NON.REC-stand.up-PST 16.CL-bed) “we stood up from bed”;

b.1. *bíimá byá-bá-táata* DEM(1.CL.PL) of(1CL.PL)-1.CL.PL-father) “these are properties of fathers”;

b.2. *betu tu-a-vúumbùk-ídí ha-thaangi thweel-é* (1PL 1PL.-NON.REC-stand.up.REL-PST 16.CL-bed leave-PAST) “we who stood up from bed, we left” (Van den Eynde & Kyota 1984: 68-70).

(4) Mukri (Central Kurdish, Iranian (Indo-European), Iran)

a. *xezîne-î pādšā=yān tātān kird-bū*
treasure-CONNECT king=3PL.AGT plundering do.PSTPFCT.COP
“they had plundered the king’s treasure”;

b. *bo=m gērā-y-î-yewe*
to-1SG.AGT narrate.PST-LIG-2SG.ATTR-PFCT
“I narrated (it) to you”;

c. *birā-le sē pirsyār=t lē de-ke-m*
brother-VOC three question-2SG from INDIC-do.PRST-1SG
“brother, I’ll ask you three questions”

d. *ser-yān we der de-hēnā* (head-3PL.A~POSS towards out INDIC-bring.PST) “they_i took their_i heads out” (Öpengin 2013: 26, 116, 215, 218).

References

- Heath, J. & L. McPherson, 2013, “Tonosyntax and reference restrictions in Dogon NPs”, *Lg.* 89, 2:, 265-296.
- Manessy, G., 1964, “Détermination et prédication en kpelle”, *BSL*, 59, 1, 119-129.
- Öpengin, E., 2013, *Clitic-affix interactions in Mukri (Central Kurdish)*. Phd Diss., Paris 3.
- Van den Eynde, K. & K. B. Kyota, 1984, “Procédés tonals de construction syntaxique en yaka (Zaire) », *Wiener Linguistische Gazette*, Institut für Sprachwissenschaft der Universität Wien, 68-73/144

The Basic Word Order Typology: An Exhaustive Study

Harald Hammarström

MPI for Psycholinguistics, Nijmegen

One of typology's most celebrated themes is the so-called *basic word order* typology, popularized by (Greenberg 1963) in a study comprising 30 languages. Since then, basic-word-order statistics from ever wider arrays of languages have been presented (Dryer 2005, Haarmann 2004, Hawkins 1983, Tomlin 1986). We are now in a position to present data from 5 112 languages, which exhausts the set of attested languages for which there exists a published description with enough information to deduce the basic word order.

We analyse the distribution of word order frequencies as composed of three factors: A *universal* (U) preference, *genealogical* (G) inheritance, and *areal* (A) diffusion. Any remaining variance is thus attributable to chance. Using orthodox sampling procedures, we have been unable to find any significant correlations between word order type and some other variable, such as population size, as has been suggested using unorthodox sampling techniques (Nettle 1999:139, Trudgill 2011:100-101).

Using known family relationships and language locations (from `glottolog.org`), we may test the size of the influence of the three factors U/G/A (in particular, given the density of the word order data, for the first time we can do detailed tests on areal diffusion). For any given hypothesis of the relative influence of the three factors U/G/A, we may measure how well it matches the actually attested word order data. These tests reveal a weak but consistent role for a universal tendency that favours verb-finality and disfavours object-initiality, fully consistent with previous work (Song 2012). Contra Dunn et al. (2011), the tendency is consistent within families.

References

- Dryer, M. S. (2005). Order of subject, object, and verb. In Comrie, B., Dryer, M. S., Gil, D., and Haspelmath, M., editors, *World Atlas of Language Structures*, pages 330–333. Oxford University Press.
- Dunn, M., Greenhill, S. J., Levinson, S. C., and Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 13 April:1–4.

- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. H., editor, *Universals of language: report of a conference held at Dobbs Ferry, New York, April 13-15, 1961*, pages 73–113. Cambridge, Massachusetts: MIT Press.
- Haarmann, H. (2004). *Elementare Wortordnung in den Sprachen der Welt: Dokumentation und Analysen zur Entstehung von Wortfolgernustern*. Hamburg: Helmut Buske.
- Hawkins, J. A. (1983). *Word order universals*, volume 3 of *Quantitative Analyses of Linguistic Structure*. San Diego: Academic Press.
- Nettle, D. (1999). *Linguistic Diversity*. Oxford University Press.
- Song, J. J. (2012). *Word Order Typology*. Cambridge: Cambridge University Press.
- Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.
- Trudgill, P. (2011). *Sociolinguistic Typology*. Oxford: Oxford University Press.

Pragmatic demarking of clefts: When and where

Corinna Handschuh

Christian Rapold

Universität Regensburg

While clefts are known to be at the origin of various grammaticalisation processes (Heine & Reh 1984, Harris & Campbell 1995, Harris 2001, 2002, 2003), the precise paths of these developments are still not understood. One such path leads from a biclausal construction with contrastive focus to a pragmatically unmarked monoclausal structure, supplanting the older unmarked construction. But exactly where, how and when this process starts and what stages it encompasses, as well as what the general tendencies are and what amounts to language specific variation remains largely uncharted territory.

In this talk, we aim at answering part of these questions by looking at those contexts in various languages where erstwhile clefts have become the obligatory, pragmatically unmarked construction for the expression of certain states of affairs, i.e. the contexts where the demarking process has been completed. More specifically, we approach the question by focusing on unexpected morphosyntactic markers that are traces of this grammaticalisation process.

The Perfect tense in Wolof (Atlantic, Niger-Congo), for instance, has been hypothesised to derive from an earlier cleft construction (Church 1981: 68 i.a., cited in Robert 1991: 36). This tense, while still a periphrastic form, functions like a single conjugation and cannot be cut up synchronically (Robert 1991: 36). In Kinyamwezi (Bantu, Niger-Congo), by contrast, the same process seems to have gone a few steps further: all the affirmative past and perfective tenses except the Remote Past have been said to derive from relative verb forms (Schadeberg 1989: 36, 39), the likely vestiges of erstwhile clefts. Drawing on data from a range of genetically and areally unrelated languages we explore the possibility of setting up a hierarchy of semantic, pragmatic and morphosyntactic contexts with respect to how conducive they are to the pragmatic demarking of clefts.

This pragmatic demarking typically goes hand in hand with formal restructuring and erosion, often leaving typological peculiarities in the resulting structure. Examples of such traces of former clefts include the following:

- agreement-markers on the final verb in most types of declarative sentences in Benchnon (Omotic, Afroasiatic) can be traced back to relative pronouns and even further back to a set of demonstratives, whereby the verb in the relative clause now seems to agree solely with its head noun (Rapold 2007)
- case-marking of predicate nominals is used for the the pragmatically salient argument (roughly corresponding to focus and/or contrastive topic) in a number of Nilotic (Nilosaharan) and Malayo-Polynesian (Austronesian) languages, often this element occurs in preverbal position in otherwise verb-initial languages (König 2008, Handschuh 2014)
- the Agar variety of Dinka (Nilotic) a similar structure appears to have grammaticalised to the unmarked sentence type. This language is analyzed as a topic-first language, the verb being in second position (Andersen 1991).

The insights gained from an in depth analysis of examples like these will allow for a more fine-grained understanding of the processes and stages of the grammaticalisation of clefts.

References

- Andersen, Torben. 1991. Subject and topic in Dinka. *Studies in Language* 15(2). 265–294.
- Church, E. 1981. *Le système verbal du Wolof* Coll. Documents linguistiques. Dakar: University of Dakar.
- Handschuh, Corinna. 2014. *A typology of marked-S languages*, vol. 1 Studies in Diversity Linguistics. Berlin: Language Science Press.
- Harris, Alice C. 2001. Focus and universal principles governing simplification of cleft structures. In Jan Terje Faarlund (ed.), *Grammatical relations in change*, 159–170. Amsterdam, Philadelphia: John Benjamins.
- Harris, Alice C. 2002. *Endoclitics and the origins of Udi morphosyntax*. Oxford: Oxford University Press.
- Harris, Alice C. 2003. Cross-linguistic perspectives on syntactic change. In *Handbook of historical linguistics*, 529–551. Oxford: Blackwell.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*, vol. 74 Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- König, Christa. 2008. *Case in Africa*. Oxford: Oxford University Press.
- Rapold, Christian J. 2007. From demonstratives to verb agreement in Benchnon: A diachronic perspective. In Azeb Amha, Maarten Mous & Graziano Savà (eds.), *Omotic and Cushitic studies. papers from the fourth Cushitic-Omotic conference, Leiden, 10-12 April 2003*, 69–88. Köln: Rüdiger Köppe.
- Robert, Stéphane. 1991. *Approche énonciative du système verbal: Le cas du Wolof* Sciences du langage. Paris: Centre National de la Recherche Scientifique.
- Schadeberg, Thilo C. 1989. Tone and history of Nyamwezi verb forms with complex final tones. *Afrika und Übersee* 72(1). 33–42.

The serial verb construction: Comparative concept and cross-linguistic generalizations

Martin Haspelmath

Since the 1970s, serial verb constructions (SVCs) have been discussed widely in African, Oceanic and many other languages in different parts of the world. This paper gives an overview of the most important generalizations about serial verb constructions that have been proposed and that do seem to hold if a sufficiently restrictive definition of the concept is adopted. The main problem of the earlier comparative literature is that the notion of a SVC has not been delimited clearly, and/or has been formulated in much too wide terms. This can be seen as a result of the confusion between comparative concepts and natural kinds: Serial verb constructions have (most often implicitly) been regarded as natural kinds (universal categories), so that phenomena in additional languages were regarded as SVCs even when they had somewhat different properties. This procedure inevitably leads to a fuzzy and very broad understanding of the concept, with a prototype (or “canonical”) structure that does not allow falsifiable claims. Here I propose a narrow definition of SVC and formulate 10 universals (cf. below) that are apparently true of all serial verb constructions in this narrow sense.

1. In all SVCs, the verbs have the same tense value.
2. In all SVCs, the verbs have the same mood value.
3. The verbs in a SVC do not have separate temporal or event-locational modifiers.
4. All SVCs are pronounced with a single intonation contour, like single-verb clauses.
5. If a SVC expresses a cause-effect relationship, or a sequential event, the order of the two verbs is tense-iconic, i.e. the cause verb precedes the effect verb, and the verb that expresses the earlier event precedes the verb that expresses the later event.
6. If there is just a single person, tense, mood or negation marker, it occurs in a peripheral position, i.e. preceding the first verb or following the last verb.
7. In all SVCs, all the verbs share at least one argument.
8. All languages with SVCs have same-subject serial verb constructions, possibly along with other types.
9. In different-subject SVCs, the second verb is always intransitive.
10. A SVC cannot have two different agents, i.e. when a non-agent is shared, then the agent must be shared as well.

Typology and extreme languages

Jeff Heath (Michigan)

jheath@umich.edu

Crosslinguistic study can be done in two styles. The Greenbergian style that dominates "typology" focuses on statistical correlations among values of universal categories. It relies on the codability of such categories, and on the assumption that a language sample innocent of extraneous skewing can be constructed. Given mounting evidence for geographical and genetic skewing and growing doubts about the universality of categories, institutionalized "typology" is not intellectually viable as a distinctive subfield (journals, associations, departments) in the long term. The history of older fields like biology and sociocultural anthropology, which passed through classificatory and typological phases, may be instructive. One option is to merge with historical linguistics. A more interesting one is to infiltrate "theoretical" linguistics in the hope of creating a scientifically responsible synthesis. This entails shifting to a Whorfian style, focusing on direct confrontations of English/SAE with extreme languages that have been insignificant statistical outliers in the Greenbergian style, and examining entire grammatical systems rather than individual features. Deep nonconfigurationality, even though rare and (therefore) usually unmentioned in typology textbooks, becomes central in this theoretical turn.

COMPARATIVE AND DIVERSITY LINGUISTICS: WHERE NEXT?

What does the future hold for comparative and 'diversity linguistics'? This talk explores some of the trends that seem most likely, over the years to come, to shape both sides of our field: historical/comparative linguistics, and typology/universals.

One obvious shift is the growing prominence of new and predominantly quantitative methodologies brought in from outside linguistics, in particular from the biological sciences. These are not linguistic analyses proper, of course, but powerful tools for 'number crunching' analysis of *any* comparative data — once converted into numerical or state format, in large-scale databases. But to what extent might these inherently non-linguistic tools end up leading the linguistics itself?

Especially in a context where the threat of language extinction is becoming ever more a reality, are language documentation efforts, for example, destined to be refocused towards 'big data' ends? Major typological databases now in preparation aspire to analyses and data-sets that are as exhaustive as possible across the world's language diversity — but can thus end up correspondingly shallower in the analysis of any given language. What are the pros and cons, and what will the effects be on linguistics? And what place might there be for documentation of another form, focused on that other dimension of vanishing linguistic diversity: at the dialectal level within any one family, with all that that can tell us of its origins?

In other ways too, diversity linguistics will surely not be immune to the transformations that are revolutionising other disciplines. Genome-wide and ancient DNA analyses, along with similar great strides in archaeological science, are together vastly enriching our understanding of the relationships of descent, contact and migrations between the populations that speak (or spoke) the language lineages we study.

How can diversity linguistics adapt to, and best join forces with, these rapidly changing fields? As one illustration, consider how cross-disciplinary approaches to language data have until now focused so heavily on just one side of the comparative linguistic coin: language *families* and their 'phylogenetic' divergence through space and time. Is it not high time that linguistics developed a more structured and principled approach towards the (pre)historical contexts of linguistic *areas* and their convergence, too?

Exemples

- (2) a. **Il a goûté la **prison** dans sa jeunesse.*
'He tasted the prison in his youth.'
- b. *Il a goûté à la **prison** dans sa jeunesse.*
'He tasted the prison in his youth.'
- (3) a. **Il a fouillé ses **souvenirs**.*
lit. 'He searched his memories.'
- b. *Il a fouillé dans ses **souvenirs**.*
'He searched in his memories.'
- (4) a. *Ils fouillent le **bois** à la recherche du fugitif.*
'They searched the woods to find the fugitive.'
- b. *Ils fouillent dans le **bois** à la recherche du fugitif.*
'They searched in the woods to find the fugitive.'
- (5) a. *Le garçon a fouillé ses **poches**.*
'The boy searched his pockets.'
- b. *Le garçon a fouillé dans ses **poches**.*
lit. 'The boy searched in his pockets.'

References

- MELIS, Ludo. 1990. *La voie pronominale : la systématique des tours pronominaux en français moderne*. Paris / Louvain-la-Neuve : Duculot. (Coll. Champs Linguistiques).
- HERSLUND, Michael. 1997. « Passive and Antipassive in a Functional Description of French Reflexive Verbs ». In *Journal of Linguistics*. 19. Hermes. pp. 75-92.
- HERSLUND, Michael. 2007. « L'antipassif réfléchi en français et en danois ». In ROUSSEAU André, BOTTINEAU Didier and Daniel ROULLAND (eds.). *L'énoncé réfléchi*. Rennes : Presses Universitaires de Rennes. (Coll. Rivages Linguistiques). pp. 175-186.
- HERSLUND, Michael. 2008. « Le français comme langue exotique: Approche typologique ». In RAEMDONCK, Dan van (ed.). *Modèles syntaxiques : la syntaxe à l'aube du XXIe siècle*. Bruxelles / Bern / Berlin [etc.] : P.I.E. Peter Lang. (Coll. Gramm-R. 2). pp. 19-42

A Semi-Automatic, Semi-Supervised Method for Small Data Document Clustering

Andres Karjus, Liisi Veski

University of Tartu

Document clustering, topic modeling and information retrieval as a field in general has made considerable progress in the past few decades. However, most of its statistical, largely unsupervised methods are developed with large datasets of reasonably large documents in mind. While there have been advances in applying such methods for clustering shorter texts like tweets (cf. Guo, Diab 2012), operating with limited training data remains an open challenge. This study presents one solution, applicable in situations where relatively little textual data are available, yet the data is too big for traditional qualitative analytic methods, such as close reading, or where additional computational insight is desired. Although the case study is nested within the framework of history of ideas, the methods discussed are applicable to linguistic research as well, particularly questions regarding semantics and context. While language-independent in principal, this approach does require certain digital resources (a lemmatizer and a Wordnet) to be available for a given language. The prototype (written in the R language (R Core Team 2014)) being used for this research is intended to be developed into a publicly accessible open-source library.

The case study focuses on the concepts of *nationalism* and *nation* in the 1930's, a time where such matters were hotly debated across Europe. The data consists of a set of essays by the Estonian scholars and politicians of the time, recently made digitally available. The study seeks to map the diversity in the understanding and usage of these two concepts, as exemplified by their contexts (the company they keep). Also, can meaningful clusters of authors and/or texts be formed by computational methods, and could they add insight to the already present knowledge in the domain of history? To this end, the essays are mined for small windows of text (1-3 sentences) containing said terms. Therein lies the challenge, as the resulting dataset is fairly small, compared to what generally well-performing topic modeling algorithms would expect.

One way to tackle the small data challenge is to leverage various forms of annotations, be it semantic or morphosyntactic, in order to aid the automatic clustering or topic modeling process. Considerable research within information retrieval has focused on using the Wordnet databases (currently available for over 70 languages) for such purposes (cf. Wei et al 2015 for an overview). The present study follows suit, but from a slightly different angle. Each non-stopword in the mined word vectors is automatically lemmatized linked to its matching synset in the Estonian Wordnet, followed by K-means clustering of the words, using their hyperonymy links as a measure of similarity (using the metric of Wu, Palmer 1994). This allows "generalizing" the word vectors, as words are replaced with their arguably more general hyperonyms (the K does need to be manually tuned for reasonable results), reducing data sparseness and allowing to cluster the text by their more general topics, without having exercised any (more data-hungry) automatic topic modeling. TFIDF weights are then used to score inter-document similarity. The current proof-of-concept program (evaluated on a subset of manually clustered documents) outperforms a baseline TFIDF model as well as LDA (Blei et al 2003). A number of problems do occur, such as imperfect automatic lemmatization, word sense disambiguation for Wordnet synset matching (due to the small dataset, disambiguity is currently resolved manually) and selecting for an optimal K; these will be adressed in the presentation.

References

Blei, D., A. Ng, and M. Jordan. 2003. Latent Dirichlet allocation. Journal of Machine Learning Research, 3:993–1022.

R Core Team 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Wei, Tingting, Yonghe Lu, Huiyou Chang, Qiang Zhou, Xianyu Bao, A semantic approach for text clustering using WordNet and lexical chains, Expert Systems with Applications, Volume 42, Issue 4, March 2015, Pages 2264-2275.

Guo, Weiwei, Mona Diab. 2012. Modeling sentences in the latent space. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1 (ACL '12), Vol. 1. Association for Computational Linguistics, Stroudsburg, PA, USA, 864-872.

Wu, Zhibiao and Martha Palmer. 1994. Verbs semantics and lexical selection. In Proceedings of the 32nd annual meeting on Association for Computational Linguistics (ACL '94). Association for Computational Linguistics, Stroudsburg, PA, USA, 133-138.

ALASKAN RUSSIAN: A SPECIAL RESULT OF LANGUAGE CONTACT

Mira B. Bergelson (Higher School of Economics, Moscow, Russia)

Andrej A. Kibrik (Institute of Linguistics RAN and Moscow State University, Moscow, Russia)

This paper is intended to contribute to a typology of language contact. Alaskan Russian is a remnant dialect of Russian, still spoken in Alaska by a small number of elderly individuals and representing the variety of Russian that developed in Alaska in the 18th-19th centuries.

Russia colonized a vast territory during the same historical period as the major Western European powers expanded their influence all around the world. The difference of Russian colonization was that it was primarily terrestrial, as opposed to the maritime expansion of the Western countries. As a result, Russian “interacts with over a hundred other languages, many of them genetically totally unrelated to Russian” (Comrie 1987: 322). Contact phenomena resulting from such interaction in Eurasia are summarized by Perexval'skaja (2008).

One major exception to the mostly Eurasian spread of the Russian influence is Russian America, an overseas colony possessed by Russia from mid-18th century till 1867. See Krauss 1996 on the kinds of Russian linguistic influence visible in the present-day Alaska. The best known language that evolved from a contact between Russian and a native American language is Copper Island Aleut, see e.g. Golovko 1997. A very different linguistic variety is still present in the village of Ninilchik, established by the Russian American Company in the Kenai peninsula in mid-19th century for the company's retirees and their families. With the exception of several ethnically Russian men, the Ninilchik community from the very beginning consisted of creoles – the officially recognized social class of people of a mixed Russian-native descent. The native component of the gene pool mostly came from the Alutiiq Eskimos, although it is unlikely that full blood Eskimos were ever found among the Ninilchik residents. Ninilchik Russian existed as an isolated Russian variety for many decades, and had very little contact with any other form of Russian for over 100 years.

Ninilchik Russian was studied by Conor Daly, who pointed out the dramatic change in the category of gender as compared to standard Russian; see the title of his unpublished paper Daly 1986. Later studies include Kibrik 1998, Bergelson and Kibrik 2010, Golovko 2014, Bergelson et al. in progress. Many features of Alaskan Russian can be attributed to a contact with native American languages, through the Russian-native bilingualism of the creoles during the Russian America time.

There are many indications to the fact that Ninilchik was not the only location where Alaskan Russian was used. We believe that Alaskan Russian, as we see it in modern speakers, developed as a special variety of Russian belonging to the class of creoles throughout all of Alaska during the Russian America time, and Ninilchik is just one of the very few pockets where it survived after the transfer of Alaska to the USA in 1867. To give just one example, the common Russian word *хлеб* [xl'ep] ‘bread’ is mostly known in Ninilchik Russian in the originally genitive form *hl'éba*. The reason is that in colloquial language this lexeme is most commonly used in the genitive: *kusók hl'éba* ‘a piece of bread’, *day hl'éba* ‘give (me) some bread’, *n'et hl'éba* ‘there is no bread’, etc. As a result, the genitive form has become conventionalized in many speakers as the basic one. But how old is this development? An answer is offered by the evidence from native Alaskan languages that only had contact with Russian in the 19th century. In Upper Kuskokwim, an Athabaskan language spoken many hundreds miles away from Ninilchik, the loan word *hlema* ‘bread’ has been registered, which corroborates the early and geographically non-local character of the innovation in question.

In this paper we review a number of innovated features of Alaskan Russian, in its Ninilchik incarnation, including:

- Phonetics: phoneme inventory, phonetic realizations of consonants, tempo of speech, and other prosodic features
- Morphosyntax: gender, case, aspect
- Lexicon: loan words, semantic patterns.

To the extent possible, we trace these peculiarities of Alaskan Russian to the special kind of language contact that took place in the 19th century Russian America and led to the formation of this unique variety of Russian.

References

- Bergelson, Mira B., and Andrej A. Kibrik. 2010. The Ninilchik variety of Russian: Linguistic heritage of Alaska. In: A. Mustajoki, E. Protassova, N. Vakhtin (eds.) *Instrumentarium of linguistics: Sociolinguistic approaches to non-standard Russian*. Slavica Helsingiensia, Vol. 40. Helsinki: University of Helsinki, 320–335.
- Bergelson, Mira B., Andrej A. Kibrik, and Wayne Leman. Ninilchik Russian dictionary. In progress.
- Comrie, Bernard. 1987. Russian. In: B. Comrie (ed.) *The world's major languages*. London and Sydney: Croom Helm, 329–347.
- Daly, Conor. 1986. *Evonaj mat' ves' noc television karaulil – His mother watched TV all night long: On the loss of gender as a grammatical category in Alaskan Russian*. Ms. 31 pp.
- Golovko, Evgeny V. 1997. Mednovskix aleutov jazyk. In: A.P. Volodin, N.B. Vaxtin, A.A. Kibrik (eds.) *Jazyki mira: Paleoaziatskie jazyki*. Moscow: Indrik, 117–125.
- Golovko, Evgeny V. 2014. Russian as a minority language: A case from Alaskan old-settlers communities. In: V.F. Vydrin and N.V. Kuznecova (eds.) *Ot Bikina do Bambaljumy, iz varjag v greki. Ekspedicionnye etjudy v chest' E.V. Perexval'skoj*. Saint Petersburg: Nestor-Istorija, 141–154.
- Kibrik, Andrej A. 1998. Nekotorie foneticheskie i grammaticheskie osobennosti russkogo dialekta derevni Ninilchik. In: V.F. Vydrin and A.A. Kibrik (eds.) *Jazyk. Afrika. Ful'be*. Saint Petersburg - Moscow: Evropejskij dom, 36–52.
- Krauss, Michael. 1996. The Russian language in Alaska and in Alaskan native languages. In: S. Wurm, P. Mühlhäusler, D. Tyron (eds.) *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*. Berlin: Mouton de Gruyter, 1209-1212.
- Perexval'skaja, Elena V. 2008. *Russkie pidzhiny*. Saint Petersburg: Aletejja.

A look at Jeju language and ideologies of diversity-oriented linguistic research

Soung-U Kim, PhD candidate
SOAS London, Department of Linguistics

Until very recently, the Korean language had been regarded a language isolate, leading to bold statements in grammars such as Sohn (1999: 12) who states that “[d]espite [...] geographical and socio-political dialectal differences, Korean is relatively homogeneous, with excellent mutual intelligibility among speakers from different areas” (cf. similar statements in Song 2012). It is this talk’s motivation to break with this notorious idea of a ‘proud, homogeneous and monolingual nation’ (as explained in King 1998, Seth 2011: 25), and present a more diverse image of the Korean linguistic landscape by introducing Jeju language spoken in Jeju Province, South Korea to a wider linguistic audience.

Jeju has traditionally been classified as a dialect of Korean (Pangenyenguhoe 2001), similar to the classification of Ryukyuan languages as dialects of Japanese (cf. Heinrich 2012), although many sources mention the unique position of Jeju as being much more conservative and therefore closer to Middle Korean (King 2006, Lee and Ramsey 2011, Stonham 2011, Yeon 2012). In fact, Jeju shares only a seemingly minor part of its lexicon with Korean, and preserves a more complex vowel inventory, a different distribution of periphrastic and morphological constructions, a greater degree of irregularity in nominal case inflection, a different system of pronouns as well as a verbal morphology with its very own politeness and evidentiality system (see Kang 2007, or Lee 1978).

With these clear differences between Korean and Jeju, it is no wonder that on-going studies (William O’Grady, University of Hawai’i and Changyong Yang, Jeju National University, p.c.) point towards a degree of mutual intelligibility between Korean and Jeju comparable to that of Russian and Polish. Similarly, recent movements such as UNESCO’s classification of Jeju as a ‘critically endangered language’ (Moseley 2010, UNESCOPRESS 2010) have marked a change in ideology motivated by foreign, scientific activities (e.g., Southcott 2013). With intergenerational transmission long interrupted, and with all Jeju speakers over the age of 75, people are slowly joining forces to revitalise the linguistic diversity of this region, trying to counter the devastating effects of nationalist language ideologies that have a tendency to emphasise commonalities between idioms of a region, and mask differences between them (cf. Kroskrity 2000 on language ideologies).

Against this backdrop, I would like to give a sociolinguistic overview on Jeju language as it is spoken in Jeju Province of South Korea today, and present some linguistic evidence on the differences between Korean and Jeju, based on data from my own fieldwork on Jeju Island. In a similar vein, I would like to raise awareness for linguistic diversity in South Korea and East Asia in general, arguing that the scientific practice of linguists can and should be the expression of our own diversity-oriented language ideologies (Bobaljik et al. 1996). By discussing the Jeju case, I would like to emphasise on our potential of influencing linguistic diversity through the outcomes of our own work as linguists (Duchêne and Heller 2007, Wolfram 2008), and argue that ‘diversity linguistics’ should be more explicitly developed as an ideological stance and practice.

References

Bobaljik , Jonathan David, Rob Pensalfini and Luciana Storto.1996. *Papers on Language Endangerment and the Maintenance of Linguistic Diversity*. Cambridge, MA: MIT Working Papers in Linguistics. Vol. 28.

Bobaljik , Jonathan David, Rob Pensalfini. 1996. Introduction. In Bobaljik et al. (eds.), *Papers on Language Endangerment and the Maintenance of Linguistic Diversity*. Cambridge, MA: MIT Working Papers in Linguistics. Vol. 28.1-24.

Duchêne, Alexandre and Monica Heller (eds.). 2007. *Discourses of endangerment: ideology and interest in the defence of languages*. London /New York: Continuum.

Heinrich, Patrick. 2012. *The making of Monolingual Japan : language ideology and Japanese modernity*. Multilingual Matters.

Kang, Young-Bong. 2007. *Ceycwue* [Jeju language]. Seoul: Kwuklipminsokpakhmulkwon [National Folk Museum], Jeju Special Self-Governing Province. Volume 1.

King, Ross. 1998. Nationalism and Language Reform in Korea: The Questione della Lingua in Precolonial Korea. In Pai, Hyung Il and Timothy R. Tangherlini (eds.). *Nationalism and the construction of Korean identity*, 33-72. Berkeley, CA: Institute of East Asian Studies, University of California.

King, Ross. 2006. Dialectal Variation in Korean. In Sohn, Ho-Min (ed.), *Korean Language in Culture and Society*, 264-280. Hawai'i: University of Hawai'i Press.

Kroskrity, Paul V. (ed.). 2000. *Regimes of language: ideologies, politics and identities*. Santa Fe: Oxford: School of American Research Press. 35-84.

Lee, Ki-Moon and Samuel R. Ramsey. 2011. *A History of the Korean Language*. Cambridge: Cambridge University Press.

Lee, Sung-Ryeong. 1978. *Ceycwuto Pangenuy Hyengtaeloncek Yengwu* [Morphological study of Jeju dialect]. Seoul: Thap.

Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO Publishing. Online version: <http://www.unesco.org/culture/languages-atlas/en/atlasmap.html> [accessed on 2014-08-29]

Pangenyenguhoy (ed.). 2001. *Pangenhak Sacen* [The Dictionary of Dialectology]. Seoul: Thaehaksa.

Seth, Michael J. 2011. *A History of Korea - From Antiquity to the Present*. Plymouth: Rowman and Littlefield Publishers.

Sohn, Ho-Min. 1999. *The Korean Language*. Cambridge University Press.

Song, Jae Jung. 2012. South Korea: language policy and planning in the making. *Current Issues in Language Planning* 13(1). 1-68.

Southcott, Darren. 2013. Jeju language must be saved, says linguist. Jeju Weekly Online, [last updated 2013-11-08, 11:32:05], accessible under <http://www.jejuweekly.com/news/articleView.html?idxno=3659> [accessed 2014-08-29]

Stonham, John. 2011. Merger of Vowels in Cheju Korean: Methodological implications for doing fieldwork on endangered languages. *The Linguistic Association of Korea Journal* 19(1). 23-40.

UNESCOPRESS. 2010. New interactive atlas adds two more endangered languages. Available online at http://www.unesco.org/new/en/media-services/single-view/news/new_interactive_atlas_adds_two_more_endangered_languages/#.VAGjgGNuW3M

Wolfram, Walt. 2008. Language diversity and the public Interest. In Kendall King, Natalie Schilling-Estes, Jia Jackie Lou, and Barbara Soukup (eds.), *Sustaining Linguistic Diversity: Endangered and Minority Language and Language Varieties*. Washington, DC: Georgetown University Press. 187-202.

Yeon, Jaehoon. 2012. Korean dialects: a general survey. In: Tranter, Nicholas, (ed.), *The Languages of Japan and Korea*. Abingdon and New York: Routledge, pp. 168-185.

Salience and transitivity

Seppo Kittilä (University of Helsinki)

The past few decades have witnessed a clear focus on functionally and/or semantically based definitions of transitivity. Probably the most seminal of these has been the list of transitivity parameters suggested by Hopper and Thompson (1980), but similar definitions have been proposed also by Givón (1995: 76) and Næss (2007) among numerous others. These definitions have in common that features such as affectedness and volitionality are seen as central for transitivity. In addition, pragmatic aspects (e.g. topicality and information structure) have also been considered in transitivity studies (see, e.g., Thompson 2002). Formally, the number and case (or similar) marking of arguments along with verbal cross-reference are seen as the main mechanisms of transitivity.

In my presentation, a somewhat different definition will be proposed, even though the proposed definition shares features with other functional-typological approaches to transitivity. The argument put forward in this presentation is that linguistic transitivity is best explained by (the lack) of salience; salient events are coded by highly transitive constructions, while a lower degree of salience results in a decrease in linguistic transitivity. Salience is explicitly noted also by Givón (1995: 76), but he does not elaborate on the notion in any detail. In this talk, a closer look at salience will be taken. A salient event is defined as an event that meets the following criteria:

1. Salient events have a clear starting point and they lead to an observable result
2. Salient events and their result can be observed by others than the agent (and patient) him/herself
3. The agent is aware of the event him/herself
4. A salient event has a definite agent and a definite patient and both participants are relevant
5. The speaker has direct evidence for the denoted occurrence of an event.

Any deviation from this prototype may result in a transitivity decrease. In my talk, I will discuss a range of different transitivity alternations explainable by salience. These include inherently less transitive events, non-realized events (failed action, negation, habituals, desideratives, irrealis mood), less affected patients, less voluntary/less salient causation and aspects of animacy/definiteness. As noted above, the explanation proposed in this paper has features in common with earlier transitivity definitions. The main difference to earlier accounts of transitivity lies in the fact that salience alone suffices for explaining all the major transitivity alternations; it, for example, enables us to combine semantic and pragmatic approaches to transitivity.

References

- Givón, Talmy. 1995. *Functionalism and grammar*. Amsterdam/Philadelphia: John Benjamins.
- Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language* 56: 251-299.
- Næss, Åshild. 2007. *Prototypical transitivity*. Amsterdam: Benjamins.
- Thompson, Sandra A. 2002. 'Object complements' and conversation: towards a realistic account. *Studies in Language* 26(1): 125-163.

Systems and Uses of Articles: Patterns of Variation

Binary systems of (indefinite and definite) articles are specifically found in Romance, Germanic, Balto-Slavic and the Balkan languages and are frequently considered as an areal feature of European languages (Haspelmath, 2001; van der Auwera, 2011). As is shown in Dryer (WALS, 2005), however, this areal restriction only holds true, if the syntactic and semantic properties of the fully grammaticalized article systems of the European languages listed above are more or less taken as a standard for the identification of article systems elsewhere. Subsuming also affixes and demonstratives with a double (gestural and non-gestural) use under the term *definite article*, Dryer shows that outside of Europe ‘definite articles’ in this larger sense are also frequently found in central Africa, in the Pacific and in Mesoamerica, even if the semantic opposition of the two relevant forms is frequently described as ‘specific vs. non-specific’ (Mosel & Hovdhaugen, 1992; Moyse-Faurie, 1997) or not at all. Defining the extension of definite articles in this way, however, raises the question whether the relevant expressions still share any (or enough) semantic properties to be subsumed under this category or ‘comparative concept’ (cf. cf. Davis et al. 2014).

It is the main goal of this paper to discuss parameters and patterns of variation in the use of ‘definite and indefinite articles’ in those languages to which this comparative concept may usefully be applied on the basis of both syntactic and semantic criteria. For ‘definite articles’ these parameters of variation can *inter alia* be specified in terms of distribution, of semantic properties of the accompanying noun and in terms of their relationship to and degree of differentiation from demonstratives. A fine-grained comparison (micro-typology) between languages along these parameters allows us to distinguish extremely rich systems of article use, as found in Greek, from various degrees of more restricted use (French > German > English) down to incipient systems, as found in spoken Finnish and in Sorbian, where definite articles are developing under the influence of neighboring Germanic languages (cf. Chesterman, 1991; Heine & Kouteva, 2006).

Going beyond this fine-grained typology of definite articles in primarily European languages, I will also look at the larger picture, where the term is used much in the sense of Dryer and includes the article system of Polynesian languages as well as a variety of instances mentioned in descriptive grammars outside of Europe. In the course of this extended comparison the following questions will be addressed and answered:

- What kinds of subdivisions can we make on semantic grounds in such a larger set of somewhat heterogeneous phenomena (+/- specific, two types of definiteness, etc.)?
- What exactly is it that changes in the processes of grammaticalization postulated by Greenberg (1978) and Hawkins (2004)? What is part of the core meaning of definite articles? (Cf. Lyons, 1999; Gisborne, 2012; Elbourne, 2012)

References

- Davis, Henry, Carrie Gillon & Lisa Matthewson (2014). “How to investigate linguistic diversity: Lessons from the Pacific Northwest.” <http://ling.auf.net/lingbuzz/001985>.
- Dryer, Matthew (2005) “Definite articles”, in Haspelmath et al. (eds.) *WALS*. Oxford: OUP, 154-157.
- Elbourne, Paul (2012). *Definite Descriptions*. Oxford: Clarendon Press.
- Haspelmath, Martin (2001). “The European Linguistic Area: Standard Average European”, in Haspelmath, Martin et al. (eds.) (2001) *Language Typology and Language Universals*, vol. 2. Berlin: de Gruyter.
- Hawkins, John (2004). *Efficiency and Complexity in Grammars*. Oxford: OUP.
- Heine, Bernd & Kouteva, Tania (2006). *The Changing Languages of Europe*. Oxford: Oxford University Press.
- Lyons, Christopher (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Mosel, Ulrike & Hovdhaugen, Even (1992). *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- Moyse-Faurie (1997). *Grammaire du Futunien*. CDP Nouvelle Calédonie.

Antonyms and word-level negation

Maria Koptjevskaja Tamm & Matti Miestamo

Typological research on negation has focused most prominently on standard negation, i.e. the basic negation strategies in declarative clauses, and some work has also been done on other aspects of clausal negation as well as on indefinite pronouns in the scope of negation. Negation at the level of words, i.e., derivational affixes expressing negation as well as case markers with negative semantics, has so far not figured in systematic typological studies, but it has received some attention in theoretical literature on semantics and morphology. Zimmer (1964) discusses “affixal” negation primarily in English and a couple of other Indo-European languages, but also comments on a few non-Indo-European languages and even suggests some cross-linguistic generalizations. Subsequent work (e.g., Horn 1989) is similarly restricted in its cross-linguistic scope.

From the semantic point of view, the issue of word-level negation is closely connected to antonymy. Antonymy and types of opposition have been a popular topic in semantic theories (see Horn 1989), where the central distinction is between contrary and contradictory opposites. The two types differ as to whether they allow a third possibility in-between: contradictory opposites are either-or (*dead* vs. *alive*), whereas in contrary opposites there is a middle ground between the two poles (*small* vs. *big*). Linguistically, antonyms can be expressed by unrelated lexemes (lexical antonyms) like the examples cited above, or by means of overt negation (*happy* vs. *unhappy*, *possible* vs. *impossible*). Lexical and morphological antonymy do not necessarily exclude each other. E.g., Russian has regular triads of the kind *bol'soj* ‘big’ – *malen'kij* ‘little’ – *nebol'soj* ‘NEG-big’, and even tetrads, such as *dobryj* ‘kind’ – *zloj* ‘mean’ – *nedobryj* ‘NEG-kind’ – *nezloj* ‘NEG-mean’. Despite all the attention that antonymy has received from semanticists, work in a broader cross-linguistic comparative perspective is lacking.

This talk presents a pilot study of antonymy and its expression by both lexical and overt morphological means. We will focus on antonymy in property words (adjectives), more specifically in such forms that can be used as adnominal modifiers. Our main interest will be in finding correlations between semantic and formal properties of antonyms. From the formal point of view, we will pay attention to the type of marking (e.g., prefix vs. suffix), to the number of different word-level negators in a language, whether these markers can be used on other word classes than property words and how they are related to other negative markers in the language. Taking in semantics, we will observe what types of opposition (contrary vs. contradictory, scalar vs. non-scalar etc.) and which domains of property scales (evaluation, size, dimension, temperature etc.) are expressed by lexical antonyms vs. each attested type of overt morphological marking, i.e. whether the linguistic evidence allows us to classify antonyms into cross-linguistically relevant types. Does the existence of a lexical antonym exclude the possibility of morphological marking? Do the markers exclude one another on the same lexical item? Are there semantic principles governing such blocking effects? Can triads and/or tetrads be found in addition to pairs? Our pilot sample includes 15 languages from different families and geographical areas. The data comes from dictionaries and grammars and, most importantly, from a questionnaire sent to language experts. As this is a pilot study of a domain previously unexplored in language typology, our main goal is to sketch different ways of approaching this intriguing domain from a broader cross-linguistic perspective.

NEED-ing HAVE, and BE-ing NEED-less: Turkish evolving away from imposed patterns

In this talk, I address the question (also pursued in some typological and theoretical literature) of the relationship (if any) between a transitive (i.e. Accusative-assigning) verb that corresponds to the English verb *to need*, and a transitive verb that corresponds to the English verb *to have*, where the possessor is in the Nominative, and the possessee is in the Accusative (or, in a language that does not have case morphology, where neither one of the arguments is the complement of an adposition). I then go on to discuss an interesting example of an attempt to artificially impose a *transitive* verb corresponding to *to need* on Turkish, which did not have one, and the relative failure of that attempt.

Harves & Kayne (2008) mention that Isačenko (1974) discusses the counterparts of the verbs *Have* and *Be* in Slavic, and raises the following question: “Could it be sheer coincidence that those Slavic languages which have become H(ave)-languages...have some modal verbs which are unknown to Russian, a B(e)-language?” (Isačenko 1974: 25). Harves & Kayne use the informal term “H-language” to refer to a language that has an overt counterpart to English *have*, in the sense of having an overt verb expressing predicative possession, and they use the term “B-language” to refer to a language that lacks transitive *have* and expresses predicative possession using *be*. They answer the question which Isačenko raised by saying that the correlation he noticed is indeed not a coincidence: They propose that only in H-languages, i.e. in languages that have a transitive verb corresponding to *have*, is it possible to incorporate a nominal *need* into such a verb (or rather into its silent counterpart), thus deriving a transitive verb *to need*. Thus, only those languages that have a transitive verb *need* also have a transitive verb *have*. This implicational relation is uni-directional; i.e. languages that do have a transitive verb *have* may lack a transitive verb *need*. The authors claim that their implicational relation (from transitive *need* to transitive *have*) holds not only in Slavic, but more generally in Indo-European. In later work (Harves & Kayne 2012), they extend their claim to a universal one.

At least since the publication of Harves & Kayne (2012), some counterexamples have come to light that show that the implicational relation they posited can't be an absolute universal. But it does appear to be a very strong tendency. In this talk, I address modern Turkish, which appears to be a counterexample: Its official monolingual dictionary includes a transitive verb corresponding to *need*, i.e. *gereksemek*, which is said to assign Accusative to its complement. Yet, Turkish has no transitive verb corresponding to *have*; it is a B-language, not an H-language.

This verb is a new creation by the Language Academy of Turkey; it was imposed on the language in the context of the language reform undertaken when the Turkish Republic was founded. Neither spoken Turkish, nor written Ottoman had a transitive verb *need*. The result of this prescriptive creation is that of a failure: the new transitive verb is only used in writing, and is found very infrequently even then. Interestingly, another, related, non-transitive (reflexive) verb with essentially the same meaning was created alongside the transitive verb (*gerek^{sin}mek*), which assigns the Dative; that second verb is used more widely in the written language, and can even be found in the spoken language. A third formation, i.e. a nominalized form of that new verb, along with an existential copula, is the one among these three which is used most often: *Possessee + Dative — gerek^{sinim}* (nominalized ‘*need*’) + *Genitive — existential copula*. This mirrors the already-existing expression of needing, i.e. a possessed nominal of *need*, with an existential or general copula. Thus, Turkish is not a counterexample to the generalization by Isačenko and by Harves & Kayne, and does offer some support to Harves & Kayne's proposal to explain that generalization. At the same time, this situation shows that prescriptive attempts at new creations are likely to fail, if typological properties of a language's syntax are not taken into account—in this instance, the fact that Turkish is a B-language rather than an H-language, and thus is likely to reject a *transitive* verb *to need*, while it is more accepting of a *non-transitive* verb *to need*, as well as being likeliest to accept new creations which involve new lexical items, without changing syntactic patterns which are already in existence (such as the nominalization pattern in this instance).

The contribution of (e)WAVE to continent-wide typology

Bernd Kortmann (Freiburg)

The electronic World Atlas of Varieties of English (eWAVE, Version 2.0 released in November 2013), part of the Leipzig family of open-access online atlases, and its companion print volume (Kortmann/Lunkenheimer 2012) offer a range of interesting insights into continent-wide typology since they allow us to separate geographically restricted from much more widespread, variety-type specific features of Englishes (i.e. features characteristic of L1 varieties, L2 varieties, English-based Pidgins or Creoles). In the talk this will be demonstrated for the 29 eWAVE varieties of English in Africa, Australia and the Pacific. Since there are 12 Pidgins and Creoles among these, the talk will also offer a quick comparison with APiCS and, more generally, address the current possibilities and, above all, limitations of cross-survey comparability concerning these two typology-driven electronic surveys of morpho-syntactic variation.

Kortmann, B./K. Lunkenheimer, eds. 2012. *The Mouton World Atlas of Variation in English*. Berlin/New York: De Gruyter Mouton.

Kortmann, B./K. Lunkenheimer, eds. 2013. *The electronic World Atlas of Varieties of English 2.0*. [eWAVE 2.0]. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://ewave-atlas.org>

Kortmann, B. 2014. "Comparison and comparability: The WAVE perspective." In: S. Mergenthal and R. Nischik, eds. *Anglistentag 2013 Konstanz: Proceedings*. Trier: Wissenschaftlicher Verlag Trier. 323-335.

Kortmann, B./V. Schröter. forthcoming. "Varieties of English." In: R. Hickey, ed. *The Cambridge Handbook of Areal Linguistics*. Cambridge: Cambridge University Press.

Typological accommodation in Ket: The case of subordinate clauses

Ket is a polysynthetic language in Central Siberia, the last surviving member of the Yeniseian family, known for being quite complex and hard to pigeonhole within a single typological account (cf. Anderson 2004, Kotorova & Nefedov 2006). The majority of structural features complicating a clear-cut typological analysis of Ket are the result of a peculiar process of structural mimicry, or ‘typological accommodation’ in Vajda’s (2009) terms. Unlike ‘metatypy’ or ‘grammatical calquing’, accommodation does not result in typological replacement, but rather in partial mimicry of a different language type. Due to the long-term contact with the radically different Samoyedic, Tungusic and Turkic languages, the Yeniseian languages have gradually adapted themselves to the structural type of the surrounding languages, while preserving the core features of their grammar that clearly distinguish them from the rest of Central Siberia. The process of typological accommodation in Ket has been described in Vajda (2009) for both the phonological (e.g. phonemic prosody) and morphological levels (e.g. originally prefixing verb structure became restructured as suffixing).

The aim of this paper is to show that this peculiar phenomenon can also be observed at the syntactic level, namely in formation of adverbial and relative clauses.

Unlike prototypical polysynthetic languages that are largely devoid of any overt subordination (cf. Heath 1975, Mithun 1984, Baker 1996), Ket has a number of distinct structural options to signal adverbial subordination through the use of postposed relational morphemes attached fully finite verbs as in (1). This is also a common pattern of forming adverbial clauses shared by the neighboring languages, but they employ non-finite forms (mainly participles) instead of finite verbs (cf. Grišina 1979; Anderson 2003, 2006).

- (1) *bure ū binət-dijte baŋlɔrɔn*
 bude ū b{in}⁷-{b³}-n²-{q}ut⁰-dijte {du⁸}-baŋ⁷-l²-a¹-don⁰
 his strength self⁷-3N³-PST²-finish⁰-N.ADESS 3M⁸-ground⁷-PST²-3CORE¹-fall⁰
 ‘He fell down because he is tired (lit. his strength is finished).’ (Grišina 1979: 40)

Such functional-structural parallel between participles in the surrounding languages and finite verbs in Ket is likewise attested in relative clauses. As shown in Pakendorf (2012), Turkic, Tungusic and Uralic languages share a common relativization pattern involving preposed participial relative clauses with a ‘gapped’ relativized noun phrase. This closely resembles the major relativization pattern in Ket (cf. Nefedov 2012), the only difference being that Ket usually makes use of finite verbs instead of participles as in (2).

- (2) *tāŋ thasa keʔt*
 [tāŋ d{u}⁸-ha⁷-s⁴-a⁰] keʔd
 [hair 3M⁸-PERPENDICULAR⁷-NPST⁴-cut⁰] person
 ‘A man cutting branches’

As we can see, formation of adverbial and relative clauses in Ket clearly imitates that of the surrounding languages and does not conform to the expected ‘polysynthetic’ pattern. At the same time, Ket adverbial and relative clauses resist accommodating a participle-like morphology and remain fully finite which reflects the general tendency among polysynthetic languages not to have truly non-finite forms (like participles, infinitives, converbs) (cf. Nichols 1992, Baker 1996).

This tendency to retain a fully finite verb in subordinate constructions structurally similar to those with non-finite verbs in the other languages of the area is a further evidence in support of Vajda’s (2009) claim about the hybrid nature of the Ket grammatical structure where alongside an overlay of areal features the core features have remained intact.

References

- Anderson, G. (2003): Yeniseic languages from a Siberian areal perspective. In: Edward Vajda and Gregory D. S. Anderson (eds.), *Studia Yeniseica*. (=Sprachtypologie und Universalienforschung 56.1/2). 12–39.
- Anderson, G. (2004): The Languages of Central Siberia: Introduction and Overview. In E. Vajda (ed.) *Languages and Prehistory of Central Siberia*. Amsterdam: John Benjamins, p. 1-119.
- Anderson, G. (2006): Towards a typology of the Siberian linguistic area. In: Yaron Matras, April McMahon, and Nigel Vincent (eds.), *Linguistic Areas – Convergence in Historical and Typological Perspective*, 266-300. Palgrave Macmillan.
- Baker, M. (1996): *The polysynthesis parameter*. New York. Oxford University Press.
- Grishina N. (1979): *Padežnye Pokazateli i Služebnye Slova v Strukture Složnogo Predloženiya*. Unpublished Candidate Degree Dissertation. Tomsk.
- Heath, J. (1975): Functional relationships in grammar. *Language* 51, 89-104.
- Kotorova E. & Nefedov A. (2006): Tipologičeskie kharakteristiki ketskogo yazyka: veršinnoe ili zavisimostnoe markirovanie? In: *Voprosy yazykoznanija*, 2006, № 5, c. 43-56.
- Mithun, M. (1984): How to avoid subordination. *Proceedings of the tenth annual meeting of the Berkeley Linguistics Society*, 493-523.
- Nefedov, A. (2012): Relativization in Ket. In: *Clause Linkage in Cross-Linguistic Perspective. Data-Driven Approaches to Cross-Clausal Syntax*. Eds. Volker Gast and Holger Diessel. Berlin: Mouton de Gruyter, p. 191-224.
- Nichols, J. (1992): *Linguistic Diversity in Space and Time*. Chicago University Press: Chicago.
- Pakendorf, B. (2012): Patterns of relativization in North Asia: towards a refined typology of prenominal participial relative clauses. In: *Clause Linkage in Cross-Linguistic Perspective. Data-Driven Approaches to Cross-Clausal Syntax*. Eds. Volker Gast and Holger Diessel. Berlin: Mouton de Gruyter, p. 253-283.
- Vajda, E. (2009): Yeniseic substrates and typological accommodation in central Siberia. In: Tom Güldemann, Patrick McConnell & Richard A. Rhodes (eds.). *The languages of hunter-gatherers: Global and historical perspectives*. Cambridge: Cambridge University Press.

Pramod Kumar

Why are we linguistically so diverse?

In recent years, there has been an interdisciplinary research attempt to investigate human evolution in terms of its origin, migration, and language. Undoubtedly we have linguistic affinity but there is linguistic diversity also. Linguistic affinity or diversity is assessed through the evidence from population genetics, archaeology, and historical linguistics. It is, moreover, well accepted in the scientific world that modern humans are descendants of a single human evolution, i.e. out of Africa. This raises the question: If we are descendants of a common evolution, why do we have so much linguistic diversity? The present paper makes an attempt to answer the question of linguistic diversity in light of the common origin of language.

On five counter-to-fact grammatical categories

Tania Kuteva, Bas Aarts, Gergana Popova and Anvita Abbi

Abstract

The following three categories were placed on top of a list of categories which are in need of clarification in a newly introduced call-for-contributions rubric “*What exactly is...?*” in a recent issue of the journal *Linguistic Typology* 17 (2013), due to the fact that they are “unclear in current descriptive and theoretical practice”: avertive, frustrative, apprehensional (Plank 2013: 267).

These are the categories we are planning to focus on in the present paper. On the basis of preliminary observations and ongoing research we will propose that the linguistic phenomena usually referred to as the avertive, the frustrative, and the apprehensional belong not to three but to five distinct grammatical categories, all of which present a case of what has been termed elsewhere *semantically elaborate grammatical categories* (Kuteva 2009, 2010). Semantically elaborate grammatical categories can be regarded as semantically “rich” categories since they relate to more than one conceptual-semantic domain simultaneously. Kuteva (2009) exemplifies this kind of category by means of the avertive, a grammatical category recently identified across languages (Kuteva 1998, Kuteva 2001, Heine and Kuteva 2002; cf. also Abbi 1975; 1977; 1980: 42-43 on what she termed the *non-precipitative aspect*). The avertive is used only in past contexts and in Kuteva 1998, Kuteva 2001, Heine and Kuteva 2002, it is treated as a linguistic expression standing for a verb situation which was on the verge of taking place but did not take place (“was on the verge of V-ing but did not V”). In other words, the avertive involves at least three conceptual-semantic domains: temporality (pastness), aspectuality (imminence), and modality (counterfactuality).

That the semantics of the avertive is elaborate – i.e. rich in specificities – becomes clear when we compare the avertive to another grammatical category, which was also identified across languages only very recently, the proximative. The proximative has been noticed in a number of individual languages but has been traditionally considered a specific verb construction rather than a grammatical category. An exception to this practice is Comrie (1976: 64-5) and Comrie (1985: 95), who has not only pointed it out (under the names of *prospective*, and *immediate future*, respectively), but has, moreover, acknowledged that the form in question expresses a grammatical distinction. Later, Heine 1992 identified the proximative (which he first called an “almost”-aspect) as a full-fledged grammatical category across languages. The proximative defines a *temporal phase located close before the initial boundary of the situation described by the main verb*. It indicates a moment shortly before the possible occurrence of the given verbal situation, with (crucially) *no* implication that the situation actually occurred. Yet another essential characteristic of the proximative is that it can be used in both past and non-past contexts. In other words, the proximative is a purely aspectual gram, its essential semantic characteristic being *imminence*. The most obvious argument in favour of treating the avertive as semantically more elaborate than the proximative is that the semantics of the former (past-plus-imminent-plus-counterfactual plus-perfective) subsumes the semantics of the latter (imminent).

Semantically elaborate categories pose problems for linguistic description. We will show that the challenges posed by these categories are best met by adopting the method of intergenetic grammaticalization (for the description and application of this method, see Heine and Kuteva 2002, Heine and Kuteva 2007). Using the method of intergenetic grammaticalization we will propose five semantically related Tense-Aspect-Mood categories, all of which share counterfactual meaning:

- (i) **Apprehensional** – counterfactuality of realization of undesirable verb situation;
- (ii) **Avertive** – counterfactuality of realization of imminent, past verb situation where the verb situation is viewed as a whole (i.e. perfective);
- (iii) **Frustrated initiation** – counterfactuality of realization of initial stage of past verb situation;
- (iv) **Frustrated completion** – counterfactuality of realization of completion of past verb situation;
- (v) **Inconsequential** – counterfactuality of presence/ full degree/ stability of expected result of past verb situation.

Of these it is only the apprehensional that is relatively well-studied (see Vuillermet 2013a, 2013b, in particular).

The other four categories to be investigated in the present study – avertive, frustrated initiation, frustrated completion, and inconsequential – have either not been given any recognition in the specialized literature as grammatical structures at all or they have been subsumed under one and the same cover category, or alternatively, depending on author, there have been proposals to lump various combinations of these categories into different “umbrella” categories. The reason behind this extreme confusion is: (a) The intrinsic challenge to assign a straightforward categorical status to the particular linguistic structures under discussion

here, namely semantically elaborate grammatical categories, and; (b) The lack of consistency in the practice of linguistic categorization in general (depending on the theoretical framework employed, the same set of linguistic facts may receive radically different interpretations).

In the present paper we will argue that using the notion of an abstract prototype or the notion of *Gesamtbedeutung* (core meaning) – in this particular case of what we refer to as counter-to-fact Tense-Aspect-Mood semantically elaborate categories – would not get us far and would, in fact, result in unnecessary vagueness and imprecision. We will propose – instead – an account of these categories in terms of precise Aristotelian categorisation, whereby two (or more) distinct categories may converge on a number of properties and yet have strict boundaries. This proposal thus fleshes out – in a new area, namely the morphosemantic domain of verbal Tense-Aspect-Mood – the notion of *intersective gradience*, which Aarts 2004, Aarts 2007 and Aarts et al. 2004 introduced for the first time with respect to word classes as well as phrasal and clausal syntactic structures.

References:

- Aarts, Bas. 2004. Modelling linguistic gradience. *Studies in Language* 28 (1): 1-49.
- Aarts, Bas 2007. *Syntactic gradience: The nature of grammatical indeterminacy*. Oxford: Oxford University Press.
- Aarts, Bas, David Denison, Evelin Keizer and Gergana Popova (eds.) 2004. *Fuzzy grammar: A reader*. Oxford: Oxford Linguistics Press.
- Abbi, Anvita (1975). *Reduplication in Hindi. A generative semantic study*. Cornell University.
- Abbi, Anvita 1977. 'Reduplicated Adverbs in Hindi' in *Indian Linguistics* Vol 38:3, Pp. 125-135.
- Abbi, Anvita 1980. *Semantic grammar of Hindi: A study of reduplication*. New Delhi. Bahri Publications Private Ltd.
- Comrie, Bernard 1976. *Aspect*. Cambridge: Cambridge University Press.
- Comrie, Bernard 1985. *Tense*. Cambridge: Cambridge University Press.
- Heine, Bernd 1992. Grammaticalization chains. *Studies in Language* 16: 2, 335- 368.
- Heine, Bernd & Tania Kuteva, 2002. On the evolution of grammatical forms. In: Wray, Alison (ed) 2002. *The transition to language*. Oxford: Oxford University Press, Pp. 376-398.
- Heine, Bernd and Tania Kuteva 2007. *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
- Kuteva, Tania 1998. On identifying an evasive gram: Action narrowly averted. *Studies in Language*. 22: 1, Pp. 113 - 160.
- Kuteva, Tania. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.
- Kuteva, Tania. 2009. Grammatical categories and linguistic theory: elaborateness in grammar. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS.
- Kuteva, Tania, 2010. Grammaticalization and the evolution of grammar: on one particular kind of grammatical categories. Kang, Young-Se, Jong-Yuri Yoon, Jongseon Hong, Jiun-Shiung Wu, Seongha Rhee, Kyoung-Ae Kim, Dong-Ho Choi, Kee-Ho Kim and Hye-Kyung Kang (eds.) 2010. *Lectures on universal grammar and individual languages (SICOL-2010)*. Hankookmunhwasa: Seoul, Korea, Pp. 144-156.
- Plank, Frans 2013. What exactly is...? A new feature: Call for contributions. *Linguistic Typology* 17 (2013), Pp. 267-268.
- Vuillermet, M. 2013a). *A Grammar of Ese Ejja, a Takanan language of the Bolivian Amazon*. Ph.D. dissertation, Université Lumière Lyon 2, 756 pp.
- Vuillermet M. 2013b). "The apprehensional domain in Ese Ejja: Watch out there may be two distinct phenomena!" *Conference on Indigenous Languages of Latin America-VI (CILLA)*, Austin, TX, October 27-30.

On the place of 'stress' in prosodic typology

Bob Ladd, University of Edinburgh

Despite years of theorising and descriptive investigation, there is still a lot of confusion about the typology of stress, accent and tone. In this paper I suggest that the problem is primarily the result of the typological oddness of European stress systems, which has coloured Eurocentric linguists' attempts to understand the prosodic systems of other languages.

The key to the problem is that stress in most European languages is, as it were, fundamentally a 'metrical' phenomenon rather than an 'autosegmental' one. That is, its essence lies not in the assignment of a feature like [stress] to specific syllables, but in the prosodic organisation of whole words and phrases. European stress is relational and strongly 'culminative': it deals in *relative* prominence, and *requires* certain domains to have a single most prominent peak. Naturally, though, stress has phonetic manifestations. In most European languages, for example, stress orchestrates the association of the intonation contour with the segmental string, so that stressed syllables are often accompanied by intonational pitch accents. In addition, even in the absence of intonational pitch accent, individual stressed syllables may be characterised by a variety of subtle phonetic properties like greater duration, greater force of articulation, more peripheral vowels, and so on. These phonetic consequences incline Eurocentric linguists to treat stress as a pretheoretical phonetic notion, and to believe that when they perceive 'stress' they are detecting an objective phonetic property of syllables, just like perceiving whether a syllable is high- or low-toned.

But the relational and culminative nature of stress sets it apart from tone and most pitch accent systems. In particular, many non-European languages seem to get by without any notion of obligatory relative prominence. Even in languages with pitch accent, which is superficially like stress in that it picks out one syllable as more prominent than others, it is generally possible for words not to have any accented syllable at all. By contrast, a key property of European stress systems is that every word or phrase *must* have a single peak of prominence. (There are conservative Indo-European systems with underlyingly accentless stems, but surface strings always have a prominent syllable somewhere.) This is the sense in which most pitch accent is 'autosegmental' and European-style stress is 'metrical'.

Nevertheless, because stress is generally treated as being a syllable-based phonetic phenomenon like accent and tone, Eurocentric linguists are comfortable basing their descriptive statements on their own perception, unaware of the extent to which these perceptions are influenced by structural facts about the phonology of European languages. For example, because of the role of stress in governing the association of intonation contours with the segmental string, they sometimes detect 'stress' in tonal configurations that resemble European intonational pitch accents. I briefly sketch three case studies of descriptive and experimental work where confusion about stress or prominence has led to deeply misleading statements about specific languages and has fed the confusion further. These are: Kenneth Pike on stress in Mixtec; Gosy and Terken on question intonation in Hungarian; and Odé, van Heuven and colleagues on prominence in Indonesian. The goal is not to criticise these specific authors but to identify the unexamined assumptions that have driven their work, and to try to point the way to a more valid typology.

Toward a complex systems approach to the analysis of rhythmic types

Leonardo Lancia, Max Plank Institute for Evolutionary Anthropology, Leipzig, Ger
leonardo_lancia@eva.mpg.de

Georgy Krasovitskiy, University of Oxford, Oxford, UK
georgy.krasovitskiy@ling-phil.ox.ac.uk

Franziska Stuntebeck, Osnabrück University, Osnabrück, Ger
franziska.stuntebeck@gmail.com

Sven Grawunder, Max Plank Institute for Evolutionary Anthropology, Leipzig, Ger
grawunder@eva.mpg.de

We approach speech rhythm as a complex coordination phenomenon and we develop a set of conceptual and computational tools to characterize rhythmic distinctions between languages. The distinction between stress-based and syllable-based languages is a longstanding matter of debate in phonology. Although there is robust evidence that listeners are sensitive to differences between languages of different rhythmic types, no phonetic correlate of this classification has been found so far (Arvaniti, A. 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*). We argue that this lack of concrete evidence for rhythmic distinctions stems from the strongly emergent nature of rhythmic patterns in speech signals. Therefore, rather than looking for phonetic correlates of the different rhythmic patterns, we analyze the dependencies between the processes responsible for the production of stress and those involved in the production of syllables during speech utterances. To this end, we propose an original version of joint recurrence analysis (Romano, M. C., Thiel, M., Kurths, J., & von Bloh, W. 2004. Multivariate recurrence plots. *Physics Letters A*) which we employ to analyze the mutual dependency between non-stationary processes. Likewise, we adapt the cross-convergent mapping method (Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S., 2012. Detecting causality in complex ecosystems. *Science Magazine*) to detect directional influences between the processes under investigation. Utilizing this approach enables us to capture the strength and the direction of the dependency between stress-related and syllable-related processes.

We collected Pear Story narratives from 5 European languages: German and English (stress-based), French and Italian (syllable-based) and Polish as an example of language whose classification is reportedly more ambiguous. So far we analyzed data from five French and six German speakers. In order to extract time-varying quantities which depend on the production of syllables and stress, the evolution over time of the acoustic energy was measured in two different frequency bands: energy below 2.5 Hz was considered to be dependent on syllabic processes, and energy between 2.5 and 12 Hz on processes underlying the production of stress. Narratives were segmented into chunks separated by pauses or hesitations, which were excluded from the analysis.

Results collected so far reveal a stronger link between syllable related and stress related processes in German than in French, as it would be expected on the basis of the relative rhythmic types. Results from the analysis of the other recorded languages will be presented at the conference.

What does the Relabeling-based theory of creole genesis account for? Claire Lefebvre (UQAM and MPI EVA)

The aim of this paper is to review the main features of the Relabeling-based theory of creole genesis and to illustrate the wide range of facts that this approach can account for.

The very nature of the process of relabeling accounts for the fact that creoles reflect the properties of both their source languages in the way they do, and for the fact that creoles emerge in a relative short span of time.

The way relabeling proceeds accounts for a number of facts. Since relabeling is semantically driven, it accounts for the fact that categories that are deprived of semantic content (e.g. operators, structural--as opposed to semantic--case) do not acquire a label at relabeling. Since the relabeling of functional categories proceeds on the basis of free forms, the theory accounts for the fact that creoles tend to be isolating. Since relabeling proceeds on the basis of superstrate languages, it is constrained by what these languages have to offer as an appropriate form to relabel a substrate lexical entry. This accounts for the fact that creoles that share the same substrate languages but that have different lexifiers may manifest slight differences. From the way relabeling applies in creole genesis, it is also possible to derive how word order is established in creoles. Relabeling may be linguistically context-bound or context free. In the former case, relabeling triggers the superstrate word order. In the latter case, the relabeled lexical item may associate with substrate structures.

This theory is a further development of the second language acquisition approach to creole genesis. It is congruent with the Full Transfer/Full Access model of second language acquisition where full transfer can be restated in terms of relabeling (Sprouse 2006).

The relabeling-based theory of creole genesis claims that, by its very nature, relabeling cannot be the only process in creole genesis. Other processes, such as grammaticalization and leveling, also play a role in the development of creoles. These processes interact with relabeling. They are hypothesized to apply at the output of relabeling, that is on the relabeled lexicons. The relationship between relabeling and grammaticalization, as well as the way grammaticalization is hypothesized to apply in creole genesis, account for the fact that the changes observed in the development of a creole are making it more like its substrate languages (e.g. Sankoff 1991). The relationship between relabeling and leveling, accounts for the fact that some creole lexical entries may not have exactly the same properties as those of each corresponding lexical entries in the substrate languages.

Four variables may interfere with the output predicted by this theory. First, the relative homogeneity of the substrate languages plays a role in defining the importance of leveling in a given situation. The more similar the substrate languages, the less important the role of leveling. Conversely, the more disparate the substrate languages, the more important the role of leveling. Second, the amount of exposure to the superstrate language when the creole is being formed plays a role in defining a creole as more or less basilectal/acrolectal. Whether speakers of a developing creole still have access to the lexifier may play a role in the degree of its influence on the developing creole. Finally, whether speakers of a developing creole still have access to the substrate languages may play a role in the degree of their influence on the developing creole.

Agreement with sentential arguments in Adyghe

The syntactic behavior of sentential arguments, showing themselves in constructions like *I know [that he is here]* differ from that of canonical (nominal) arguments in several respects. Here belong, for instance, their incompatibility (or non-standard behavior) with particular constructions and operations (see Postal 2011, Fabrizio 2011, Letuchij 2012 for details).

Among the features which sentential arguments manifest in many languages is their inability to control number agreement. In SAE languages the matrix predicate in general cannot agree in number with its sentential argument, even if the latter includes a coordinate construction where several event are denoted. For instance, in Russian examples like (1) the option with plural agreement is ungrammatical, though it is available in structure (2) with two coordinate NPs in the argument position:

- (1) *Menja raduj-et/*-ut čto ja dopisa-l-Ø stat'j-u*
 I.ACC rejoice-PRS.3SG/*PRS.3PL that I.NOM write.over-PST-SG.M article-SG.ACC
i čto nača-l-o-s' ljet-o.
 and that begin-PST-SG.N summer-SG.NOM.
 'I am glad that I finished the article and that summer began.'
- (2) *Menja raduj-et/-ut pobed-a Spartak-a i vesn-a.*
 I.ACC rejoice-PRS.3SG/PRS.3PL win-SG.NOM Spartak-SG.NOM and spring-SG.NOM
 'I am glad that "Spartak" (a football club) won and that it is spring now.'

However, this property of sentential arguments is not a universal. In Adyghe, a polysynthetic language of West Caucasian group, a coordinate construction including two sentential arguments¹. The use of the 3PL variant with agreement can be favored/disfavored by several factors.

1. The verb form in the sentential argument

As pointed out by Serdobol'skaya (2009), if a particular verb (e.g., *š'ešənen* 'be afraid') can be used either with a nominal argument ('I am afraid of the bear') or a sentential one ('I am afraid that I will lose'), a sentential argument often bears a more peripheral marking (e.g., an instrumental case)². However, these 'peripheral forms' differ in their preferences to agreed/non-agreed variants in some speakers' judgments. There are contexts where for the form on *-n-ew* (masdar + converb/adverbial case), both variants are available, while the form on *-n-žə* (masdar + instrumental case) the plural option is preferred.

- (3) *se mašəne-r s-fə-n-ew-jə parašwətə-m*
 I car-ABS 1SG-lead-MSD-ADV-LNK parachute-OBL
sə-q^h-je-phč'exə-n-ew-jə s-a-š'e-š'həne sə-š'e-š'həne /
 1SG.ABS-DIR-DAT-jump-MSD-ADV-LNK 1SG.ABS-**3PL.IO**-LOC-fear/1SG.ABS-LOC-fear
 'I am afraid of driving a car and jumping with parachute.'
- (4) *se mašəne-r s-fə-n-ž'-jə parašwətə-m*
 I car-ABS 1SG-lead-MSD-INS-LNK parachute-OBL
*sə-q^h-je-phč'exə-n-ž'-jə s-a-š'e-š'həne/*sə-š'e-š'həne*
 1SG.ABS-DIR-DAT-jump-MSD-INS-LNK 1SG.ABS-**3PL.IO**-LOC-fear/*1SG.ABS-LOC-fear
 'I am afraid of driving a car and jumping with parachute.'

2. The syntactic position of the sentential argument

For some speakers, the variant with agreement is impossible if the sentential argument occupies the position of A of a transitive verb (e.g., a causer of a causative verb), but possible if the SA occupies an indirect object position. This seems to reflect a hierarchy of syntactic roles (see Keenan/Comrie 1977): higher semantic roles seem to prefer agreement with canonical (=

¹ See Ershova 2012 with a thorough analysis of sentential arguments in Besleney Kabardian, though no special attention is paid in this work to the agreement of the matrix predicate. The current work is based on the 2014 field trip to Bzhedug Adyghe (one of Adyghe dialects), organized by the Higher School of Economics and Russian State University for Humanities.

² Adyghe has only two cases coding nuclear (nominal) arguments of the verb: absolutive and oblique (ergative). Instrumental-marked NPs cannot control the verbal agreement, contrary to absolutive- and oblique-marked ones.

nominal) arguments, while for lower semantic roles agreement with non-canonical arguments is possible.

Note that the verb form in the sentential argument is the same in (5) and (6).

(5) *mame a- æ -š^həne maš^hjəne s-fə-me
 mummy 3PL.A-CAUS-fear car 1SG.A-lead-COND
 parašwət-ž^h-jə sə-q^h-je-p^hč^h’exə-me
 parachute-INS-LNK 1SG.ABS-DIR-DAT-jump-COND

Intended: ‘It frightens my mother if I drive a car and if I jump with a parachute.’

(5) mame a-š^h’e-š^həne maš^hjəne s-fə-me
 mummy 3PL.IO-LOC-fear car 1SG.A-lead-COND
 parašwət-ž^h-jə sə-q^h-je-p^hč^h’exə-me
 parachute-INS-LNK 1SG.ABS-DIR-DAT-jump-COND

‘My mother is afraid if I drive a car and if I jump with a parachute.’

In the talk, some other factors will be considered.

Therefore, in Adyghe, sentential arguments differ from nominal ones by their restricted ability to control agreement (sometimes agreement with SAs is impossible, while agreement with nuclear NP arguments is always possible). However, this border between SAs and NPs is not that sharp as in SAE languages: the possibility of variants with agreement is conditioned by many factors, but not prohibited entirely.

References

Ershova, Ksenia A. 2012. *Konstrukcii s sentencial’nymi aktantami v besleneevskom dialekte kabardino-čerkesskogo jazyka* [Constructions with sentential arguments in Beslenev Kabardian]. Moscow: Russian State University for Humanities BA diploma work.

Fabrizio, Claudia. 2011. “The syntax of Latin Nominal Infinitives: a manifestation of semantic alignment?” *SILTA (Studi Italiani di linguistica teoria e applicata)*. 2011.

Keenan, Edward L., and Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63-99.

Letuchij, Alexander B. 2012. O nekotoryx svojstvax sentencial’nyx aktantov v russkom jazyke [On some properties of sentential arguments in Russian]. *Voprosy jazykoznanija* 5. 57-87.

Postal, Paul M. 2011. *Edge-Based Clausal Syntax*. Cambridge (Mass.): MIT Press.

Serdobol’skaya, Natalia V., and Motlokhov, Alexander V. 2009 Semantika konstrukcij s sentencial’nymi aktantami v adyghejskom jazyke [Semantics of constructions with sentential arguments in Adyghe]. In Yakov G. Testelets (ed.). *Aspekty polisintetizma: očerki po grammatike adyghejskogo jazyka*. Moscow: RGGU. 498-559.

Language usage, language processing and typology

Steve Levinson, MPI Nijmegen

Recent work in the L&C department in MPI Nijmegen has explored the processing implications of the core ecological niche for language learning and use, namely interactive conversation. It turns out that the rapidity of turn-exchange puts extreme requirements on predictive comprehension and speedy production, reflected e.g. in the trouble kids have to approach adult norms. This strong functional pressure must have implications for language typology. But what exactly? This paper explores what we have recently found out about differing processing in different word orders, and the ways in which the tough processing requirements of conversation can be buffered.

Verbalization and insubordination in Siberian languages

Andrej Malchukov (Mainz)

In my talk I will discuss ‘insubordination’ in Siberian languages, involving reanalysis of nonfinite verbs as finite predicates. The examples include noun-verb continuum in Tungusic, convergence between finite verbs and possessive constructions in Eskimo languages, perfect and habitual tenses in Chukchi, out-of-focus verb forms in Yukaghir, and the “finite” form in Nivkh. I conclude that insubordination manifests a general areal ‘drift’ in North-Eurasian languages, although its results may be partially different reflecting differences in the “input” structures.

References

Evans, Nicholas. 2008. Insubordination and its uses. In: I. Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*, 366—431. Oxford: Oxford University Press.

Malchukov, Andrej. 2004. *Nominalization/verbalization: constraining the typology of transcategorial operations*. Munich: Lincom.

Expectations and Common Ground as grammatical categories:

The Tundra Yukaghir case

Recent research in information structure (IS) has cast doubt on the universality of information structure categories (Matić & Wedgwood 2013). In this paper, I am going to develop this idea further by analysing possible partitions of IS space and exemplify the proposed theory with an analysis of the highly idiosyncratic IS system in Tundra Yukaghir, a north-eastern Siberian isolate.

The set of standard assumptions in the IS research can be subsumed as follows: (a) Common Ground delivers partially defined propositions that constitute the background; (b) the job of grammars is to signal the locus of saturation of these partially defined propositions, i.e. the ‘focus’; (c) grammars must differentiate between the following types of saturation – narrow, broad and sentence focus (Lambrecht 1994 and many others). There is no doubt that this kind of system works well in the description of some European languages. However, many languages seem not to be sensitive to focus size and the degree of specification of the background. In order to try and accommodate at least some (though certainly not all) variation in the IS categories across languages, I will propose an additional dimension of IS, based on possible expectations generated in discourse. If the development of discourse is conceived as a sequence of questions under discussion (QUD, Roberts 2012), then at least the following types of expectations related to QUDs can be hypothesised: (a) eventualities entailed by QUD and therefore expected (EE); (b) eventualities plausible given the current QUD and therefore possible (PP), and (c) eventualities underspecified as to the current QUD (U). These three categories form a Horn scale <EE, PP, U>, such that the use of the type on the right implicates that the meaning of the type(s) on the left does not apply.

I shall argue that the so-called Tundra Yukaghir focus system encodes the difference between EE, PP and U rather than focus size and background specification. In particular, the S- and O-focus sentences denote EE eventualities, neutral sentences with the particle *mər*= PP eventualities, while those sentences that contain neither ‘focus’ marking nor particles convey U-type events. These sentence types directly encode EE, PP and U as *procedural meanings* in the sense of Relevance Theory: they provide the hearer with instructions, not about the content of the partition of the context set that is to be updated, but rather in which way inferences are to be drawn given the current QUD. Using the EE form equals the instruction “interpret as an answer to the preceding QUD”, whereas U form implicates dissociation from the QUD. PP confirms expectations or presents an eventuality as corresponding to expectations given the current QUD. I will show that these general types of instructions can have multifarious interpretive effects.

Tundra Yukaghir thus instantiates an IS system which has little in common with the standard European type, despite appearances. It is conceivable that other languages are sensitive to yet other restrictions and divisions (as seems to be the case e.g. in Toura, as described by Bearth 1992). I conclude with the suggestion that a realistic typology of IS should devote its energy to identifying differences of this kind rather than looking for structural identity where there is none.

References

- Bearth, T. 1992. Constituent structure, natural focus hierarchy and focus types in Toura. *FL* 26: 75-94.
Lambrecht, K. 1994. *Information Structure and Sentence Form*. Oxford: OUP.
Matić, D. & D. Wedgwood. 2013. The meanings of focus. *JL* 49: 127-163.
Roberts, C. 2012. Information structure: Towards an integrated formal theory of pragmatics. *S&P* 5.

Corpus-based Typology : perspectives for cross-linguistic comparison

The goal of this presentation is to discuss the possibility of conducting typological research based on spoken corpora in various under-described languages, in complementation with current methods which rely mostly on grammars.

The wealth of data collected and annotated within endangered-languages projects for instance, makes corpus-based typology a rich alternative or complement to grammar-based typology. However, corpus design and reflections on the nature of annotations are crucial, and so is the elaboration of a methodology for cross-linguistic research based on first-hand data.

Through the analysis of the hypotheses underlying the architecture of two collaborative pilot projects, CorpAfroAs¹ and CorTypo², and through the assessment of the achievements and limitations of those projects and others³, I will establish some of the conditions for optimal use of spoken corpora in lesser-described languages, in view of typological research. This will lead me to address the following issues:

- the tension between language-internal annotation of corpora, and comparability of categories for typology,
- the amount and type of information necessary for the interrogation of corpora by general linguists or typologists unfamiliar with the language.

CorpAfroAs and CorTypo provide different methods for cross-linguistic interrogation of the corpus: CorpAfroAs relies on Grammatical Sketches including language-internal definitions of the glosses used for annotation, for each of the languages of the multilingual corpus. It is the task of the typologist to decide what cross-linguistic concepts or categories can be related to the language-internal ones, and to put together the information allowing to search the corpus for relevant data.

CorTypo includes a comparative database interfaced with the corpus, which is annotated on a language-internal basis. In this database, language-internal categories (e.g. "perfective", or "indefinite") are structured (via exclusive/complementary distribution method) into language-internal functional domains (e.g. "aspect", or "reference"), that can in turn be compared across languages: each FD entry contains all the constructions and morphemes grammaticalized within this domain in the language, and their functions. Since not all languages have grammaticalized all functional domains, comparison relies on the number and contents/structure of functional domains grammaticalized in each language (Frajzyngier 2013).

Whatever the nature of the hypotheses underlying cross-linguistic comparison, it is important that corpora should be annotated on the basis of language-internal categories, and that the comparative component should be a separate module. For instance, one could create a database based on comparative concepts (Haspelmath 2010), or semantic maps (Haspelmath 2003), and it should be possible to interface those databases with multilingual corpora, including CorpAfroAs and CorTypo. The methodological conditions for this, in terms of corpus design, will also be discussed in the presentation.

Ultimately, this presentation aims at fostering collaborative work and future projects, by discussing the theoretical issues at stake for corpus-based typology, in order to propose architectures that facilitate the use of annotated spoken corpora in view of typological analyses.

References

- Haspelmath, M. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison." In: Tomasello, Michael (ed.) *The new psychology of language*, vol. 2. Mahwah, NJ: Lawrence Erlbaum, 211-242.
- Haspelmath, M. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86, 3, 663-687.
- Frajzyngier, Z. (2013) Non-aprioristic typology as a discovery tool. In Thornes, Tim, Erik Andvik, Gwendolyn Hyslop and Joana Jansen (eds.), *Functional-Historical Approaches to Explanation: In honor of Scott DeLancey*. John Benjamins: Amsterdam-Philadelphia.

¹ 2007-2011; <http://corpafroas.huma-num.fr>

² 2013-2016; <http://cortypo.huma-num.fr/>

³ *Relative Frequency of Verbs, Nouns and Pronouns*, <http://www.eva.mpg.de/linguistics/research/typological-surveys/the-relative-frequencies-of-nouns-pronouns-and-verbs-cross-linguistically.html>; *Referentiality*, <http://www2.uni-erfurt.de/sprachwissenschaft/referentiality/index.html>.

Inflectional complexity in creole languages: Evidence from the *Atlas of Pidgin and Creole Language Structures*

Inflectional morphology has been a key ingredient in assessing complexity in creole languages. It has been claimed that synchronically creoles strikingly show a lack of inflectional morphology, thus giving rise to morphologically extremely simplified languages. (McWhorter 2001, Good 2012, Duval-Markussen 2014, Siegel et al. 2014). In this paper, I will challenge this view by drawing on data of the *Atlas of Pidgin and Creole Language Structures* (Michaelis et al. 2013). I will consider the following morpheme types: tense-aspect markers, case markers, and definite articles:

- (1) Ternate Chabacano (Spanish-based, Philippines; Sippola 2013)

Ta yudá éle su marido
IPFV help 3SG 3.SG.POSS husband
'She helps her husband.'

- (2) Korlai (Portuguese-based, India; Clements 2013)

Pedru su kadz tidoy kadz Pedru su
Pedru GEN house both house Pedru GEN
'Pedru's house' 'both of Pedru's houses'

- (3) Haitian Creole (French-based, Caribbean; Fattier 2013)

kay la avyon an ti fi a
house DET plane DET little girl DET
'the house' 'the plane' 'the little girl'

The main misconception which has given rise to the idea that creoles lack inflectional morphology is the fact that scholars have implicitly or explicitly based themselves exclusively on spelling conventions in deciding whether a morpheme is an affix or a free morpheme (see e.g. Siegel et al. 2014). But all morphemes written as separate words in the cited examples above can be interpreted as affixes. I will consider two criteria for affixhood: (i) uninterruptibility and (ii) morphonological idiosyncrasies. In example (1), nothing can intervene between the imperfective marker and the verb stem. The same holds for the genitive case marker and the possessor in example (2). In example (3), the definite article shows morphonological variation depending on the phonological shape of its host. Instances of tense-aspect affixes are abundant in the creoles in *APiCS*, and affixed case markers and definite articles also occur repeatedly. Thus, I argue that creoles have lost most of their lexifiers' inflectional marking, but have gained at the same time a considerable degree of complex morphology through unusual accelerated grammaticalization processes.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45,1: 31-80.

Daval-Markussen, Aymeric. 2014. First steps towards a typological profile of creoles. *Acta Linguistica Hafniensia* 45,2: 274-295. DOI:10.1080/03740463.2014.880606

- Fattier, Dominique. 2013. Haitian Creole structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/49>.
- Good, Jeff. 2012. Typologizing grammatical complexities or why creoles may be paradigmatically simple but syntagmatically average. *Journal of Pidgin and Creole Languages* 27:1-47.
- Clements, J. Clancy. 2013. Korlai structure dataset. In: Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) s. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/40>
- McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5(2-3). 125–166.
- Michaelis, Susanne, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). 2013. *The Atlas of Pidgin and Creole Language Structures*. Oxford: Oxford University Press.
- Siegel, Jeff & Szmrecsanyi, Benedikt & Kortmann, Bernd. 2014. Measuring analyticity and syntheticity in creoles. *Journal of Pidgin and Creole Languages* 29,1: 49–85.
- Sippola, Eeva. 2013. Ternate Chabacano structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/44>.

Natural speech vs. controlled-stimulus narratives: referential density in Chechen

Zarina Molochieva

Bickel (2003, 2006) shows that referential density, variation in the ratio of overt arguments to grammatically available arguments, varies significantly across languages. His work raises two questions which the present paper aims at answering: (1) What causes cross-linguistic differences in referential density -- structural properties of grammar, situational and pragmatic factors such as text length, discourse traditions, social situation (Bickel 2006)? (2) Can referential density be calculated on natural speech or does it require a controlled stimulus? Bickel uses exclusively narratives produced in response to a stimulus, the Pear Story movie (Chafe 1975).

This paper uses narratives in Chechen (Nakh-Daghestanian, Caucasus) as a test. Chechen is a good test language because it has a rich evidential system and differences are reported in the frequency of evidential usage between highland dialects (rural, more isolated, smaller communities) and lowland dialects (larger speech communities, some urban, well connected). The frequency of evidential categories in highland dialects depends on the interpersonal relationship of the speakers, with closely acquainted speakers able to presuppose and assume more about the hearer's knowledge and therefore using relevant evidentials.

This paper presents a pilot study comparing referential density in Bickel's ten Pear Film narratives and two spontaneously produced texts of about 200 clauses each. The speakers of the spontaneously produced texts are rural elderly speakers (75 and 82) of different dialects, one highland and one lowland. The lowland speaker addresses and interacts with a non-close acquaintance who speaks a different dialect, while the highland speaker speaks to a very close acquaintance from the same highland dialect. Both texts are third-person narratives about extended events with motions, actions, and plot, similar to Pear Film narratives. The Pear Film texts were collected from younger diasporic Standard (lowland) Chechen speakers and mostly speaking to people they do not know very well. Since referential density as reported by Bickel (especially 2006) and the conditions for evidential usage in Chechen both involve speakers' and hearers' attention to and knowledge of referents, the prediction is that the highland texts will have lower referential density than the lowland texts.

The pilot study confirms this prediction. Referential densities are: Pear stories ~0.54 (Bickel 2006), spontaneous lowland speech 0.55, spontaneous highland speech 0.50. Highland speech has lower referential density. Lowland spontaneous speech and Pear Film narratives have similar referential density. Referential density is not greatly discrepant overall; any of the three values would keep Chechen in its same place in Bickel's table (2006). This indicates that referential density can also be usefully measured on natural texts (assuming basic genre consistency as here).

Bickel (2003:733) suggests that, in languages with low referential density the speakers may pay more attention to the event, while in languages with high referential density speakers attend more to the referents. This study suggests that the matter is more complex: In Chechen, in general and in my texts, it is motion events which structure narrative, and motion events are greatly elaborated using serialization of motion verbs, spatial preverbs, etc., suggesting that speakers are paying a great deal of attention to the motion events. In particular this is true in the highland text, which has the lower referential density. The pilot study will be expanded to approximately ten texts by the time of the conference.

References:

- Bickel, Balthasar. 2003. Referential density in discourse and syntactic typology. *Language* 79, 708-36.
Bickel, Balthasar. 2006. Referential density in typological perspective. Plenary lecture at the Leipzig Spring School on Linguistic Diversity. March 22, 2006.
Chafe, Wallace L. 1975. The Pear Film. <http://www.linguistics.ucsb.edu/faculty/chafe/pearfilm.htm>
[Two additional references on factual points suppressed for anonymity.]

Phonological features from a quantitative typological perspective

Large-scale typological surveys in phonology have mostly concentrated on surface segment inventories of spoken languages. While this tradition has generated useful insights into many patterns, especially areal patterns, it risks missing universal principles in the structure and evolution of inventories. Here, we present a series of analyses that build on a class of algorithms that we use to extract and quantify feature decompositions. The algorithms derive minimal sets of features of mathematically determined contrasts. These methods are independent of linguistic theory; in fact, they provide a straightforward mechanism to compare different proposals for distinctive feature systems.

We then subject the results from the feature decomposition algorithms to quantifications of the empirical performance of specific theoretical proposals (feature geometries) as well as to estimations of worldwide, areal and genealogical trends. All results are based on the PHOIBLE database, which contains over 2100 inventories from 1672 distinct languages (Moran et al. 2014).

Our findings suggest that

- while most inventories allow multiple (competing) feature analyses, 90% of the inventories allow less than 50 decompositions, making results manageable
- while a restricted set of features (based on a geometry with 37 features) allows coverage of all known segment inventories in the database, the phonetic and phonemic implementation in segments appear to constitute an extremely large inventory of which we have do not know the limits
- phonetically informed and universally constrained feature geometries allow more efficient segment coding in languages than arbitrary and language-specific feature sets
- there is evidence for a universally preferred combination of segment inventory size and feature numbers, centered on about 36 segments and 12 features
- the number of features in phonological inventories tend to be stable genealogically
- certain features are diachronically preferred (e.g. high, front, labial) and dispreferred (e.g. labiodental, round, fortis)

Overall, these results show that algorithmically derived feature decompositions provide a fruitful but little-exploited terrain for phonological typology. Our analyses reveal strong universal constraints on the organization and evolution of sound systems. All these constraints are probabilistic, not categorical, in line with much recent work on the nature and emergence of phonology (Blevins 2004, Mielke 2004, Sandler et al 2011, Collier et al. 2014).

References

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Collier, Katie, Balthasar Bickel, Carel P van Schaik, Marta B Manser & Simon W Townsend. 2014. Language evolution: syntax before phonology? *Proceedings of the Royal Society B: Biological Sciences* 281(1788). 20140263.

Mielke, Jeff. 2004. *The Emergence of Distinctive Features*: The Ohio State University dissertation.

Moran, Steven, Daniel McCloy & Richard Wright (eds.). 2014. PHOIBLE Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://phoible.org/>.

Sandler, Wendy, Mark Aronoff, Irit Meir & Carol Padden. 2011. The gradual emergence of phonological form in a new language. *Natural language & linguistic theory* 29(2). 503–543.

Laws of in-law languages

Maarten Mous, Leiden University Centre of Linguistics

The paper provides a typology of in-law languages in the world. In South and East Africa, in the Caucasus, in Mongolia, in Australia, and in South East Asia registers of respect developed around the avoidance of the name of certain in-laws and related vocabulary. There are commonalities in which words to avoid (when is word considered similar - sound, morphology) and which strategies to avoid them (form and semantic manipulations, lexicon) and their properties (form manipulations mainly of initial of stem or word; preference for certain sounds in form manipulations; semantic grouping in lexical replacement; lexical structuring in applying a same strategy). These commonalities provide insights in speakers' "view" on language structure and may help recognising the effects of such phenomena in earlier times and reconstructing language history. The paper also addresses the social characteristics (who respects whom, when, at what sanction and how is the register learned).

Expressing (non)existence in some Oceanic languages
Claire Moyse-Faurie (Lacito-CNRS)

On the sad occasion of the closing down of the Linguistics Department of the MPI, I will talk about be's and not be's¹. Many (not) be(e)'s are indeed flying around in Oceanic languages. Some only refer to pure existence ('be', 'exist'), others aim at localizing the existence ('be at'), either in a neutral way, or by specifying the posture ('be standing, sitting, lying'; Lichtenberk 2002). Moreover, existential verbs are well known for expressing possession (Lyons 1967).

Each of these 'be's' has its own personality, that is, specific compatibilities with tense-aspect and negative markers, choice of subject (impersonal vs. personal, specific vs. nonspecific, animate vs inanimate). And of course, each of these currently observable 'be's' has its own history, even if some of them do fly across several countries.

My talk will mostly focus on the verbs expressing non-existence or absence, even though the dichotomy between "existence vs non-existence" is not that strong in the Oceanic cultural world, where there is a spatial contiguity between the world of Death (under the sea) and the world of Living. [Analogously the MPI Linguistic Department will not totally disappear; it will just go on living in a different way, or in different places.]

Oceanic languages have more existential verbs than non-existential ones. For example, Wayan Fijian has three existential verbs, and only two negative ones (Pawley 2000), as is also the case in some Kanak languages: one means 'be absent', the other 'not exist, be none, not have'. In other Kanak languages, such as Xârâcùù, there is only one negative verb for these two values. The forms used for the various negative types (denial, negative verb, negative verbal marker, negative imperative), may be identical or similar, but they can also be very different (Moyse-Faurie and Ozanne-Rivierre 1999). Faced with such a variety of negative forms, it is difficult to identify a diachronic cycle such as the one described by Croft (1991) for existential negative verbs. We can only identify the change from a specific existential negative verb to a preverbal negative marker as for instance in Kanak languages, such as Nyelâyu, Tîrî or Xârâcùù.

In some languages, it is possible to negate the verb of existence in order to express non-presence, apart from the existential negative verb. This is the case in East Futunan, in which the negative verb *le'e* conveys a purely abstract non-existence meaning (1), while the combination of the existential verb *iai* with the negative marker *se* insists on the absence of something in a specific place (2):

- (1) *Na le'e ni puaka i Futuna i le temi mu'a.*
PAST not.exist NSPEC.PL pig in Futuna in SPEC time before
"In the past, there were no [existing] pigs in Futuna."
- (2) *Na se iai ni puaka i Futuna.*
PAST NEG exist NSPEC.PL pig in Futuna
"There were no pigs [living] in Futuna".

Both constructions, however, require the use of non-specific articles.

References

- Croft, William. 1991. The evolution of negation, *Journal of Linguistics* 27, 1:1-27.
Lichtenberk, Frantisek, 2002. Posture verbs in Oceanic. In John Newman (ed.), *The Linguistics of Sitting, Standing and Lying*, Typological Studies in Language 51, Amsterdam/Philadelphia, Benjamins, 269-314.

¹ Plagiarizing Andy Pawley's 2000 title of his article on copulas in Wayan Fijian!

- Moyse-Faurie, Claire and Françoise Ozanne-Rivierre, 1999. Negation in New Caledonia and Loyalty Islands languages. In E. Hovdhaugen and U. Mosel (eds.), *Negation in Oceanic Languages*. Typological Studies, Lincom Europa, Lincom Studies in Austronesian Linguistics 2, 57-79.
- Pawley, Andrew, 2000. Two be's or not two be's. On the copulas of Wayan Fijian. In Bill Palmer and Paul Geraghty (eds), *SICOL, Proceedings of the Second International Conference in Oceanic Linguistics vol. 2, Historical and descriptive Studies*. Canberra, Pacific Linguistics, 297-314.

MOOD IN DARGWA LANGUAGES

Dargwa languages are characterized by the rich systems of verb forms expressing modality. I distinguish direct mood (indicative) which expresses tense, and indirect moods. The system of indirect moods comprises two groups of forms:

a) modality which involves locutors (the speaker and the addressee): marked optative, unmarked optative, conditional optative, jussive, probabilitive, imperative, rogative (from Latin *rogare* 'to ask'), and prohibitive.

b) epistemic modality: conditional mood, intentional mood, concessive mood, hypothetical mood, irrealis.

Each of these forms will be addressed in the paper in regard to its meaning, morphological and periphrastic marking, paradigms, negative counterparts, and usage. Different Dargwa languages are compared in respect of the quantity, structure and usage of indirect mood forms.

Tsammalex

Tsammalex is a multilingual lexical database on plants and animals that includes an image and video repository (Naumann et al 2014). It focuses on the Kalahari region of Southern Africa and on the Dogon region in Mali, West Africa. Tsammalex is a platform for scholars of lexicology in vernacular languages and other users (biologists, language planners, communities, and individual users) to share images, to search for biological species, and to compare lexical data. To date, the sample contains about 60 languages and over 15,000 lexical items, which provide names for roughly 1,000 different species described with 4,000 images and video.

Due to our research expertise, Tsammalex is currently focused on Africa. Nevertheless, datasets that comply with the Cross-Linguistic Linked Data (CLLD) framework (Forkel et al, 2014) are scalable to worldwide geographic and linguistic diversity. CLLD datasets are also compatible with other MPI-EVA published typological databases, e.g. WALS, WOLD, PHOIBLE, so that different databases can be queried in tandem from a central CLLD portal webpage.

Furthermore, the functionality that Tsammalex extends from other CLLD applications includes data on ecoregions taken from the World Wildlife Fund's (WWF) *Terrestrial Ecoregions of the World* (Olson et al 2001). Map layers include ecoregion borders that have been simplified using the modified Visvalingam algorithm implemented by the tool mapshaper¹ to allow for better rendering performance. WWF contains very detailed information about ecozones, e.g. the *Northern Congolian forest-savanna mosaic* is 705006 km² in area, is categorized as *Tropical & Subtropical Grasslands, Savannas & Shrublands*, and its status is endangered.²

In this talk we present a brief overview of Tsammalex, including highlighting how image repositories are used in different regions by researchers to document African languages. We illustrate the data collection potential of how Tsammalex can bring together resources collected at MPI-EVA and other projects such as the Dogon Languages Project (Heath et al 2014). We also show how Tsammalex reaches out into other scientific communities.³ And by making available quantitative data on languages and where they are spoken in regard to habitat, Tsammalex provides the infrastructure necessary to undertake studies on the pan-African distribution of words, on dialect variation,⁴ and on linguistic and anthropological studies that investigate the distribution of aspects of spoken languages and environmental factors. If time allows, we will report on such studies.

¹ <http://mapshaper.org/>

² <http://www.worldwildlife.org/ecoregions/at0712>

³ For example, <http://paw.plantsanimalswords.org/>

⁴ For example, by creating Gabmap input from CLLD applications. See: <https://www.gabmap.nl/>

References

Forkel, Robert et al. 2014. CLLD: CLLD toolkit for cross-linguistic databases. ZENODO. 10.5281/zenodo.11854.

Heath, Jeffrey et al. 2014. The Dogon Languages Project. Online: <http://dogonlanguages.org/>.

Naumann, Christfried & Sell, Lena & Jaulgey, Noémie (eds.) 2014. A lexical database on plants and animals (preliminary version). Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://tsammalex.clld.org>, Accessed on 2014-11-28.)

Olson, D.M., E. Dinerstein, E.D. Wikramanayake, N.D. Burgess, G.V.N. Powell, E.C. Underwood, J.A. D'Amico, I. Itoua, H.E. Strand, J.C. Morrison, C.J. Loucks, T.F. Allnutt, T.H. Ricketts, Y. Kura, J.F. Lamoreux, W.W. Wettengel, P. Hedao, and K.R. Kassem. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* 51:933-938.

The Prospective as a cross-linguistic Grammatical Category

As a cross-linguistic phenomenon, the category of prospective, also defined as 'immediate/imminent future', 'near/nearest future', 'close future' or 'proximative' (Heine&Kuteva 2002), has attracted the attention of general linguists not very long ago. In linguistic reference literature, this term is still understood only as the future aspect meaning of present tense forms of Russian perfective verbs (Metzler Lexikon Sprache 2010). Contrary to this, Lars Johanson opposes the prospective to interterminal, adterminal and postterminal aspects and defines the Turkish future tense form *-acak* as a prospective one, thus arguing that Turkic languages do not have future tense forms, but prospective aspect ones (1975). Bernard Comrie defines the prospective aspect through its opposition to Perfect. He also opposes the prospective to the future and to expressions of intention and mentions the combinability of English prospective forms (*be going/about to do*) with tense markers (1976: 64-65). Later, typological investigations carried out by Östen Dahl (1985) and Joan Bybee (1994: 271-273) found evidence for postulating a cross-linguistic category of 'prospective'.

For Turkic, Nevskaya 2005 gave a preliminary description of prospective means and main sources of their development. These are as follows: 1) actional constructions with infinitives of the lexical verbs and various auxiliary verbs bearing tense morphology; tense markers getting omitted by certain forms, which then express tense reference themselves and function as near futures (Intention>Prospective); 2) constructions of direct speech with volitive forms of the lexical verb (Intention>Prospective); 3) synthetic intraterminal aspect forms of low focality (mainly participles) that tend to develop into prospectives and, finally, into futures (Intraterminal>(Obligation)>Prospective).

In September 2013, Irina Nevskaya and Agnes Korn organized a symposium in Frankfurt with the title: *The Prospective as a Grammatical Category: Evidence from Turkic, Iranian and beyond*. Its participants gave first descriptions of this category for a number of Altaic, Uralic and Iranian languages. Many of the languages under this survey are endangered ones and not sufficiently described.

In the presented lecture we will first address the following problematic issues connected with Prospective: the need to define this category in a stricter way and delineate it from various categories belonging to epistemic, deontic or dynamic modalities that refer to anticipated or planned (i.e. potential) actions (future tenses, intention, ability, possibility, obligation, volition, desiderative, suppositional mood, etc.). Secondly, we will sum up the preliminary results of our research on the main formal and semantic types of prospective language means in Altaic languages giving a broader Eurasian perspective.

Literature

- Bybee, J. Perkins, R. & Pagliuca, W. (1994): *The Evolution of Grammar. Tense, aspect and modality in the languages of the world*. Chicago: The University of Chicago Press.
- Comrie, B. (1976): *Aspect*. New York: Cambridge University Press.
- Dahl, Ö. (1985): *Tense and Aspect Systems*. London, New York: Basil Blackwell.
- Heine, B. & Kuteva, T. (2002): *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Johanson, L. (1971): *Aspekt im Türkischen*. Uppsala: Almqvist & Wiksell.
- Metzler Lexikon Sprache. (2010). München: Knaur Verlag.
- Nevskaya, I. (2005): The Typology of prospective in Turkic languages. In: *Sprachtypologie und Universalienforschung*. Focus on: Sentence types and sentence structures. Gippert, J., Erdal, M. & Voßen, R. (Eds.). 58/1, 111-123.

Pan-Eurasian typological clines

Johanna Nichols, UC Berkeley

EVA Linguistics Department closing conference

My work at EVA over the years (since 2002) has involved fieldwork on languages of the Caucasus and cross-linguistic work on several typological variables. This paper brings these two threads together and proposes findings and hypotheses bearing on the long-term prehistory of Eurasia.

As work progresses on complex and multifactorial typological variables – ones subdivided by enough datapoints to give a fine-grained typology and fine-grained geographical distribution – most of them seem to exhibit a clinal distribution across all of Eurasia. These include (* = variable proposed by other EVA researchers, used and/or extended here by me):

- Lexical vs. inflectional person. Where person is more lexical it is bound up with lexical and lexically listed properties of nouns (e.g. sharing gender marking, number categories, declension classes, etc. with nouns); where it is more inflectional it is closer to pure person, factored out from other categories, driving hierarchical marking, favoring inclusive/exclusive distinctions, showing up as outermost of the inflectional affixes, etc. Person is most nearly lexical in western Eurasia, most nearly inflectional in the east and especially in the Pacific Rim linguistic population.

- Verb-based vs. noun-based derivation: the proportion of a standard wordlist for which the derivational base is a verb vs. for which it is a noun. Western languages (especially European languages) have higher proportions of noun-based derivations; eastern ones, of verb-based derivations.

- The proportion of verbs in a standard wordlist that use causativization as realization of the causative alternation (Nichols, Peterson, Barnes 2004). This distribution is distorted in interesting ways by sociolinguistic and processing factors, but granting these factors the large-scale pattern is infrequent causativization in the west, frequent causativization in the east.

Other promising candidates for east-west clinal distributions include:

- Chiefly within southern Eurasia, proportion of light verb constructions in the same standard wordlist for the causative alternation. *(Stilo)

- Proportion of telic vs. ingressive lexicalization of predicates involving a lead-up phase, a state-changing event, and a resultant state. In the west these are almost always lexicalized as telic (usually achievements); in the east as ingressives.

- The ratio of verbal vs. nominal inflectional synthesis. Verbal synthesis is higher in the east, nominal in the west.

- Case-based vs. role-based or hierarchy-based constraints on syntactic rules. *(Bickel 2003, 2004)

Other complex variables are just coming into consideration.

- Various phonological properties. *(Blasi & Moran)

This paper presents a close typological survey of Eurasian languages plus the rest of the Northern Pacific Rim, firming up the distributional picture. I also discuss how the known reconstructable protolanguages fit into the picture and how and why typological enclaves (Bickel & Nichols 2003) sometimes distort it; and how and why, for some variables, distributions differ between northern and southern Eurasia.

Langues, dynamiques et pertinences

Robert Nicolai

À partir de quelques exemples concernant le contact des langues et des populations l'objectif de cette contribution est de s'intéresser au thème évolutif dans le temps de la saisie de la « dynamique des langues et de sa préhension » à travers l'activité et l'activisme des acteurs de la communication et à travers les procès de mise en signification qu'ils développent.

Il s'agit donc de prendre en considération deux procès parallèles : celui que les locuteurs (acteurs séculiers) actualisent dans leur utilisation des langues et celui que les descripteurs des langues (acteurs réguliers) mettent en œuvre dans le compte rendu qu'ils en font, étant bien entendu que ceux-ci sont aussi ceux-là.

La réflexion se focalisera sur la question de la sélection et de la modification des pertinences retenues par ces acteurs, sur celle de l'élaboration théorique et des transformations introduites en rapport.

Cela introduit à une approche des modalités de construction sémiotique qui contribuent à la transformation des langues, et à une approche des modalités d'élaboration des modèles. Ce qui met nécessairement en regard le « hier » et le « aujourd'hui ».

Soit donc : questionner les langues, notre pratique des langues et nos façons de les décrire hier et aujourd'hui. Qui ? Comment ? Pourquoi ? Pour qui ?

Morphophonological Diversity in the Arawak Family: The Diachronic Origins of the Mehinaku Absolute Markers

The aim of this presentation is to discuss a diachronic account of selected aspects of the morpho-phonology of nominal possession in Mehinaku, an Arawak language spoken in central Brazil (see [AUTHOR] forthcoming 2015). The phenomena explained here include the systematic occurrence of word-final main stress in Absolute and Genitive forms of nouns, the nasalization of stem-final vowels in a sub-class of the Genitive forms, as well as some vowel quality changes in the paradigms of Inalienable nouns that have defied coherent treatment in comparative and historical Arawak studies.

I demonstrate how, through the application of regular sound changes (in special, $*f' > \emptyset / _ i$) and the action of phonological constraints active in the phonology of Mehinaku, such as general restrictions on vowel sequences, the superficially diverse set of exponence strategies displayed by the modern language - stress shift, vowel quality changes (ablaut), nasalization and suffixation - can be shown to derive from a simple concatenative system, one reconstructed for the proto-language (cf. Payne 1991) and more representative of other modern Arawak languages. The simpler reconstructed system employs a single suffix for coding the Absolute form of Inalienable nouns and a restricted set of suffix allomorphs for the marking of the possessed status of Alienable nouns. The particular developments postulated in our account are supported both by a consideration of the synchronic regularities holding of Mehinaku phonology and by comparative evidence on the development of the Central Branch of the Arawak family to which this language belongs.

REFERENCES

- [AUTHOR]. forthcoming 2015. On the Realization of Nominal Possession in Mehinaku: A Diachronic Account. *International Journal of American Linguistics* 81 (1): 119-132.
- Payne, David. 1991. A Classification of Maipuran (Arawakan) Languages Based on Shared Lexical Retentions. *Handbook of Amazonian Languages, Vol. 3*, ed. D. Derbyshire and G. K. Pullum, pp. 355-499. Mouton de Gruyter.

Genetic perspectives on prehistoric language contact: a way forward or a dead-end?

Over the previous decades, studies of language contact have emerged as an independent discipline of linguistics, with a journal dedicated to this topic and ever more numerous workshops, conferences, and symposia being convened to discuss various aspects of contact-induced linguistic changes. One might therefore expect that by now all is known about the how and why of language contact, and that only the finer details concerning individual case studies need to be filled in. And yet, basic issues such as which changes are expected to result from language shift and what kind of impact to expect in cases of long-term bilingualism are still debated (Lucas 2014; Ross 2013; Thomason 2003). This is partly due to a lack of historical documentation of the vast majority of language contact situations, with the exception of the relatively few well-documented languages of Europe and East Asia. In the absence of written records, researchers can only make use of the linguistic changes they find to deduce information about the nature of the contact process, but as yet no widely accepted diagnostic tools exist (Ross 2003; Ross 2013: 15-16). One way forward might thus be to obtain independent information about the prehistoric contact situation from other fields such as archaeology (Heggarty 2014) and, as I have previously argued, molecular anthropology (Pakendorf 2011, 2014a). However, given the rather disparate temporal dimensions and potentially divergent modes of transfer of genes and languages, the question remains whether insights gained from molecular anthropological investigations can really advance our knowledge of the processes and results of language contact. In this talk I will critically review the results of a decade of interdisciplinary research into population and language contact in Africa and Siberia conducted largely at the MPI of Evolutionary Anthropology (Pakendorf 2007, 2014b, to appear; Barbieri et al. 2012, 2013; Beyer & Schreiber 2013; Bostoen & Sands 2012; Duggan et al. 2013) to investigate the question whether such approaches indeed present a way forward for contact linguistics or are rather a dead-end.

References

- Barbieri, C. et al. (2012): *Molecular Biology and Evolution* 29(4):1213-1223.
- Barbieri, C. et al. (2013): *European Journal of Human Genetics* 21: 430-436.
- Beyer, K. & Schreiber, H. (2013): In Léglise I. & Chamoreau, C. (eds): *The Interplay of Variation and Change in Contact Settings*, 107–134.
- Bostoen, K. & Sands, B. (2012): In Brenzinger, M. & Fehn, A.-M. (eds): *Proceedings of the 6th World Congress of African Linguistics Cologne 2009*, 129-140.
- Duggan, A. et al. (2013): *PLoS ONE* 8 (12): e83570. DOI: 10.1371/journal.pone.0083570.
- Heggarty, P. (2014): In Bowerman, C. & Evans, B. (eds): *The Routledge Handbook of Historical Linguistics*, 598-626.
- Lucas, C. (2014): In Bowerman, C. & Evans, B. (eds): *The Routledge Handbook of Historical Linguistics*, 519-536.
- Pakendorf, B. (2007): LOT Dissertation series 170. Utrecht: LOT.
- Pakendorf, B. (2011): In: Gibson, K. & Tallermann, M. (eds): *The Oxford Handbook of Language Evolution*, 573-580.
- Pakendorf, B. (2014a): In Bowerman, C. & Evans, B. (eds): *The Routledge Handbook of Historical Linguistics*, 627-641.
- Pakendorf, B. (2014b): In Robbeets, M. & Bisang, W. (eds): *Paradigm Change in the Transeurasian Languages and Beyond*, 287-310.
- Pakendorf, B. (to appear): In Gardani, F. et al. (eds): *Borrowed Morphology*.
- Ross, M. (2003): In Hickey, R. (ed): *Motives for Language Change*, 174-198.
- Ross, M. (2013): *Journal of Language Contact* 6: 5-47.
- Thomason, S.G. (2003): In Joseph, B.D. & Janda R.D. (eds): *The Handbook of Historical Linguistics*, 687-712.

The influence of the Lori (Borujerdi) dialect on the Borujerd Jewish dialect

Shiva Piryaee¹
(piryaeeshiva@gmail.com)

This paper describes research on the dialect of the Borujerd² Jews, which is spoken by Jewish people living in Borujerd city.

Since many Borujerd Jews have emigrated from Iran (especially to Israel), there are just a few speakers left in this city. This critical issue will lead this dialect to become extinct soon in the coming years.

As Yarshater³ has mentioned before in his article, this dialect is of a Median type, very similar to the one spoken in Nehavand.

According to him “as the Iranian Jewish communities have as a rule preserved the original dialects of their localities, one would expect a Lori dialect among the Jews of Borujerd”.

So, this paper is a linguistic study on how the Borujerdi dialect might have affected the Jews’ dialect. Because of the increasing decline in the numbers of the Jewish population in this city, especially in recent years, there are just a few native speakers of this dialect now living there.

In fact, this dialect is in dramatic danger of extinction, unless prompt action is taken to preserve, record and analyze it linguistically.

Key words: Jews’ dialect, Borujerdi dialect, Lori

¹ Iranian dialect researcher in the Center for the Great Islamic Encyclopedia (CGIE); the researcher is a native speaker of the Borujerdi dialect.

² A city located in Lorestan province, Iran.

³ Ehsan Yarshater, “The Dialect of Borujerd Jews”, *Archaeologia Iranica et Orientalis*, Gent, 1989 (1029-1046).

Scope under language contact

Maria Polinsky
Harvard University

English doubly-quantified sentences readily admit scope ambiguities:

(1) A shark attacked every pirate. (*a* > *every*, *every* > *a*)

Contrasting with English, Chinese is assumed to lack inverse scope (cf. the Isomorphic Principle; Huang 1982, Aoun & Li 1989). The sentence in (2) should therefore only describe a situation with a single boy. Note that sentence-initial numeral phrases like *yi-ge-nanhai* require the existential predicate *you* 'be/have', which is obligatory:

(2) 有 一只 鲨鱼 攻击了 每一个 海盗 (simplified Chinese)

有 一隻 鯊魚 攻擊了 每一個 海盜 (traditional Chinese)

You yi-zhi shayu gongji-le mei-yi-ge haidao.

have one-CL shark attack-PFV every-one-CLF pirate

'A shark attacked every pirate.' (*one* > *every*, **every* > *one*)

The status of Chinese inverse scope, both why and whether it is disallowed, has come under recent scrutiny. For example, Zhou & Gao [Z&G] (2009) tested speakers in Beijing on the interpretation of doubly-quantified sentences such as (3):

(3) 每个 强盗 都 抢了 一个 银行 (simplified Chinese)

每個 強盜 都 搶了 一個 銀行 (traditional Chinese)

mei-ge qiangdao dou qiang-le yi-ge yinhang

every-CL robber DOU rob-PFV one-CLF bank

'Every robber robbed one bank.'

Z&G conclude that despite its dispreferred status, an inverse scope interpretation is in fact available in Chinese. Their study and its results face a serious entailment problem, however, because their doubly-quantified sentences featuring *every* over *one/a* at surface structure felicitously describe the supposed inverse interpretation solely on the basis of surface scope: if one bank was robbed by every robber then it is trivially the case that for every robber there is a bank that he or she robbed; *one/a* > *every* entails *every* > *one/a*. Hence Z&G's conclusion that Chinese has inverse scope is not well founded.

The current study tests the status of Chinese inverse scope by focusing on the interpretations available for sentences like (2) where *one/a* scopes over *every* at surface structure, a feature that avoids the entailment problem inherent to Z&G's materials. By comparing the responses from native speakers of Chinese (N=72) and native speakers of English (N=144), we show that Chinese in fact does not allow inverse scope. Further, our results 1) suggest that the Chinese prohibition on inverse scope does not straightforwardly emerge from numeral semantics or bi-clausal structure of sentences with *you*, and 2) demonstrate that in English the numeral *one* competes with the indefinite article *a* and yields a strong specificity inference.

We then compare the scope grammars of native speakers with the scope grammar of heritage speakers of Mandarin whose dominant language is English (N=26) and heritage speakers of English whose dominant language does not have inverse scope (N=22). The goal of this comparison is to determine whether or not scope interpretations are transferred from one language to another under contact. Our results show that heritage speakers of do not alter their scope under the influence of their dominant language, regardless of the direction of possible transfer toward simplification or complication. This finding has implications for the theory of scope representations as well as for our understanding of the transfer between languages in bilingual systems.

Kildin Saami documentary sociolinguistics: Investigations on the variation and change in function words

Michael Rießler (Freiburg), Niko Partanen (Freiburg), Evgenia Zhivotova (Leipzig)

This paper is a result of collaborative work between language documentation, language technology and corpus-based sociolinguistics and includes both applied and theoretical research aspects. We aim at contributing to diversity linguistics by improving current methodology for the building of comprehensive databases for future research on endangered and little-studied speech communities. Applying *language technology* in *language documentation* helps us to create more systematically annotated *corpora*, rather than eclectic data collections. Ultimately, the multimodal corpora we create will be useful for *quantitative* investigations into synchronic variation and change in Kildin Saami.

Kildin Saami (Glottolog: kild1236) is a seriously endangered and under-described language spoken actively by no more than a few hundred speakers on the Kola Peninsula in North-West Russia. Although the number of speakers is decreasing rapidly, recent language planning and revitalization attempts have opened new domains of language use, for instance in print, radio and social media (Rießler 2014). Changes induced by this language shift to Russian have been documented on all levels of linguistic structure in Kildin Saami: phonology, morphology, syntax, discourse-pragmatics and lexicon (Rießler 2007; Rießler 2009; Blokland and Rießler 2011) and variation and change in function words specifically are dealt with in Karvovskaya (2011), Zhivotova (2014a), Zhivotova (2014b), and Kotcheva and Rießler (2015). Rießler and Karvovskaya (2013) have argued (based on the use of focus particles) that not only shift-induced language attrition is responsible for synchronic variation and change in Kildin Saami, but language planning and the introduction of new (mostly written) domains of language use, typically occupied by language activists with puristic attitudes, can lead to “revitalization-induced” language change, too.

Our investigation into variation and change in function words is based on a comprehensive corpus of spoken or written text modi of formal or informal registers originating from speakers for whom we also have comprehensive speaker biographies. Our transcribed (in Standard Kildin Saami orthography) spoken text data as well as the written text data are stored in a similar XML format, using ELAN.¹ The program allows for transcription, calculation of basic frequency statistics and the creation of concordances in addition to extracting, coding and preparing statistical analysis using R.² Different to many endangered language documentation projects, which annotate manually or semi-manually, we apply more automated corpus data annotation, specifically a part-of-speech tagger based on finite state transducer technology and programmed in collaboration with the Center for Sámi Language Technology³ (cf. Blokland, Gerstenberger, Fedina, Partanen, Rießler, and Wilbur 2014+).

The quantitative evaluation of data from our corpus provides significant evidence for proving preliminary claims by Zhivotova (2014a) and Zhivotova (2014b) that the use of borrowed vs. native function words is not dialectal or the result of individual speaker preferences, but determined by the choice of text modi or registers: whereas the native formatives are used consistently in formal written texts, borrowed formatives occur most typically in informal spoken texts. Furthermore, the degree of speakers’ language loss is also significant: fully competent speakers use borrowed function words much less regularly, which is significant even in informal speech.

We can also show that “purist variants” (Rießler and Karvovskaya 2013) are relevant for sociolinguistic description. In small speech communities undergoing revitalization, such variants might, in fact, be the potential survivors of language variation and change. Therefore, the documentary linguist needs to be sensitive to all kinds of variation in the data in order to make the documentation as complete and useful as possible.

Version of November 30, 2014; this work is licensed under a Creative Commons Attribution–NoDerivatives 4.0 International Licence. Licence details: <http://creativecommons.org/licenses/by-nd/4.0/>

¹<https://tla.mpi.nl/tools/tla-tools/elan/>

²<http://www.r-project.org>

³<http://giellatekno.uit.no>

References

- Blokland, R., C. Gerstenberger, M. Fedina, N. Partanen, M. Rießler, and J. Wilbur (2014+). “Language Documentation meets Language Technology”. Abstract submitted to The First International Workshop on Computational Linguistics for Uralic Languages, Tromsø, 16 January 2015.
- Blokland, R. and M. Rießler (2011). “Saami-Russian-Komi contacts on the Kola Peninsula”. In: *Language contact in times of globalization*. Ed. by C. Hasselblatt, P. Houtzagers, and R. van Pareren. Studies in Slavic and General Linguistics 38. Amsterdam: Rodopi, pp. 5–26.
- Karvovskaya, E. (2011). “Focus particles in endangered languages”. In: *Proceedings of Conference on Language Documentation and Linguistic Theory 3*. Ed. by P. K. Austin, O. Bond, D. Nathan, and L. Marten. London: Department of Linguistics, SOAS, pp. 161–170.
- Kotcheva, K. and M. Rießler (2015). “Clausal complementation in Kildin, North and Skolt Saami”. In: *Semantic functions of complementizers in European languages*. Ed. by K. Boye and P. Kehayov. Empirical Approaches to Language Typology. Berlin: de Gruyter Mouton. In press.
- Nagy, N. and M. Meyerhoff (2013). “Extending ELAN into quantitative sociolinguistics”. ePoster presented at ICLDC 3, University of Hawai‘i at Mānoa, 28 February–3 March 2013.
- Rießler, M. (2007). “Kildin Saami”. In: *Grammatical borrowing in cross-linguistic perspective*. Ed. by Y. Matras and J. Sakel. Empirical Approaches to Language Typology 38. Berlin: Mouton de Gruyter, pp. 229–244.
- (2009). “Loanwords in Kildin Saami, a Uralic language of northern Europe”. In: *Loanwords in the world’s languages. A comparative textbook of lexical borrowing*. Ed. by M. Haspelmath and U. Tadmor. Berlin: de Gruyter Mouton, pp. 384–413.
- (2014). “Vom Matthäusevangelium zur Wikipedia. Medien für das bedrohte Kildinsaamische”. In: *Medien für Minderheiten*. Ed. by R. Kaluweit, M. Álba, and P. Metzger. Freiburger Romanistische Arbeiten 4. Freiburg: Rombach. In press.
- Rießler, M. and E. Karvovskaya (2013). “Purism in language documentation and description.” In: *Keeping languages alive. Documentation, Pedagogy and Revitalization*. Ed. by M. Jones and S. Ogilvie. Cambridge: Cambridge University Press, pp. 83–97.
- Zhivotova, E. (2014a). “Function words in a situation of language contact. Kildin Sámi and Russian”. poster presented at The Young Linguists’ International Conference, Stockholm, 28 June 2014.
- (2014b). “Function words in a situation of language contact. Kildin Sámi and Russian”. Conference paper presented at The 2nd Silesian Meeting of Young Scholars, Szczyrk, Poland, 6–7 November 2014.

This paper aims to determine the capacities of morphology, syntax, and semantics of anticausative constructions in Kartvelian and to find out to what extent they operate together.

Kartvelian anticausatives typically represent bivalent predicates derived from causatives by means of (i) deleting the CAUSER from the proposition and promoting the CAUSEE to the subject position with the semantic properties of an EXPERIENCER-like subject whose control over the action is substantially decreased; (ii) changing the case-marking pattern from tense-dependent ERG-NOM-DAT (in the aorist group of tenses) to the tense-independent one, viz. DAT-NOM; (iii) changing the cross-reference model from active one to inactive in the morphology of the verb, including the application of the so-called “relative version” marker. These operations result in the emergence of prospective-like intentional statives which differ from their parent causative verbs both formally (verbal morphology, case-marking alignment) and semantically (actional class of the verb, semantic roles assigned to the arguments). Such verbs represent the only option to downgrade a causative and can by no means be viewed as passives, as (i) the CAUSER argument is not recoverable and (ii) the actional properties of the verb change (DYN > STAT).

Examples from Georgian illustrate this type of verbal derivation:

- (1) *gogona-Ø bič'-s leks-s Ø-a-c'er-in-eb-s*
 girl-NOM boy-DAT verse-DAT IO₃-VER_L-write-CAUS-SM-S₃SG

‘The girl makes the boy write a verse’

- (2) *bič'-s Ø-e-c'er-in-eb-a leks-i*
 boy-DAT IO₃-VER_R-write-CAUS-SM:STAT-S₃SG.INACT verse-NOM

‘The boy **is very much like to and is about to** write / **cannot help** writing a verse’

This phenomenon is paralleled in two sister languages of Georgian, viz. Megrelian and Svan, and has so far been scarcely discussed in the special literature (cf. e.g. two lines on *v-e-c'er-in-eb-i* in Šaniže 1953: 381), most likely, due to (i) the questionable status of such constructions in Standard Georgian and (ii) the general tendency to primarily focus on phenomena in the Kartvelian languages other than Georgian only if they have been described in respected grammars of Georgian.

The talk will provide a comparative account of the phenomenon in issue against the background of typological approaches to valency (cf. e.g. Baker et al. 2012; Dixon & Aikhenvald 2010; Plungian 2011) and focus upon the semantics of the arguments of the anticausative derivatives.

Abbreviations

CAUS	causative	NOM	nominative
DAT	dative	S ₃ SG	3 rd person singular subject marker
DYN	dynamic	SM	series marker
ERG	ergative	STAT	stative
INACT	inactive	VER _L	locative versionizer
IO ₃	3 rd person indirect object marker	VER _R	relative versionizer

References

- Baker, M., Safir, K., & J. Sikuku. 2012. Sources of (a)symmetry in Bantu double object constructions, in: *Proceedings of WCCFL 30* (Santa Cruz)
- Dixon, R. & A. Aikhenvald (eds.). 2010. *Changing Valency. Case Studies in Transitivity*. Cambridge University Press.
- Plungian, V. 2011. *Vvedenije v grammatičeskiju semantiku: Grammatičeskije značenija i grammatičeskije sistemy jazykov mira* [Introduction to Grammatical Semantics: Grammatical Meanings and Grammatical Systems in the Languages of the World]. Moscow: RGGU Press.
- Šaniže, A. 1953. *Kartuli gramat'ik'is sapuzvlebi. I. Morpologia* [Foundations of the Georgian Grammar. I. Morphology]. Tbilisi University Press.

The impact of language contact on Mosetén and Pirahã

Jeanette Sakel (University of the West of England, Bristol, UK)

When I set out to conduct fieldwork on the indigenous South American language Mosetén, I sought out elderly speakers from remote areas, hoping to record a variety of the language not influenced by Spanish. I soon realised that this effort was futile. Spanish is part of the linguistic reality of Mosetén and has penetrated many lexical and grammatical areas of the language. However, even indigenous languages with restricted contact and bilingualism exhibit contact phenomena from other languages. The Amazonian language Pirahã, spoken in a predominantly monolingual society, displays considerable contact influence from Portuguese.

Yet, the contact phenomena we find in both languages are quite different from one another. Contact with Spanish has had a major impact on Mosetén lexicon and grammar. Mosetén has borrowed many function words (Matras & Sakel 2007) in their morpho-phonological form (so-called *matter* loans). Other areas of grammar have also been remodelled to follow the patterns of Spanish, such as the gender agreement system (as a so-called *pattern* loans).

In Pirahã, on the other hand, contact influence is restricted to *matter* loans in the lexicon. These are all novel concepts introduced through contact with outsiders and present instances of *insertion* in terms of Haspelmath and Tadmor (2009). While these loans seem ubiquitous in spoken Pirahã they have not yet permeated other aspects of the language.

Despite the frequency of loans on the surface of a language, language contact phenomena can be very different depending on the intensity of the contact (Thomason and Kaufman 1988), the level of bilingualism in the community, language attitudes and other factors such as taboos against matter loans (Matras & Sakel 2007). Pattern loans at the level of the grammar will only occur in situations with bilingual speakers, while lexical pattern loans do not always face the same restrictions (Sakel 2007).

In the case of two Amazonian languages Mosetén and Pirahã, it is important to take into account the history of colonisation and, ultimately, the extent of the contact with other languages to fully understand the contact phenomena we find in the languages today. Despite matter loans appearing in both languages, considerably more categories are affected by the contact in Mosetén, which is characterised by ubiquitous bilingualism and a long-term contact history with Spanish.

The two Amazonian languages Mosetén and Pirahã are on opposite ends of the contact experience in terms of the extent of the contact and the degree of bilingualism of the speakers. Yet, both have – at least on the surface – considerable contact influence from Spanish or Portuguese. It is unlikely that the ‘pristine’ language not influenced by other languages I was originally looking for exists. Matter loans can affect languages even without close contact or bilingualism. Also, while some Amazonian languages have very few matter loans from other languages, the less visible patterns of the language may tell a different story about the contact influence from other languages.

Symmetry and asymmetry in voice systems: Evidence from Tagalog and German sentence production

The nature of the voice system of the Austronesian language Tagalog has been debated since over a century now. It has been described as being an accusative, an accusative-like, or an ergative system. Recently, Foley (2008) proposed that Tagalog and other Philippine languages belongs to their own type, the Philippine type. This kind of voice system is described as being “symmetric” in that there are no unmarked voice forms. This contrasts with “asymmetric” voice systems, like the more familiar nominative-accusative or ergative-absolutive systems, that have marked and unmarked voice forms. In German, semantic agents are preferentially chosen as subjects and active voice forms are morphologically unmarked; if the patient is to be chosen as subject, the verb has to be marked for passive. In the “symmetrical” voice system of Tagalog, on the other hand, the sentence-initial verb always overtly marks the semantic role of the “subject” NP via voice affixes so that there is no morphologically unmarked voice form.

This paper investigates whether the grammatical classification of voice systems in “symmetric” and “asymmetric” is also paralleled in language processing. It is hypothesized that cognitive processing load dynamics during sentence production in “symmetrical” voice systems will be similar for different voices as all of them are equally marked; processing load dynamics during the production of marked and unmarked voices in “asymmetrical” voice systems, however, should differ from each other as distinct processes might have to be carried out to prepare the production of voice forms differing in their markedness.

Differential processing load in sentence production in Tagalog and German was investigated using measurements of the size of speakers' pupils as they described line drawings of simple transitive events. Pupil size changes can be linked to cognitive resource allocation (Laeng et al. 2012, Sara 2009) and therefore to the “cognitive effort” it takes speakers to plan and prepare the production of a sentence. Pupillometric data from two eye-tracking experiments on Tagalog and German were analysed to investigate the time course of cognitive resource allocation during sentence production. In German, pictures were described using either active or passive sentences. In Tagalog, pictures were described using either actor voice (where the semantic agent was “subject”) or undergoer voice sentences (where a variety of non-agents, such as patient or beneficent, were the “subject”).

Pupil responses indicate processing differences between voice types in both “asymmetrical” (German) and “symmetrical” (Tagalog) voice systems. In German, speakers' pupil size increased earlier when they produced passive sentences as compared to active sentences, indicating that processing load increases earlier when producing a marked compared to an unmarked voice form. In Tagalog, pupil size increased more during the production of undergoer voice compared to the production of actor voice. However, there were no differences in the shape of the pupil response curves. This suggests that the same cognitive processes are carried out during the production of the equally marked actor and undergoer voice forms in Tagalog (although the production of undergoer voice forms appears to be more effortful), whereas different processes are carried out during the production of (unmarked) actives and (marked) passives in German.

This paper contributes to research on the typology of voice systems by investigating the manifestations of a grammatical distinction between “symmetrical” and “asymmetrical” voice systems in language processing, by analysing the time course of cognitive resource allocation while speakers produce sentences with different voice forms in Tagalog and German.

Finding out what really drives change: The importance of multivariate analysis in historical linguistics

In addition to tracing language change through time, a fundamental objective of Historical Linguistics is to provide explanations, i.e. to identify what exactly triggers and drives individual changes. Reanalysis, frequency effects, functional load and language contact are just some of the causes that have been shown to play an important part in processes of change, but what all instances of language change have in common is that they imply a stage of variation between two (or more) alternatives. Whilst these alternatives are functionally equivalent or at least overlap in part their functional spectrum, speakers nevertheless do not necessarily choose between them at random; a variety of contextual factors, both linguistic and extralinguistic, make it more or less likely for a speaker to pick one or the other of the available alternatives every time he has the choice. It is these contextual factors that can eventually determine the fate of the competing variants; for instance, if one of the alternatives is preferred in a specific high-frequency context, this may cause it to become the default choice and eventually oust its competitor(s).

Sometimes it is relatively clear which factor triggers the choice of one variant over the other, but in the majority of cases it is less evident, as (a) the relevant factors may not have any intuitively obvious relation to the structure we are interested in, (b) the choice is not predictable, as it is merely *more likely* that one of the variants will occur in a specific context, and (c) not every intuitively apparent correlation between a specific context and a specific variant is statistically significant. Therefore, the only scientifically sound way to identify relevant contextual factors is a multivariate analysis in which the widest possible range of factors is considered and checked for relevance. Whilst this method is widely used in Sociolinguistics, it has only recently been applied to diachronic corpora, partly due to the fact that written texts are, rightly, not usually considered to be representative of actual language usage, which is where most changes occur; this problem can, however, be resolved at least in part by using corpora consisting of linguistic material with a high degree of ‘communicative immediacy’ (*Nähesprache*, cf. Koch & Oesterreicher 1985), such as personal letters and diaries that approximate the spoken language.

In this paper, the usefulness of multivariate analysis to identify the factors driving language change is exemplified using a diachronic corpus of personal documents written in Spanish between the 16th and the 20th century, examining variation between a number of competing modal and adverbial constructions. Some rather unexpected factors, such as the grammatical person of the subject or the construction’s phonological environment, are shown to have played a significant role in the evolution of these areas of Spanish grammar over the past centuries, triggering the gradual demise of one variant or the rise to dominance of another.

Finally, it is argued that by compiling a cross-linguistic inventory of potentially relevant contextual factors drawn from multivariate analyses of a wide range of languages, it will be possible to establish which kind of factor is likely to play a role in a particular type of grammatical development, which in turn may help to provide more reliable explanations for changes that have taken place in languages of which we have no diachronic records.

Reference

Koch, Peter & Wulf Oesterreicher. 1985. “Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte”, *Romanistisches Jahrbuch* 36: 15-43.

Noun to verb ratio and word order

Frank Seifart^{1,2}, Jan Strunk¹

¹University of Amsterdam, ²Max-Planck-Institute for Evolutionary Anthropology

In this paper we investigate the relation between Basic Word Order, as a structural-typological feature of languages, and patterns in the use of nouns and verbs in discourse, from a cross-linguistic perspective. We test the hypothesis that discourse in verb-final languages displays a higher noun-to-verb ratio than in non-verb-final languages, i.e. that speakers of verb-final languages use a larger number of noun tokens relative to verb tokens, as well as a larger number of different nouns (i.e. noun types) relative to the number of different verbs (i.e. verb types).

We investigate this hypothesis using corpora of spoken language from nine genealogically and areally diverse languages (see Table). For each language, our corpora comprise a total of around 30,000 words, from between 63 (Baure) and 8 (Texistepec) individual texts (mostly narratives, some procedural texts and conversations). All data are annotated manually for part-of-speech by experts on the language. This set of languages includes three verb-final languages and three languages with nonfinal position of the verb. We find that indeed speakers of verb-final languages tend to employ more noun types and tokens relative to verb types and tokens than speakers of non-verb-final languages (at left, in blue).

LANGUAGE	FAMILY	LOCATION	WORD ORDER	V-(NON)FINAL
Chintang	Tibeto-Burman	Nepal	SOV	final
Lamunxhin Even	Tungusic	Siberia	SOV	final
Bora	Boran	Peru	SOV/OSV	final
Hoocąk	Siouan	North America	SOV	final
N uu	Tuu	South Africa	SVO	nonfinal
Baure	Arawakan	Bolivia	VSO	nonfinal
Texistepec	Mixe-Zoquean	Mexico	VSO/VOS	nonfinal
English	Indo-European	North America	SVO	nonfinal
Dutch	Indo-European	Europe	SVO	nonfinal

We suggest that the key to explaining this finding lies in the relative informativeness of nouns vs. verbs as a function of the number of different nouns vs. verbs used in a text: using many different nouns in one text means that these nouns are relatively informative, or specific, while using only few nouns means these nouns are probably relatively general. When for structural-typological reasons verbs come at the end of a clause, speakers are prompted to provide more information about the contents of the clause in parts-of-speech that precede the verb, namely nouns, in order to create coherent discourse. This is in line with Polinsky's (2012) finding that languages with different Basic Word Orders have different ratios of nouns and verbs in their lexica. However, by using corpus data, we provide a direct account for a specific relation between grammatical structure (word order) and usage patterns.

Reference

Polinsky, Maria. 2012. Headedness, again. *Theories of Everything*. In Honor of Ed Keenan, ed. by Thomas Graf, Denis Paperno, Anna Szabolcsi, and Jos Tellings, 348–359. *UCLA Working Papers in Linguistics* 17. Los Angeles: UCLA Department of Linguistics. http://www.linguistics.ucla.edu/faciliti/wpl/issues/wpl17/papers/40_polinsky.pdf.

Classifiers in Kilivila: introducing referents and keeping track of them

Gunter Senft

MPI for Psycholinguistics Nijmegen

After a brief introduction of the system of classifiers (Malinowski's 'classificatory particles') in Kilivila, the Austronesian language of the Trobriand Islanders of Papua New Guinea, I will outline their main functions, especially the functions of introducing referents in discourse and keeping track of these referents, thus preserving coherence in discourse. With respect to the organization of discourse and conversation, the referent introducing function and the discourse deictic, anaphoric reference function of these classifying formatives are of special interest - both from a language specific point of view and from a cross-linguistic and more comparatively oriented point of view. A number of research questions illustrate how this brief case study could stimulate further research on how classifier systems in other languages – including sign languages – fulfil these functions of introducing referents and keeping track of them and which other (additional) means in these languages are used to fulfil them.

Leipzig, May 2015

Between language and thinking

Abstract

If a grammar is a modular mental mapping system from semantic to linguistic structures (Seuren 1996), then the question arises of how the former come into being. Slobin (1987) speaks of ‘thinking-for-speaking’, Levelt (1989) of ‘microplanning’. Lexical selection is a central part of that process, with all the syntactic and morphological consequences thereof. But there is more. A semantic structure that is input to the grammar must specify the scope of logical and nonlogical scope-creating operators, so that the grammar can place them in the sentence in such a way that the hearer will be able to grasp the intended scope relations. Moreover, the ‘packaging’ of the information concerned, including topic-comment structure and sentence intonation, must satisfy conditions of context, situation and rhetorical preference. Yet hardly any integrated research into this area of sentence production and comprehension has so far been conducted, despite the vast amounts of work on ‘information packaging’ in pragmatics, on sentence intonation in phonetics, on positive and negative polarity items in formal semantics, or on NEG Raising in grammar—to mention just a few problem areas that still lack a solution but may well find one in terms of an integrated theory of thinking-for-language or microplanning. In my view, the answer to all or many such questions may well be found in the assumption that semantic structures are language-specifically refashioned during the thinking-for-language (or microplanning) phase before they are passed on to the grammar. The packaging of central thought content for linguistic expression and the unpacking of incoming linguistic goods for delivery to central thought seems to be quite an industry, about which little is known as yet. Specific examples, arguments and analyses are presented.

References

- Levelt, Willem J.M. (1989). *Speaking. From Intention to Articulation*. Cambridge MA: MIT Press.
- Seuren, Pieter A.M. (1996). *Semantic Syntax*. Oxford: Blackwell.
- Slobin, Dan (1987). Thinking for speaking. In: Jon Aske, Natasha Beery, Laura A. Michaelis and Hana Filip (eds), *Berkeley Linguistics Society. Proceedings of the 13th Annual Meeting. Febr. 14–16, 1987*. Berkeley CA: Berkeley Linguistics Society: 435–45.

Pieter A.M. Seuren
 Max Planck Institute for Psycholinguistics
 Nijmegen
 The Netherlands
 pieter.seuren@mpi.nl

Encoding direct object in Enets

Our paper discusses the ways a direct object can be encoded in Enets, a Uralic, Samoyedic language (we use data from both Enets dialects, Forest Enets = FE, Tundra Enets = TE). The study is based on a corpus of texts of ca. 40 hrs, all of which were collected and transcribed with the support of HRELP and MPI.

When encoding a 3rd person direct object, Enets has the following options to choose from: either to encode it with an object cross-reference affix on the verb, or to encode it with an overt noun marked by Nominative or Oblique, or to use both. Object cross-referencing in the verb is not obligatory and involves cross-referencing the number of the object only (e.g. FE *pɔnʲiŋa-d* [use(ipfv)-2SG.S] ‘you use’, *pɔnʲiŋa-r* [use(ipfv)-2SG.SOsg] ‘you use it’, *pɔnʲiŋ-i-z* [use(ipfv)-SOpl-2SG.SOpl] ‘you use them’). Enets Nominative and Oblique cases are formally different only in two cases: first, when a noun belongs to an alternation inflectional class and thus uses two distinct stems for Nominative and Oblique (e.g. FE *mɛʔ* ‘tent.NOM’, *mɛz* ‘tent.OBL’), and second, when a noun has a possessive affix (e.g. FE *mɛ-l* [tent-NOM.SG.2SG] ‘your tent (nom)’, *mɛ-t* [tent-OBL.SG.2SG] ‘your tent (obl)’). Anyway, there is always a choice: to use or not to use the object cross-reference, to use Nominative or Oblique case for direct object. This paper studies the grounds beyond the choice.

Cross-referencing 3rd person direct object in the verb is almost obligatory when the object NP is absent from the clause, as in (1). Cross-reference is not used when the direct object has a destinative marker, as in (2). With a non-destinative direct object, the choice of the verbal form has an evident correlation with definiteness (this trend is stronger in TE than in FE): a definite NP is normally cross-referenced, as in (3a), while an indefinite NP normally is not, as in (3b).

Oblique form of non-possessive nouns is obligatorily used in the case of generic direct objects, as in (4a), though it is also attested with non-generic NPs, cf. (4b). Nominative form is used to mark prominent NPs expressing key participants in the discourse structure. Cf. (4c), where the speaker emphasizes the object the tale is about, or (5) where the crucial objects for the purification ritual are introduced in a story about this ritual. This distribution between Nominative and Oblique direct objects is worth particular attention, since it goes against the general cross-linguistic tendency of the differential object marking to use the nominative form for less prominent objects (e.g. Aissen 2003).

The choice of Nominative vs. Oblique form of possessive forms is mainly based on the number of the noun and on the person of the possessor: plural direct objects take Nominative possessive affixes, while singular direct objects with the 1st person possessor take Nominative affixes, as in (6a), and singular direct objects with possessors of the 2nd and the 3rd persons take Oblique affixes, as in (6b). The second important factor for possessive forms is the distinction between imperative vs. non-imperative clauses: in imperative clauses Nominative is normally used even when it is not expected from the number and person factor, as in (7). For possessive destinative forms two other factors come into play: in FE unexpected Oblique can optionally be used in other-benefactive contexts, as in (8a), in TE unexpected Oblique can optionally be used with specific direct objects, as in (8b).

Since the parameters involved in nominal and verbal encoding of direct object are not completely independent from each other, their mutual distribution is also worth studying. E.g. the parameter of prominence is related to the parameter of definiteness, as definite NPs tend to be prominent, and (4b) vs. (4c) illustrate both the choice of the case of the noun and the choice of the verbal cross-reference.

References

Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *Natural Language & Linguistic Theory* 21. 435–483.

Abbreviations

1, 2, 3 – 1st, 2nd, 3rd person, ABL – ablative, CONN – connegative, DEST – destinative, DU – dual, FOC – focal, FUT – future, HAB – habitual, IMP – imperative TAM-series, ipfv – imperfective, NEG – negative verb, NOM – nominative, OBL – oblique, pfv – perfective, PL – plural, PRF – perfect, PST – past TAM-series, PTCL – particle, PTCP.ANT – anterior participle, PTCP.SIM – simultaneous participle, S – subject indexation series, SG – singular, SOpl – subject-object indexation series for plural object, SOsg – subject-object indexation series for singular object.

Examples

- (1) *ese-jʔ* *tʃi* *peri* *baziʔ-ubi-zaf* FE
 father-NOM.SG.1SG here always tell(pfv)-HAB-3SG.SOsg.PST
 ‘My father always told about him.’
- (2) *ugaʎe-zɔ-d* *teza-da-d* TE
 coal-DEST.SG-OBL.SG.2SG bring(pfv)-FUT-2SG.S
 ‘You will bring coal.’
- (3) a. *kunɔ-xɔa* *tʃike* *bese* *naʔa-da-zutʃ* TE
 where-FOC this money send(pfv)-FUT-3PL.SOsg.PST
 ‘They’d better send this money somewhere.’
- b. *bese* *nʎe-ʔ* *teza-dɔʔ* TE
 money NEG-3PL.S bring(pfv)-FUT.CONN
 ‘They will not give money.’
- (4) a. *ɔdiz* *ooŋa-ʔ* FE
 verdure.OBL eat(ipfv)-3PL.S
 ‘They eat grass.’ (speaking of a species of birds)
- b. *ɔbu-xoɔ* *ɔdiz* *koɔ-xiʔ* FE
 what-FOC verdure.OBL find(pfv)-3DU.S
 ‘They found a plant.’ (the plant was not mentioned)
- c. *iŋi-ziʔ* *ko-ʔ* *ɔdiʔ* FE
 certainly-3DU.SOsg find(pfv)-CONN verdure.NOM
 ‘Of course, they found the plant.’ (the aforementioned plant, important for the story)
- (5) *djuuʔ* *tɔʔ* *mua* TE
 fat.NOM there do(pfv).3SG.S
 ‘She put some fat there (into the basin).’
- <...>*sen* *mii-gɔa* *tabuʔ* ... *tabuʔ* *sɔʔɔ,*
 how_much what-FOC hair.NOM hair.NOM take_out(pfv).3SG.S
 ... *tʃinadi* *satʃeza-za*
 now light(pfv)-3SG.SOsg
 ‘She took out some wool, and then she lighted it all up,’
- (6) a. *kunʎ* *poga-jʔ* *berta-da-u* FE
 how fishing_net-NOM.SG.1SG throw(pfv)-FUT-1SG.SOsg
 ‘How will I leave my net?’
- b. *poga-d* *tetti-r* FE
 fishing_net-OBL.SG.2SG measure(pfv)-2SG.SOsg
 ‘You have measured your net.’
- (7) *abaa,* *pɔʎidʎu-r* *tɔza-z* FE
 elder_sister awl-NOM.SG.2SG bring(pfv)-2SG.SOsg.IMP
 ‘Give me your awl.’
- (8) a. *kɔru-zo-niʔ* *ta-ʔ* FE
 knife-DEST.SG-OBL.SG.1SG give(pfv)-2SG.S.IMP
 ‘Give me a knife!’
- b. *tɔzɔ* *tʃikɔ-xɔzɔ* *kasa-zɔ-niʔ* *mua-zɔʔ* TE
 so this-ABL.SG man-DEST.SG-OBL.SG.1SG take(pfv)-1SG.S
 ‘Then I got married (lit. took a husband for myself).’

1-3 May 2015 (Leipzig, Germany)

Closing conference of the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology

Abstract

Paul Sidwell

Dept. Linguistics, CAP, Australian National University

Title: **A comprehensive phylogenetic analysis of the Austroasiatic languages**

The author has been investigating the classification and phylogeny of Austroasiatic using computational methods (with assistance from Simon Greenhill and Russell Grey) since 2009. The first results were presented publically at the 2009 ICAAL meeting in Bangkok and reported in print in Sidwell and Blench (2011) and Sidwell (2013), and results of later trials in Sidwell (2014). Since late 2014 the dataset has been extended and revised to cover 200 words for more than 100 languages. Each time that the data has been analysed (producing both NeighborNet analyses and phylogenetic trees) the results have consistently indicated a strongly branching tree with only weak or contradictory indications of nesting among the branches.

The earliest experiments were done with just enough data – representing 2 or 3 language per branch – to test the branching hypotheses and to obtain general indications of the typology of the tree. The most recent analyses have sought to comprehensively cover all known Austroasiatic branches and sub-branches. Additionally, the use of more extensive vocabulary allows us to experiment with subsets of the data, permitting testing of various linguistic hypotheses, plus testing aspects of the phylogenetic methods themselves.

The experiments appear to confirm that computational phylogenetics can deliver robust results comparable to more traditional comparative methods, as well as providing other useful indications of language evolution (such as contact effects). The present results have both converged with, and helped to refine, the comparative reconstruction.

References

- Sidwell, Paul. 2014. Austroasiatic Classification. In *The Handbook of Austroasiatic Languages*, Mathias Jenny and Paul Sidwell (eds.). Leiden, Boston: Brill. Pp. 144-220.
- Sidwell, Paul. 2013. Issues in Austroasiatic Classification. *Language and Linguistics Compass* 7/8 (2013): 437–457.
- Sidwell, Paul. and Roger Blench. 2011. The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis. In *The Dynamics of Human Diversity*. N. J. Enfield (ed.). Canberra: Pacific Linguistics. pp.315-344.

How to lose an argument in Nama: Transitive and intransitive verbs in a Papuan language of Southern New Guinea

Jeff Siegel

This paper describes two different valency reduction strategies and their functions in Nama, a Papuan language of the Nambu Branch of the Morehead-Upper Maro family. There are basically three types of verbs in Nama – transitive, agentive intransitive and patientive intransitive. These differ according to the number of arguments and how they are morphologically marked.

Transitive verbs mark two arguments: verb final sets of suffixes index the person and number of the agent (A). These follow various suffixes that indicate tense and aspect. Verb initial sets of prefixes index the person and number of the patient (P), which is the semantic undergoer or beneficiary. The agent NP is indicated by an ergative suffix or an ergative pronoun:

yemofem fá ekmayènd
yèmo-fem fá e-kèm-ai-énd
3.ERG-NSG 3.ABS 3NSG.P-lie.down-PERF.REMOTE-3NSG.A
'They laid them down (e.g. the posts) long ago.'

Nearly all verbs can lose the agent argument to become agentive intransitive verbs (unergatives, unaccustatives or reflexives). This is done with the addition of the detransitivising prefix *á-* or *a-*. In such cases, the argument is treated as an actor or theme, and has the unmarked absolutive form. But it is indexed on the verb by the same set of suffixes as those for agents of transitive verbs (A). Thus there is a split system with regard to nominative-accusative and ergative-absolutive marking. Special prefixes indicate the absence of a grammatical patient:

fá nákmayènd.
fá n-á-kèm-ai-énd
3.ABS ØP-DETRANS-lie.down-PERF.REMOTE-3NSG.A
'They were (in the process of) lying down long ago.' or
'They laid themselves down long ago.'

A small set of verbs (including **kèm* 'lie down') can also become patientive intransitive verbs. In this type of verb (which includes postural and copular verbs), the NP argument again has the same absolutive form as the patient in transitive sentences but is treated as an experiencer and is indexed by the same sets of prefixes as those for patients of transitive verbs (P). Patientive intransitive verbs are differentiated from the other verb types by the absence of agent/actor indexing suffixes and a separate set of tense/aspect suffixes:

fá ekmangèrmèn
fá e-kèm-angèrmèn
3.ABS 3NSG.P -lie.down-PERF.REMOTE
'They were (in the state of) lying down long ago.'

The paper gives more examples of each type of intransitive verb and their various functions.

A new look at the Kayanic branch: data from Merap

This paper discusses the Kayanic branch of the North-Borneo phylum and on the light of the little known language of Merap spoken in the Indonesian Borneo, tries to define a lower level branch. It will describe synchronically the phonology of Merap and diachronically the changes that may have occurred from Proto Malayo-Polynesian and how these changes are shared by other languages of the area through a comparative analysis.

Very little is known about the this language spoken by a group of alleged Kayanic people in the today's province of Kaltara (Kalimantan Utara) in Indonesia. The Merap number possibly around 1000 people spread in three main villages in the area of the Malinau and Tubu Rivers in the North-east part of the Indonesian Borneo and are all surrounded by non Kayan people, that is speakers of Kenyah, Punan Tubu and some Dusunic languages who all consider Merap as unintelligible for its very striking phonologic features.

The importance of Merap phonological features was mentioned in Guerreiro (2009) and details of its history are found in some ethno-historical descriptions by Sellato (1995, 2001), Cesard (2009), Okushima (2008) and Kaskija (2012).

Merap people consider themselves as the original Kayan people though their original place is located in the upper part of the Bahau River occupied by Kenyah groups. Indeed their language has much in common phonologically with Pua' and Ngorek and lexically with proper Kayan languages like Uma' Leken, though a relation with some Kenyah variants is documented too.

Within the most recent classifications of Borneo languages in the Kayanic branch are listed Kayan proper, Muller-Schwaner Punan, Murik and Modang. It is proposed that the Modang branch in Ethnologue and in other Borneo classifications has to be expanded to include Ma' Pnaan and Punan Segah, Merap, Bahau, and other languages whose speakers claim the upper Bahau as their original place like the Hueng Bau, the Ngorek and the Pua'. Provided that Murik now spoken in Sarawak, is a variant of Ngorek, it will be tested the possibility of including it as well in this newly defined Modang-Bahau branch.

Some of the features that will be discussed in this paper are the lack of occurrence of the stops p, t, k in final position, the diphthongization of final vowels u and i, the absence of velar nasals in final position, the presence of an uncommon voiceless unreleased palatal stop in final position, the palatalization of final nasals, and the process of monosyllabization of bisyllabic roots.

In Merap, the Proto-Malayo-polynesian voiceless obstruents in final position change systematically in glottal stop preceded by vowels that are in-glided as it can be observed in PMP *puk ‘hair’ > M *bauəʔ*, PMP *anak ‘child’ > *neiəʔ*, PMP *paʔit ‘bitter’ > M *paʔiaj*.

The vowels u and i in final position are diphthongized into ou and ei and the velar nasal is realized into an aspirate h like in PMP *ejung ‘nose’ > M *ruəh*. This last example shows also that the process of monosyllabization of bisyllabic roots is systematic. Other examples are *tlau* ‘three’, *mlah* ‘red’, *hni* ‘mother’, *kiu* ‘elbow’, *teiʔ* ‘calf of leg’, *ngan* ‘arm’ that descend from bisyllabic PMP roots.

Some of these changes are shared by languages like Òma Lóngh Kenyah known for its idiosyncratic features like the voiceless unreleased palatal stop in final position, the change of nasal velars in final position in aspirates (Soriente 2006). The monosyllabization is shared with Ma’ Pnaan (Soriente 2012) and with other Kayanic languages like Wehea and Busang (Guerreiro 1996).

The discussion of Merap phonologic features, the very first step of a language documentation project *in fieri*, will contribute to a description of this undocumented language and on the other hand to better understand its relation to the other languages of the area.

System of orientation in Dolgan
Eugénie Stapert, Leiden University

This paper investigates the differences in the system of orientation between Dolgan and Sakha. Dolgan is a Turkic language spoken by approximately 7,000 people on the Taimyr Peninsula in north-central Siberia, and is linguistically very closely related to its neighbouring language Sakha. However, there are some significant differences between the two languages, both in grammar and in lexicon. One of these differences concerns the system of orientation.

In Sakha there are four different terms to indicate the cardinal directions north, west, south and east. In Dolgan, on the other hand, there are only two terms, the first one (*muora*) grouping together north and west, and the second (*ta:s*) grouping together south and east. The terms themselves are found in Sakha as well, but with a different meaning, suggesting that the conceptual shift from a four-way to a two-way system was accompanied by a semantic shift in the terms that were employed to describe this new system.

The aim of this paper is to investigate how these differences between Dolgan and Sakha may be explained and in particular whether they could have arisen as a result of contact with the neighbouring ethnic groups (Evenks and Nganasan).

Intragenetic Typology: The case of Romance

(Patrick O. Steinkrüger; Humboldt-University of Berlin)

In this talk I try to apply the concept of intragenetic typology (Kibrik) to Romance languages. At a first glance, the putting together of typological and diachronic criteria to classify languages seems to be contradictional, but it makes sense, as it will be demonstrated during the talk. Since, traditionally, linguistic families are classified following primarily phonological and lexical criteria, also creole languages based on Romance (Michaelis et al. in the *APiCS*) will be taken into consideration.

Two main topics will be discussed: A) We try to select “core features” and “peripheral features” of Romance, e.g. inventory of phonemes and the complexity of syllable structure, number of genders, systems and use of the pronoun paradigm, position of the article, the position and combination of object clitics (in different constructions), DOM, auxiliary selection, cleft structure of questions, negative concord, (only some of the features are considered in Bossong 2008). Additionally, the method of implicational argumentation will be mentioned (see the attempts by Körner). B) The sampling of linguistic systems; what are the relevant “languages”? The so called *Ortsmundarten* (e.g. local varieties as Mirandese), regional substandards (mostly lingua francas used in historical areas as Piedmontese) or standard languages (almost *Plansprachen* as “Italiano”, “Français” or “Rumantsch Grischun”). It is important to mention these differences, because there is for example no standard language with double subject agreement (as in Northern Italian varieties).¹ There is also the cartographic challenge for typological studies to represent the distribution of certain features by points (as in the *WALS*) or by areas (currently and historically) on the earth.

Specifically formulated questions referring to Romance are: Does a “Romance type” exist ? (a question raised by Coseriu). And are there (soft/hard) implications observable (see Körner)? What are the main typological shifts from Latin to Romance? Interesting questions for general linguistics are: What can be considered generally the criteria for defining language families: only historical and diachronic criteria based on sound and lexical material and probably some morphological similarities (the “19th century-method”), exclusively typological ones based on structural similarities (also caused by divergence in different contact situations) and finally the consideration of extralinguistic criteria as prestige and power of cultures (e.g. Arabic, English, different Greek, Indic, Latin (incl. “Charlemagne area”), Tibetan, Ottoman, Persian, Russian, Sinitic, Spanish, etc. areas) which influence e.g. sound-systems (including by scripts), patterns of word-formation and syntactic structures.

Selected references

Bossong, Georg (2008): *Die romanischen Sprachen: eine vergleichende Einführung*. Hamburg: Buske.

Comrie, Bernard (2008): “Basque, Romance, and areal typology: what do we learn from the World Atlas of Language Structures?”, in Hans J. Döhla et al. (eds.), *Lenguas en diálogo: El iberorromance y su diversidad lingüística y literaria, Ensayos en homenaje a Georg Bossong*. Frankfurt am Main: Vervuert, 55-63.

Kibrik, Aleksandr E. (1998): “Does intragenetic typology does make sense?”, in: Winfried Boeder et al. (eds.), *Sprache in Raum und Zeit: Beiträge zur empirischen Sprachwissenschaft*. Tübingen: Narr, 61-68.

Körner, Karl-Hermann (1982): *“Concordança del participi passat” im Katalanischen und die syntaktische Typologie der romanischen Sprachen*. Berlin: Steiner.

Ramat, Paolo & Elisa Roma (2007) (eds.): *Europe and the Mediterranean as Linguistic Areas. Convergencies from a historical and typological perspective*. Amsterdam / Philadelphia: Benjamins.

Sandfeld, Kristian (1930): *Linguistique balkanique; problèmes et résultats*. Paris : Klincksieck.

¹ It is also astonishing, considering the fact that Romance languages are generally very well described, that until now it does not exist any descriptive grammar of spoken French.

What does speech add to points and vice versa? Preschoolers and their surrounding environment in three cultures

Elena Lieven, University of Manchester
Sabine Stoll, University of Zurich

Pointing is one of the earliest means of communication between children and their surrounding interactional partners. As a result, most research has focused on the interpretation of early points and the expressed mental states of prelinguistic infants. Less is known how pointing is used in early language development. In this study we compare the points and their accompanying speech acts of 2-and-3-year-old children learning Russian, German and Chintang. We use longitudinal naturalistic data comparing the behavior of the children with the behavior of their surrounding adults. We found three main results:

First, pointing is a frequent phenomenon both for the children and adults of all three cultures. However, there were significant differences in the amount of pointing by the children in the three cultures. Chintang children pointed least - at approximately 12 times per hour. This means on average one point every 5 minutes. For German children, there are approximately 30 points per hour which adds up to 1 point every 2 minutes. This is probably due to differences in the number of interactional partners in the different cultures. Chintang children have many more interactional partners meaning that they are unlikely to be as involved in the dyadic exchanges that take place in the German or Russian contexts. It is interesting that adults in all three cultures pointed at similar rates.

Second, in all three cultures a variety of points such as finger points, hand points, head points, etc. was used. Finger points were the most frequent types of points in all the cultures, both for adults and children. A comparison revealed that the children of the three languages behaved significantly differently from each other with Chintang children using significantly fewer finger points than the children of the other 3 languages. Adults, however, showed similar distributions in the types of points.

Third, the relative frequencies of speech acts going along with points do not differ significantly between the cultures. Both children and adults use points most with declaratives followed by imperatives and questions. However, children and adults differ across all three cultures with children using hardly any imperatives with their points. Children of all three cultures also use a considerable number of points without any utterance suggesting that the points substitute for language. This is not found in the adult data. For all three cultures, deictics are by far the most frequent part of speech used while pointing, followed by NPs. However, there is a difference in that both Russian adults and children are significantly more likely to accompany their points with VPs than in either of the other two cultures. These results show that on the one hand, there are some culture- and language-specific differences in the frequency and the type of points used by children but that this is less the case for the adults. At the same time, however, there are striking similarities in the use of points and accompanying speech acts in languages and cultures, which differ as widely as Chintang, Russian and German. These results suggest that detailed analyses are needed to work out the universals and particulars of the combination of speech and gesture across cultures and how this develops.

Spatial interrogatives

CROSS-LINGUISTIC ASPECTS OF AN UNDERSTUDIED PARADIGM

Thomas Stolz, Bremen/Germany

The bulk of the studies dedicated to the grammar of space (Svorou 1993) focus on the linguistic representation of spatial relations in declarative sentences. On this basis, hypotheses are formulated as to the conceptual organization of space-related cognitive systems of human beings (Stolz/Letsrade/Stolz 2014). Universals are postulated and/or typological differences are assumed. Given that the declarative-borne insights are cognitively rooted, it is tempting to assume that all parts of grammar which make reference to spatial categories are organized according to the same principles, i.e. the distinctions made in the context of declaratives must also hold for other sentence modalities.

To test the tenability of this idea, it is necessary to look at a different sentence type - for instance, interrogative clauses (Siemund 2001). The talk marks the beginning of a cross-linguistic project which investigates the properties of spatial interrogatives with special focus on general location. Spatial interrogatives are content questions in the sense of Cysouw (2007). In somewhat antiquated English, the ternary set which involves the spatial interrogatives *where?*, *whither?*, and *whence?* is an example of the paradigms which are scrutinized in the project. As the contributions to Chisholm/Milic/Greppin (1984) show spatial interrogatives are still largely a linguistic *terra incognita*. In my project, it is tested whether or not the distinctions made within the paradigm of spatial interrogatives correspond to those which are typical of spatial relations in declarative sentences. The languages of my sample are classified according to the patterns of correspondence of the distinctions in declarative and interrogative sentences. These findings are evaluated further in terms of a theory of space in language.

References

Chisholm, William & Milic, Louis T. & Greppin, John A. C. (eds.). 1984. *Interrogativity: A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam, Philadelphia: John Benjamins.

Cysouw, Michael. 2007. Content interrogatives in Asénica Campa: corpus study and typological comparison. *International Journal of American Linguistics* 73 (2), 133-163.

Siemund, Peter. 2001. Interrogative constructions. In: Martin Haspelmath et al. (eds.), *Language Typology and Language Universals*. Volume 2, 1010-1028. Berlin, New York: Walter de Gruyter.

Stolz, Thomas & Lestrade, Sander & Stolz, Christel. 2014. *The crosslinguistics of zero-marking of spatial relations*. (= *Studia Typologica* 15). Berlin, Boston: De Gruyter Mouton.

Svorou, Soteria. 1993. *The Grammar of Space*. Amsterdam, Philadelphia: Benjamins.

Double ‘again’-expressions

A derivational affix can occur not alone, but accompanied with another lexical or morphological marker with the same or similar meaning. There are three possibilities of the resulting meaning of such a combination: 1) a “redundant” interpretation (multiple exponence in a broad sense), when the resulting meaning is the same as the primary meaning of a single marker; 2) a “recursive” interpretation, when the primary meaning is applied to the stem twice consecutively; 3) an “iconic” interpretation, when the meaning of multiplicity or intensification is added to the primary meaning. This issue is discussed in the paper based on the data of ‘again’-expressions (Wälchli 2006) (such as *re-* in Romance), cf. (Kulikov 1993) on similar issues on causatives.

All three possible interpretations of ‘again’-expressions are attested across languages, cf. (1) for the “redundant” interpretation, (2) for the “recursive” interpretation and English expression *again and again* for the “iconic” interpretation.

(1) TEPEHUA-HUEHUETLA (Totonac, Mexico), (Kung 2007: 294–298)

?entons maa pastak-**choko-pala**-kan-li

then RPT think-**REP1-REP2**-INS-PFV

‘Then they **rethought** it’. – redundant interpretation (‘again V’)

(2) NOON (Atlantic, Senegal), (Soukka 2000: 171–172)

hay ‘come’ – hay-is ‘come back’ – hay-sis (<is-is> ‘come back again’ – hay-sis-is ‘come back for the third time’ – recursive interpretation (‘again [again V]’)

Moreover the interpretations can compete, cf. (3)–(4) from Russian:

(3) Pročitala <read> vsě i vernulas’ eščë raz sjuda, čtoby **opjat’ pere-chitat’** <again REP-read>

‘She read everything and returned here again to read it one more time’. (Internet) – redundant interpretation

(4) Otpustiv ego, Aratov pere-čel <REP-read> pis’mo, brosil ego na pol... No pogodja nemnogo podnjal i **opjat’ pere-čel** <again REP-read>... (Russian National Corpus)

‘Having let him go, Aratov read the letter again, throw it to the floor... But a bit later he picked it up and re-read it one more time’ – recursive interpretation

Multiple exponence seems to be the most interesting case. It is widely attested and well studied across inflectional markers and it is postulated to be a rare feature across derivational ones (cf. Caballero, Harris 2012). It is the case for the morphological markers, but if one takes into account also the combinations of morphological markers with lexical ones, this phenomenon seems in contrast to be typical. At the same time it is not the same as in case of inflection: derivational markers (or a derivational marker and a lexical one) that are used together tend to be synonymous, but not to have a fully identical meaning, so the resulting meaning is also not absolutely predictable. An interesting case for ‘again’-expressions is the situation when two markers of the ‘again’-domain are combined – a repetitive one (‘one more time’) and a reditive / restitutive one (‘to return into initial point / state’), as in (1) above. In this case the reditive marker tends to be linearly closer to the stem (or to be morphological if the morphological marker and the free lexeme are combined) and the resulting meaning tends to be repetitive and not reditive. It can be related to the fact that the reditive meaning is more concrete, lexically restricted and diachronically initial, while the repetitive one is more abstract, fully productive and tends to develop from the reditive one on later stages of grammaticalization (cf. Rosemeyer 2014).

References

- Caballero G., Harris A. 2012. A working typology of multiple exponence // Kiefer F. et al. Current Issues in Morphological Theory. John Benjamins.
- Kulikov L. 1993. The "second causative": A typological sketch // Comrie B., Polynsky M. (Eds.) Causatives and Transitivity. Amsterdam-Philadelphia: John Benjamins. P. 121–153.
- Kung S. 2007. A descriptive grammar of Huehuetla Tepehua. Ann Arbor: UMI.
- Rosemeyer M. 2014. The development of iterative verbal periphrases in Romance // Paper delivered at SLE-2014, Poznań.
- Soukka M. 2000. A descriptive grammar of Noon: a Cangin language of Senegal. München: LINCOM.
- Wälchli B. 2006. Typology of light and heavy ‘again’, or the eternal return of the same // Studies in Language, 30(1). P. 69–113.

Applicative Constructions and Double Derivation in Shiwilu (Kawapanan)

Pilar M. Valenzuela
Chapman University
valenzuela@chapman.edu

This presentation offers the first account of the applicative system of Shiwilu (jeb), a polysynthetic language from Peruvian Amazonia belonging to the Kawapanan family. In so doing, it treats points of special typological and diachronic interest.

The prolific Shiwilu applicative system comprises seven constructions involving distinct affixes. These constructions can be classified into (a) those that require simple verb derivation and (b) those that require double verb derivation. Within the former group, the verb takes a semantically specialized applicative affix (*ek-* ‘with X,’ *-lapi* ‘leaving X behind,’ *-pa* ‘helping X’) or is marked by *-tu*, a more general valency modification morpheme that may play a locative applicative function. In (1) and (2) below the suffixation of *-tu* increases and reduces the verb valency, respectively. The examples in (3) illustrate the locative applicative function:

- (1) Object addition: *lanpi*’- ‘swallow’ > *lanpi*’-*tu* ‘swallow sth.’
(2) Object reduction: *apu*’- ‘abandon sth./sb.’ > *apu*’-*tu* ‘become detached’

Locative applicative: (the 3rd person is zero-marked when in object function):

- (3) a. *Peks[a]=ek wichi*’-*lli*. b. *Peksa wichi*’-*tu-lli*.
bed=**LOC** sleep-nonFUT.3SG bed sleep-**APPL**-nonFUT.3SG>3SG
‘He slept on the bed.’ ‘He slept on the bed.’

In applicative constructions involving double derivation, a suffix with an applicative-like meaning (*-i* benefactive/malefactive, *-wa* ‘approaching X,’ *-nan* ‘next to X, sparing X’) combines with a verb base that must simultaneously be marked by *-tu*. This is illustrated in (4b). Note that *-nan* occurs between the verb root and the suffix *-tu*:

- (4) a. *Ñiñi*’*wa kitek-llun*. b. *Ñiñi*’*wa kitek-nan-tu-llun*. **kitek-nan-llun*
dog bite-nonFUT.3SG>1SG dog bite-**APP-tu**-nonFUT.3SG>1SG
‘The dog bit me.’ ‘The dog bit (others) next to me, sparing me.’

The distinction between simple and double verb derivation is not among the parameters along which applicative constructions tend to vary crosslinguistically (Peterson, 2007).

Crucially, Shiwilu's only sister language, Shawi (cbt), exhibits simple derivation applicative constructions exclusively. It is proposed here that the double derivation strategy is a Shiwilu innovation that arose through the incorporation of a case-marker or another element into a verb already marked by *-tu*. In this way, Shiwilu practically doubled its number of applicative constructions.

References

Peterson, David A. 2007. *Applicative Constructions*. Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press.

Property words in Oceanic languages

Some Oceanic languages are claimed to lack distinctions between major classes of nouns, verbs, and adjectives; a phenomenon referred to as lexical flexibility (Van Lier & Rijkhoff 2013). Even though the potential lack of a noun/verb distinction is the most controversial aspect of such flexibility, this presentation is concerned with adjectives or, rather, with property words in the Oceanic family.

I present a study of 30 Oceanic languages, in which I investigated a semantically varied range of property words, including words expressing dimensions, values, age, colors, physical properties, human propensities, and experiential states (cf. Dixon 2004). An earlier study by Ross (1998) showed that Oceanic property words display extremely varied morpho-syntactic behavior – both across languages and across different constructions within one language – and that in many cases they split up into multiple classes based on differential distributional properties. The first goal of my presentation is to evaluate Ross’s typology and generalizations.

Secondly, I address two correlations that have been proposed in existing literature on property words. One, put forward by Wetzters (1996) and Stassen (1997), is that languages with a tense system favor “nouny adjectives” (i.e. property words not expressing any verbal categories when used predicatively), whereas languages without grammatical tense have “verby adjectives” (i.e. predicative property words behaving like verbs). The other generalization, proposed by Dixon (2004), holds that languages with nouny adjectives tend to be dependent-marked (at the clause level), while languages with verby adjectives are rather head-marking or zero-marking. Both claims are supported by WALS data on predicative adjectives (Stassen 2013), tense marking (Dahl & Velupillai 2013a, b), and locus of marking (Nichols & Bickel 2013). My balanced sample of Oceanic languages, however, patterns differently: many languages have tense marking and verby adjectives, and all languages with dependent marking have verby adjectives.

One explanation for these findings is the presence of multiple types of property word classes in many Oceanic languages: in addition to a large class of verby adjectives they often have one or more (very) small classes of property words with non-verby properties. More generally, the prevalence of verby property words fits into the broader pattern of lexical flexibility of Oceanic languages, in which especially the function of predication (and its associated categories of person and TAM marking) is very ‘hospitable’ towards non-event-denoting words, including property words.

References

- WALS chapters (Dryer, Matthew S. & M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Accessed on 2014-11-24):
- Dahl, Ö & V. Velupillai 2013a. The Past Tense (<http://wals.info/chapter/66A>).
 - Dahl, Ö & V. Velupillai 2013b. The Future Tense (<http://wals.info/chapter/67A>).
 - Nichols, J. & B. Bickel 2013. Locus of Marking in the Clause (<http://wals.info/chapter/23A>).
 - Stassen, L. 2013. Predicative Adjectives (<http://wals.info/chapter/118A>).
- Dixon, R.M.W. 2004. Adjective classes in typological perspective. In: R.M.W. Dixon & A.Y. Aikhenvald (eds.) *Adjective classes. A cross-linguistic typology*. Oxford: Oxford University Press. pp. 1-49.
- Ross, M. 1998. Proto-Oceanic adjectival categories and their morphosyntax. *Oceanic Linguistics* 37 (1), 85-119.
- Stassen, L. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Wetzer, H. 1996. *The typology of adjectival predication*. Berlin: Mouton de Gruyter.
- Van Lier, E. & J. Rijkhoff 2013. Flexible word classes in linguistic typology and grammatical theory. In J. Rijkhoff & E. van Lier (eds.) *Flexible word classes. Typological studies of underspecified parts of speech*. Oxford: Oxford University Press.

***‘Not-yet’-expressions in the languages of the world:
special negators or a separate cross-linguistic category***

In many languages of the world there is a special negation strategy to indicate that an action has not been accomplished or that a state has not been attained. This is illustrated by data from Indonesian in (1) below. In this language, the normal way to negate verbal predications as in (1a) is to use the particle *tiada* (or *tidak*). Nominal predications as in (1c.) are normally negated by the particle *bukan*. However, both verbal and nominal predications have to be negated by *belum* ‘not yet’ when the speaker intends to communicate that an action has not been carried out yet as in (1b) or a particular state/goal has not been reached yet as in (1d). The perfect marker *sudah* cannot be negated by *belum* or the standard negator *tidak*, cf. Sneddon (1996: 202).

The existence of unique expressions usually translated as ‘not yet’ is readily reported in grammars of individual languages as special negators that differ from the standard negator. They are also sporadically mentioned in the literature on negation cf. for instance (Payne 1985, Miestamo 2005) as well as in studies on adverbial constructions cf. van der Auwera (1998). The latter author identifies phrasal adverbials in the languages of Europe, commonly translated as ‘not-yet’, under the label *nondum* which shall be adopted here too. There is a tacit knowledge that unique *nondum* expressions are wide-spread in the languages of the world. However, a systematic study of their cross-linguistic distribution does not yet exist. Consequently, the first goal of this study is to gain a better understanding of their occurrence in the languages of the world. Other equally important goals are to get an insight about the semantics and the categorial status of expressions usually glossed as ‘not yet’.

In my current sample of 100 unrelated languages, *nondum* expressions occur in most areas of the world, but are notably absent in Europe in the form of single, bound or semi-bound, clearly grammaticalized negative temporal markers. My main sources have been descriptive grammars and parallel texts. Information about the semantics of *nondum* expression is not readily available from grammars, which is why work with texts becomes indispensable in this study. My current results allow for the following observations. *Nondum* expressions are typically encoded as affixes cf. (2) and (3) below or free particles; they tend to be unsegmentable morphological/lexical units. In terms of content, they typically indicate the non-occurrence of an otherwise expected action or state but also an anticipation about its imminent occurrence. Thus they appear to belong to both the temporal and the negative domain; however, as Contini-Morava (1989: 138) notes the negation they indicate is of limited duration. Given their cross-linguistic frequency as well as their functional and structural similarities in a number of unrelated languages, I would like to suggest that they belong to both the temporal and the negative domain and are probably best described as a category of their own, the *nondum*-category which covers both non-completion of an action or state and an expectation for that the action or state in question will materialize soon.

EXAMPLES

(1) Indonesian (Austronesian, Malayo-Polynesian) (Sneddon 1996: 196, 202)

- | | |
|---|--|
| a. <i>aku tiada berkata begitu</i>
I NEG say like that
‘I did not say that’ | b. <i>mereka belum berangkat</i>
they not yet leave
‘They haven’t left yet’ |
| c. <i>dia bukan guru</i>
3SG NEG teacher
‘she isn’t a teacher’ | d. <i>dia belum professor</i>
3SG not.yet professor
‘He isn’t a professor yet’ |

(2) Swahili (Atlantic-Congo, Narrow Bantu, East Bantu) (Miestamo 2005: 126)

- ha-wa-ja-som-a*
NEG-they-NEG.PERF-read-FV
‘They have not (yet) read.’

(3) Shor (Turkic, South Siberian), (Nasilov 2005: 229)

Aŋčī kel-gelek

Hunter come-not.yet

‘The hunter has not come yet’

REFERENCES

- Contini-Morava, Ellen. 1989. *Discourse Pragmatics and Semantic Categorization: The Case of Negation and Tense-aspect with Special Reference to Swahili*. Berlin: Mouton de Gruyter.
- Miestamo, Matti. 2005. *Standard Negation: The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin: Mouton de Gruyter.
- Nasilov, Dmitrij M. 2005. Notes on -QALAQ- in Shor. *Turkic Languages* 9:229-233.
- Payne, John R. 1985. Negation. In *Language Typology and Syntactic Description, vol I: Clause Structure*, ed. Timothy Shopen, 197-242. Cambridge: Cambridge University Press.
- Sneddon, James Neil. 1996. *Indonesian: A Comprehensive Grammar*: [Routledge Grammars. London: Routledge.
- van der Auwera, Johan. 1998. Phrasal adverbial in the languages of Europe. In Johan van der Auwera (ed.) in collaboration with Dónall P. Ó Baoill *Adverbial Constructions in the Languages of Europe*, 25-145. Berlin/New York: Mouton de Gruyter.

Toward a typology of sociatives

‘Together’-expressions or sociatives have been discussed in the typological literature mainly in the context of comitatives and reciprocals. Stolz et al. (2006: 367) are interested in ‘together’-expressions as reinforcements of comitatives (*together with*) which may serve for disambiguating comitative and instrumental roles. Nedjalkov (2007:18) makes a strict distinction between (i) sociatives (joint action as in *to eat together*) and (ii) spatial reciprocals (bringing into contact two or more objects as in *to tie the ends together*). I will argue here that this formal identity in English *together* and many other ‘together’ expressions is not an accidental homonymy and that *to tie/link/join/gather/come/call together* are sociatives rather than reciprocals. I will use the term resultative sociatives for these, since, unlike stationary sociatives such as *to eat/work/live/be together*, jointness applies only as a result of the event.

The major purpose of this paper is to argue that sociatives deserve to be studied as a typological phenomenon in their own right, and not only because they entertain strong connections to comitatives and reciprocals. Major topics of interest in sociatives include:

- Sociatives are semantically and formally complex. In many languages they can be encoded in highly different ways and there is a large array of varying shades of meaning.
- It is difficult to identify a semantic core domain in sociatives. Among other things, it is highly doubtful whether ‘joint action’ is the semantic core.
- There are widespread homonymy patterns in sociative expressions which in some respects look as if they were polysemy (they are cross-linguistically recurrent). Sociatives can contribute to a better understanding of the relationship between polysemy and homonymy.
- Sociative expressions entertain diverse relationships with different parts of speech. They can behave like nouns, verbs, adjectives, participles, adverbs, numerals, pronouns, and derivational affixes, which testifies to their diverse and complex lexical sources.
- Sociatives can be maximally ungrammaticalized and unlexicalized which is a challenge for assessing how universal they are. In some languages, however, sociative expressions have attained the frequency of highly grammaticalized elements, but sometimes without obligatorification.

Sociatives are often exemplified in reference grammars, but usually without being recognized as a typological phenomenon of their own. However, due to the high domain internal diversity of sociatives, it is not advisable to rely on sporadic examples in reference grammars alone. This is why this study mainly rests on data from a massively parallel text. Sociatives are investigated in translations of the New Testament from more than 200 languages from all continents. Parallel texts have the advantage that data collection is facilitated by algorithmic procedures. Since data can be compared cross-linguistically on the level of individual examples, semantic space can be modelled with a high degree of granularity. A major result of the study is that resultative and stationary sociatives are extreme poles of a scale which at the same time singles out different lexical sources, local vs. temporal jointness, and construction type (predicative vs. depictive).

References

- Nedjalkov, Vladimir P. 2007. Overview of the research. Definitions of terms, framework, and related issues. In Nedjalkov, V. P. (ed.) *Reciprocal Constructions 1*: 3-114. Amsterdam Benjamins.
- Stolz, Thomas & Stroh, Cornelia & Urdze, Aina. 2006. *On Comitatives and Related Categories: a typological study with special focus on the languages of Europe*. Berlin: Mouton de Gruyter.

Differential A Marking: diachronic developments and restrictions from a typological perspective

The present paper aims to investigate the diachrony of Differential A Marking (DAM), which here refers to a variation in the case-marking of the more agent-like argument of bivalent and trivalent constructions. (I follow Bickel's 2010 approach to grammatical relations.) DAM can be conditioned by the following factors:

- Referential properties of the A argument, e.g. person, number, animacy, focality
- Valency classes (defined by shared case frames)
- Clause properties: TAM values, polarity, clause type (main vs. different types of subordinate clauses), scenario (properties of co-arguments)

DAM is often restricted to specific environments; at the same time, however, many languages have more than one factor conditioning DAM and exhibit complex interaction patterns of these factors. For instance, in languages having different valency classes, further splits are often restricted to only one valency class (usually the one involving prototypical transitive constructions). Moreover, splits conditioned by information structure often only occur in a subset of A arguments (i.e. A arguments with specific referential properties) and/or only within certain TAM categories.

However, DAM cannot be accounted for in terms of universal tendencies (cf. Bickel & Witzlack-Makarevich 2008, Bickel et al. 2014), and it is often difficult to find direct synchronic functional explanations for individual instances of DAM. A diachronic investigation of the patterns, on the other hand, has the advantage of being able to account for rare idiosyncratic cases. The diachronic emergence of DAM involves, among others, the following contexts and mechanisms:

- Reanalysis of nominalized clauses as full (main) clauses
- Emergence or extension of valency classes or voice types
- Reanalysis and extension of other clausal dependents (arguments or adjuncts)
- Reanalysis of focus markers as A markers
- Morphological properties of a subset of A arguments (most prominently suppletive pronoun forms)

The present study investigates the mechanisms and contexts that give rise to DAM from a cross-linguistic typological perspective and thereby seeks to account for language-specific patterns of DAM and their restrictions.

Examples are drawn from a worldwide sample of languages. However, languages of Australia, New Guinea, the Himalayas and the Caucasus feature more prominently, since DAM patterns are found more frequently in the languages of these areas.

References

- Bickel, Balthasar. 2010. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 399–444.
- Bickel, Balthasar & Alena Witzlack-Makarevich. 2008. Referential scales and case alignment: reviewing the typological evidence. In Marc Richards & Andrej L. Malchukov (eds.), *Scales*. Linguistische Arbeitsberichte 86, Universität Leipzig, 1–37.
- Bickel, Balthasar, Alena Witzlack-Makarevich & Taras Zakharko. 2014. Typological evidence against universal effects of referential scales on case alignment. In Ina Bornkessel-Schlesewsky, Andrej Malchukov & Marc Richards (eds.), *Scales and hierarchies: a cross-disciplinary perspective*. Berlin: de Gruyter Mouton, 7–44.

Worldwide patterns of language migration

Søren Wichmann

Max Planck Institute for Evolutionary Anthropology

The paper builds on three fundamental achievements of the project known as the Automated Similarity Judgment Program: a method for dating language group divergence (Holman et al. 2011), a method for inferring homelands of groups of related languages (Wichmann et al. 2010), and a large lexical database of close to two thirds of the world's languages (<http://asjp.cld.org/>) providing the data through which dates and homelands are inferred. By combining the 'whens' and the 'wheres', distances in time and space separating a given proto-language (e.g. proto-Slavic) from a descendant (e.g. proto-East Slavic) can be computed, and, thanks to the large size of the database, thousands of migration events (where proto-Slavic → proto-East Slavic would represent one such event) can be mapped.

Although still unpublished, this method of inferring language migrations has been applied and tested many times since I first presented it in the Fall of 2010, resulting in a range of empirical insights on language migration rates during different periods of world prehistory, both generally across language families and also as regard specific cases, such as the Bantu expansion. For instance, for most of human prehistory the average language migration rate remains at a few hundred meters per year but then increases after 2000 BP—but only for agriculturalists, not for hunter-gatherers. There is a tendency to move more along the east-west than the north-south axis, a tendency which is somewhat greater for agriculturalists than hunter-gatherers. The speed of expansion is similar in different natural environments, with a few exceptions: in (1) deserts and xeric shrublands, (2) temperate grasslands, savannas & shrublands, and (3) oceans the speed of migration seems to increase in comparison to all other environments.

This paper will present the conclusive study on language migration rates based on a refined approach to the measurement of geographical distances. The robustness of the procedure for determining homelands will be tested through simulations. Finally, comparisons will be made to see to which degree results are affected when different language classifications are used to define the proto-language pairs that enter into in the migration events.

References

- Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, and Dmitry Egorov. 2011. Automated dating of the world's language families based on lexical similarity. *Current Anthropology* 52.6: 841-875.
- Wichmann, Søren, André Müller, and Viveka Velupillai. 2010. Homelands of the world's language families: A quantitative approach. *Diachronica* 27.2: 247-276.

Diversity Linguistics: Retrospect and Prospect
The Max Planck Institute for Evolutionary Anthropology
Leipzig, Germany
March 1-3, 2015

Robert S. Williams
The American University in Cairo

Abstract

The Uncunwee Documentation Project: Language Documentation in the Nuba Mountains of Sudan

In this paper I will provide an overview of the *Uncunwee Documentation Project* (UDP), a major documentation project working to document, analyze, and archive Uncunwee, an endangered and little documented East Sudanic language spoken in the Nuba Mountains of Sudan and in the Uncu Diaspora.

The UDP, funded through the Hans Rausing Endangered Languages Project, the American University in Cairo and supported by the Linguistic Department at MPI-EVA, involved an internationally-based collaborative team of seven linguists, situated at the American University in Cairo and the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. Team members included four postdoctoral linguists, one doctoral student linguist, and two language community linguists.

I will begin the talk with a short historical overview of areal documentary linguistics in the Nuba Mountains, beginning in modern times with the work of Junker and Czermak in the first quarter of the 20th century, and continuing into the present with the UDP and other ongoing documentation projects in the Nuba Mountains, including the DOBES-sponsored Tima documentation project.

With respect to the UDP, I will discuss a range of issues, including those involved with our preliminary work with Uncunwee-speaking refugees in Cairo and the challenges of working with the Uncu community in the Nuba Mountains. I will then summarize the findings and accomplishments of the project and finally speak to the prospect for further documentation work in this area.

Title: "A Typology of Language Naming Principles"

Author: Jan Wohlgemuth

Subfields: Etymology, Typology, Onomastics

Abstract:

Language names, be they endonyms or exonyms, show a variety of nomination principles, e.g. names going back to words meaning 'people', 'humans', or 'we', or involving shibboleths like the varieties of French called "langue d'oc" and "langue d'oïl". The talk will present a tentative typology of these nomination patterns as well as their areal distribution.

MPI.EVA Closing Conference on DIVERSITY LINGUISTICS: RETROSPECT AND PROSPECT
1-3 May 2015, Leipzig

Abstract

Vocalogenesis in Chadic Languages (West & Central Africa)

The about 190 Chadic languages which are spoken in the plains, mountain ranges, and riverine lowlands of the West and Central African Sahel zone, form the largest and a typologically quite diverse family of Afroasiatic. Among these the about 80 languages of the Central Chadic (aka Biu-Mandara) branch stand out in diversity and pose some unique challenges of analysis and description in terms of phonological typology and historical-comparative reconstructions. This is, among other complex features, particularly the case with regard to their vowel systems whose inventories range from between only one to six underlying/phonemic vowels of synchronic description and who display many more phonetic surface vowels.

Any attempt at phonological and lexical reconstructions on proto-language levels must start from these “underlying” vowels – based on decisions as to how many “underlying” vowels there are per language or language group (mostly between one and four). Such decisions must take into account the effects of both diachronic and synchronic palatalization and, possibly, labialization “prosodies” which may affect consonants, vowels, or both, giving rise to subtypes of *Consonant Prosody* languages, *Vowel Prosody* languages, and *Mixed Prosody* languages which show features of both *Consonant Prosody* languages and *Vowel Prosody* languages. For Central Chadic systemic reconstructions of proto-languages, the choice would appear to lie between one and three diachronic underlying vowels (or rather: vocoids?).

Modern Central Chadic languages, or even modern Chadic languages in general, can thus be assumed to have undergone “vocalogenesis” by creating a larger number of synchronic phonemic vowels from definitely lesser numbers of diachronic phonemic vowels (or vocoids?), involving approximants (corresponding to Semitic “weak radicals”) and “prosodic” colouring by palatalization and/or labialization features.

The paper uses data from selected Central Chadic languages to illustrate both the range of typological diversity within this branch of Chadic and diachronic processes of “vocalogenesis” for individual present-day languages which are spoken largely by minority groups mainly in Nigeria and Cameroon.

Are you a butterfly or a spider?

A study of animal metaphors on Minangkabau women as cultural models

Tessa Yuditha
Radboud University
Jakarta Field Station
Max Planck Institute for Evolutionary Anthropology

Abstract:

Why are humans sometimes conceptualized as animals? What kinds of animals and animal characteristics are mapped onto humans in metaphorical expressions? Further, are there any cultural models behind the animal metaphors in a particular language? A small number of animal metaphor studies have been done on a small number of languages, for example English, Spanish, Tagalog and Mandarin Chinese (see Rodríguez, 2009; Chen and Chen, 2011). By adopting the folk conception of the ‘Great Chain of Being’ where humans are placed at the top of the organization and other creatures are below, Rodríguez (2009) demonstrates how animal metaphors in English and Spanish are used to marginalize particular social groups relating to their specific behaviors. The aforementioned studies show that there are cognitive and cultural motivations behind the metaphorical conceptualizations. Amongst a variety of things, the same animals can be viewed differently in different cultures.

Minangkabau, an indigenous culture of Indonesia which is located in West Sumatra, is famous for its matriarchal system and rich literatures and tradition. The culture channels its traditional values through a variety of spoken and written genres, including *Pitaruah*, an oral poetic genre, which offers the listeners different kinds of advice. One of the *Pitaruah* talks about Minangkabau women in animal metaphors. There are five types of animals used as models, i.e. *barau-barau* ‘Bulbul bird’, *layang-layang* ‘swallow’, *lawah-lawah* ‘spiders’, *anai-anai* ‘flying termite’ and *ramo-ramo* ‘butterfly’. The animal characters are depicted and metaphorically mapped onto Minangkabau women’s characteristics to convey the traditional values of a good traditional Minangkabau woman. This paper will discuss the five animal metaphors relating to Minangkabau women by elucidating the relevant Minangkabau cultural models, such as politeness.

Keywords: animal metaphor, metaphorical conceptualization, women, cognition, cultural models, politeness.

References

Chen, S. F and Li-Chi Lee Chen. 2011. What Animals Reveal about Grammar and Culture: A Study of Animal Metaphors in Mandarin Chinese and English. Journal of National Taiwan Normal University.

Rodríguez, I. L. 2009. Of Women, Bitches, Chickens and Vixens: Animal Metaphors for Women in English and Spanish. Cultural Studies Journal of Universitat Jaume I: Language and Representation, Vol VII/2009, p. 77-100.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind.

Speech Act Realization Strategies in Persian, English, and Turkmen: A Case of Monolingual and Bilingual Learners

Amir Zand-Moghadam

Assistant Professor of Applied Linguistics, Allameh Tabataba'i University, Tehran, Iran

amir.zand.moghaadam@gmail.com

Aylar Adeh

MA Student of Applied Linguistics, Allameh Tabataba'i University, Tehran, Iran

aylar23@gmail.com

Abstract

Given the differences and similarities among languages regarding pragmatic strategies they employ, and due to the fact that similarities and differences between second and third language acquisition, particularly in terms pragmatics, have not been sufficiently investigated, the present study aimed at comparing and contrasting request and compliment strategies in three languages (namely, Persian, English, and Turkmen) and also among monolingual and bilingual learners of English. To this end, having determined the differences among the three languages in terms of request and compliment realization strategies, 59 Persian students learning English as a second language and 59 bilingual Turkmen-Persian students learning English as a third language were selected and given a WDCT. The collected data were then coded and analyzed following Blum-Kulka, House, and Kasper (1989) and Ferankia's (2013) models. Chi-square test for independence revealed that, with regard to the speech act of compliment, bilinguals generally favored "multiple head acts plus supportive moves", whereas monolinguals mostly used "single head acts only". Differences were also found with regard to the use of "supportive moves only": monolinguals used "supportive moves only" (other types of speech acts) more often than the other group. However, bilinguals preferred "head act plus supportive moves" in all situations. With regard to the speech act of request, it was revealed that bilinguals frequently used "hedging expressions" and "want statements". However, monolinguals clung to the strategy of "query preparatory" in the majority of cases. The findings of this study point to pragmatics differences among the three languages and also the advantage of bilinguals over monolinguals in terms of using appropriate speech act strategies in particular contexts. This study has implications for language researchers, teachers, and materials developers.

Keywords: pragmatics, speech act strategies, monolingualism, bilingualism

Transitivity harmony in Panoan and elsewhere

Roberto Zariquiey

The transitivity class of verbs in Panoan languages is encoded at multiple sites within the clause by means of varied processes associated with what Kittilä (2002: 67-78) calls indirect transitivity marking. Among those processes, we find so-called transitivity harmony constructions, in which two predicates in a multiple-predicate construction must match their transitivity with the each other in order to be combined. Transitivity harmony constructions are pervasive in Panoan languages, but they are also attested in languages of different families and different regions of the world.

This talk presents the first attempt to offer a cross-linguistic study of transitivity harmony constructions, based on both a detail study of Panoan data and database with 27 languages of other parts of the world. Six main parameters of cross-linguistic variation among transitivity harmony phenomena are explored: integration or clause union, transitivity matching strategy – harmony target, paradigmatic distinctions, obligatoriness and transitivity domain. It is argued here that some important correlations can be proposed among these parameters. Based on this study, I propose a number of generalizations about in which conditions transitivity harmony is more likely to develop.

References

- Hopper, Paul and Sandra Thompson (1980). ‘Transitivity in Grammar and Discourse’. *Language* 56: 251-299.
- Jacobsen, Wesley M. 1985. Morphosyntactic Transitivity and Semantic Markedness’. In W.H. Eilfort, P.D. Kroeber, and K.L. Peterson (eds.) *Papers from the Parasession on Causatives and Agentivity at the Twenty-First Regional Meeting of CLS*. Chicago, IL: Chicago Linguistic Society, 89-104.
- Kittilä, Seppo. 2002. *Transitivity: towards a comprehensive typology*. PhD dissertation. University of Urdu.
- Kittilä, Seppo. 2011. Transitivity typology. In Song, Jae Jung (ed.). *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 346-367.
- Valenzuela, Pilar M. 2003. *Transitivity in Shipibo-Konibo Grammar*. Ph.D. dissertation in linguistics, University of Oregon, Eugene.
- Valenzuela, Pilar. 2011. Multi-verb predicates and transitivity harmony in Shipibo-Konibo. In P. C. Muysken & A. Y. Aikhenvald (eds.). *Multi-verb Constructions: A view from the Americas*. Studies in the Indigenous Languages of the Americas, 185-212. Amsterdam: Brill.
- Zariquiey, Roberto. 2011. *A grammar of Kashibo-Kakataibo*. La Trobe University; Melbourne Australia.