

Abstracts

Association for Linguistic Typology
10th Biennial Conference (ALT 10)

2013 August 15-18, Leipzig

On the Universality of Auxiliary Verbs

Delimiting properties of auxiliary verbs vis à vis lexical verbs has been the topic of continuous debate in generative grammar. It has, as stated by Heine (1993: 26), “provided one of the most popular battlegrounds for disputes of linguistic theory.”

Although it has often been observed that there is no any specific language-independent formal definition that can be used to determine the characterization of any given element as an auxiliary verb (Anderson 2006: 5, Kuteva 2001:5, cf. Heine 1993: 70), the current paper argues that there is still room to find some universal properties that help us end up with the conclusion that auxiliaries and lexical verbs are two distinct types of syntactic entities. To this end, this paper describes carefully the characteristics necessary for what is to count as an auxiliary verb.

Having done that, the paper turns to argue that at least two universal properties must co-occur in order to distinguish the auxiliary verb from other syntactic categories. (i) Auxiliation should be understood as the development of constructions into markers of tense, agreement, modality, and perhaps aspect. (ii) Auxiliary verbs do not enter into a thematic relation with the arguments in the sentence, leaving this job to the lexical verbs that auxiliaries tend to occur separately from. This may constitute the standard syntactic argument that auxiliaries and lexical verbs are two distinct types of syntactic entities.

References

- Anderson, G. 2006. *Auxiliary Verb Constructions*. Oxford: Oxford University Press
- Heine, B. 1993. *Auxiliaries, Cognitive Forces, and Grammaticalization*. New York: Oxford University Press
- Kuteva, T. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press

Allocutivity vs anti-allocutivity: a closer look at non-argumental addressee- vs. speaker-encoding

This presentation deals with two symmetrical phenomena independently attested in a number of unrelated languages where they are known under different names but which can be described as the encoding of a non-argumental addressee viz. speaker in (most) main-clause predicates.

The former has been recognised in its own right and fully described as such (only) in Basque, where it is known under the term ‘allocutivity’ (Hualde and de Urbina, 2003, 242). Despite language-specific differences in use and degree of grammaticalization, careful examination of data from several unrelated languages shows that allocutivity phenomena share a number of morphological and syntactic traits (XXXX, under review).

There seems to be no label for the latter, i.e. the encoding of a non-argumental speaker, for which I would like to propose the term ‘anti-allocutivity’.

The aim of this presentation is to present data illustrating anti-allocutivity phenomena in the world’s languages, to propose a typology thereof along the lines of Fig. 1 (which exemplifies this in the case of allocutivity), and to analyse their (shared) crosslinguistic morphological and syntactic properties.

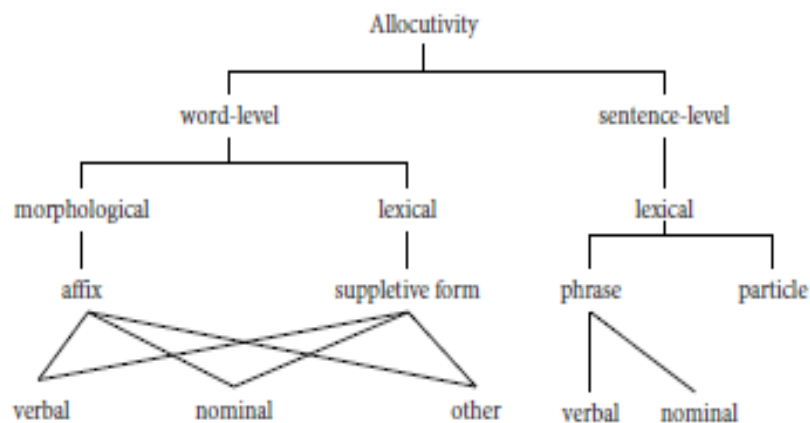


Figure 1: Types of allocutivity

Some preliminary findings include the fact that contrary to allocutivity, in the case of antiallocutivity morphological encoding on the verb by means of a grammaticalized affix is rare (possibly non-existent), whereas sentence-final particles (which in the case of an SOV language cliticise on the verb) seem to be frequent; the gender of the non-argument is not always encoded in the case of allocutivity but seems to be always encoded in the case of anti-allocutivity and there seems to be no language encoding both a non-argumental addressee and a non-argumental speaker.

References

Hualde, José Ignacio, and Jon Ortiz de Urbina (eds.) 2003. *A grammar of Basque*, Volume 9. Mouton de Gruyter: Berlin.

XXXX. under review. Verbal allocutivity in a crosslinguistic perspective.

Double marking of prominent objects: a cross-linguistic typology¹

Since the introduction of head- vs. dependent-marking as a typological parameter by Nichols (1986), not much attention has been paid to marking of arguments simultaneously by case-marking and verbal cross-referencing. Notable exceptions are Siewierska (1997) and Bakker & Siewierska (2009). The latter claim that “the likelihood of an argument exhibiting both overt agreement and case marking” declines according to the hierarchy A(gent) > P(atient) > R(ecipient). I argue that this claim must be qualified against data from a variety of genetically and areally unrelated languages where a well-defined and cross-linguistically recurrent type of non-A argument systematically exhibits double-marking.

As is well-known in Romance linguistics (e.g. Leonetti 2008), referential/specific objects (both ditransitive Rs and monotransitive Ps) in Spanish dialects and Romanian are both case-marked by an adposition and “doubled” by a cross-referencing clitic. Similar situations, when case-marking and cross-referencing of objects co-occur rather than exclude each other, can be observed in numerous languages, such as Amharic, Burushaski, Dera, Macedonian, Maithili, Mollala, Neo-Aramaic, Sentani, Thulung Rai, Usan, Yade.

I propose to classify these phenomena according to the following parameters: (i) what kind of non-A argument (P, R, or both) participate in double-marking; (ii) which factors determine double-marking (animacy, specificity, semantic role, or combinations thereof); (iii) is the same marker used both for P and R, as in e.g. Spanish, Maithili and Mollala, or these roles are distinguished in head-marking, dependent-marking, or both, as in Romanian and Burushaski; (iv) whether head- and dependent-marking can occur independently of each other (cf. Baker 2012 on Amharic).

The existence of strikingly similar patterns of double-marking of ditransitive recipients and animate/definite/specific monotransitive patients in a variety of unrelated languages suggests that this morphosyntactic pattern should be recognized as a cross-linguistic type of argument encoding. Moreover, it is obvious that double-marking of prominent objects is motivated by well-known universal functional preferences favouring overt case-marking of and overt agreement with animate/definite/thematic objects (cf. Givón 1976, Bossong 1985, Dalrymple & Nikolaeva 2011).

References

- Baker M.C. (2012). On the relationship of object agreement and accusative case: Evidence from Amharic. *Linguistic Inquiry* 43-2, 255–274.
- Bakker D. & Siewierska A. (2009). Case and alternative strategies: Word order and agreement marking. In: A. Malchukov & A. Spencer (eds.), *The Oxford Handbook of Case*. Oxford: OUP, 290–303.
- Bossong G. (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Dalrymple M. & I. Nikolaeva (2011). *Objects and Information Structure*. Cambridge: CUP.
- Givón T. (1976). Topic, pronoun, and grammatical agreement. In: Ch. Li (ed.), *Subject and Topic*. New York: Academic Press, 149–188.

¹ The work is supported by the Russian Foundation for the Humanities grant #11-04-00282a

Suffixing Preferences: Psycholinguistic Effects on Historical Change?

It has been known that suffixes are more common than prefixes cross-linguistically (Dryer & Haspelmath, 2011), and a number of accounts have been proposed, including psycholinguistic, historical, and formal accounts. Psycholinguistic accounts (Cutler et al. 1985, among others) are based on the idea that prefixes have some disadvantages over suffixes in language processing. There is, however, a weakness shared by psycholinguistic approaches: they only predict the general preference for suffixing and cannot account for the fact that the strength of suffixing preferences widely varies depending on grammatical categories. For example, while case marking shows strong suffixing preferences, person marking shows no evidence of suffixing preferences.

This study examines the possibility that a combination of a general psycholinguistic preference and historical origins can explain the current distributions of affixes. An illustrative example is discussed in Dryer (2011) for negation morphemes. In syntax, a negation morpheme more often precedes the verb than follows it, presumably because a negation morpheme that follows its scope causes a semantic garden path effect. In morphology, on the other hand, there are about the equal number of negation prefixes and suffixes. This can be readily explained if we assume that morphological negation markers come into being through the morphologization of syntactic negation words, but because of an independent psycholinguistic factor, preposed negation words are more often prevented from being morphologized. The historical and psycholinguistic factors cancel out each other and about the equal number of prefixes and suffixes results. This study is an attempt to pursue this approach in a more systematic way.

This study compared the typological frequencies of corresponding syntactic and morphological grammatical morphemes for each grammatical category, based on the literature on grammaticalization and the typological databases including Dryer and Haspelmath (2011). For example, it has been argued that gender markers typically evolve from demonstratives via the stage of definiteness markers (Greenberg, 1978). There are about the equal number of preposed and postposed demonstratives and definiteness markers; as expected, we observe suffixing preferences in gender markers in morphology.

Overall, our results confirmed that (i) there is a correlation between syntax and morphology, and that (ii) on top of that, the distribution is skewed towards postposing in morphology, as in Figure 1. Some grammatical categories, however, are outliers: object agreement markers are preposed more often than expected; case markers are postposed more often than expected.

We compare our results with accounts that do not resort to a general psycholinguistic preference such as Givón (1979), and discuss how predictions diverge.

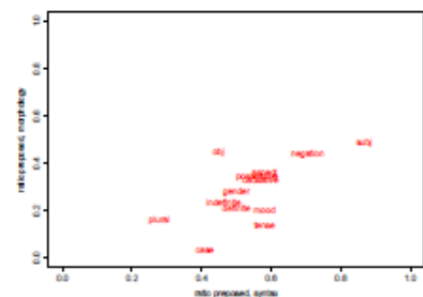


Figure 1: Ratio of preposed elements in syntax and morphology

References

- Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The suffixing preference: a processing explanation. *Linguistics*, 23, 723–758.
- Dryer, M. S. (2011). Order of negative morpheme and verb. In M. S. Dryer & M. Haspelmath (eds.)
- Dryer, M. S., & Haspelmath, M. (eds.) (2011). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. <http://wals.info/>
- Givón, T. (1979). *On Understanding Grammar*. Academic Press.
- Greenberg, J. (1978). How does a language acquire gender markers? In J. Greenberg (Ed.), *Universals of Human Language*. vol. iii: *Word Structure*. Stanford: Stanford University Press.

Suppletive kinship term paradigms in the languages of New Guinea

In some languages, the possessed forms of nouns, particularly kinship terms, may show suppletion according to possessor person. For example, in Awtuw (a Sepik language of New Guinea; Feldman 1968), ‘grandfather’ is split into two forms, *eywe* with 1st person possessors, and *yar* with 2nd and 3rd. Prior accounts of this phenomenon (e.g. Merlan 1982, Dahl & Koptjevskaja-Tamm 2001, Drossard 2004, Ortmann 2006, Vafaeian 2010) have ascribed it to the special relationship of 1st person possessors to kinship terms, leading inter alia to the use of a marked term (hypocoristic or term of address) by speakers making reference to their own kin, yielding an apparent 1st ~ non-1st suppletive alternation. However, these accounts have been based on a limited stock of examples; a fuller treatment reveals a considerably wider range of patterns.

The present study expands the existing typology on the basis on the languages of New Guinea, an area where suppletive kinship terms appear to be more common than elsewhere. Examples are drawn from 30 languages spread across 7 families/phyla (including 7 branches of the Trans-New Guinea phylum), and classified according to (i) the paradigmatic pattern of stem suppletion, and, where possible, (ii) the hierarchical relationship between forms (where relevant) -- which is the marked term and which the default term? On the first parameter, two stem alternation patterns predominate: 1st person ~ non-1st person -- by far the most common -- and 3rd person ~ non-3rd person. Within each of these there is evidence for either of the logically possible values of the second parameter. The combined typology is illustrated in (1), using the word ‘father’ in 4 languages of the Trans-New Guinea phylum. Telefol (Mountain Ok; Healey 1962) and Tainae (Angan; Carlson 1991) both display 1st ~ non-1st patterns. In Telefol the bare non-1st person stem (= 3SG form) is the default, while in Tainae it is the 1st person stem. In both cases evidence for the default status of a given stem comes from the fact that it can be used -- accompanied by possessive pronouns -- for ANY possessor person. Ekagi (Wissel Lakes; Drabbe 1952) and Usan (Madang; Reesink 1987) illustrate 3rd ~ non-3rd stem alternations. In Ekagi it is the non-3rd person stem *ajta* which is the default, while in Usan, it is the 3rd person stem which is the default. In both cases the evidence for default status comes from the clearly non-referential use of the terms in texts. Several languages also provide examples of 3-way person suppletion, but these (i) always occur as a minor pattern alongside one of the stem alternation patterns shown in (1), and (ii) there is good morphological evidence that they can all be interpreted as the concatenation of the two dominant stem alternation patterns.

(1)

	Telefol	Tainae	Ekagi	Usan
1SG	áatum	apo	1	ná-ajta-j
2SG	k-aaláb	tí-no	2	aká-ajta-j
3SG	aaláb	ka-no	3	na-kamee
stem alternation pattern:	1 st ~ non-1st	1 st ~ non-1st	3 rd ~ non-3rd	3 rd ~ non-3rd
default stem:	non-1 st	1st	non-3 rd	3rd

Thus these patterns discussed here cannot all be attributed to the special status of 1st person, *pace* previous observers. Rather, there are two binary paradigmatic splits (1st or 3rd person vs. everything else), where the markedness relationship is not fixed, but may vary across and even within languages. The resulting paradigmatic oppositions largely coincide with the patterns of subject person syncretism seen cross-linguistically in verbal inflection (Cysouw 2003, Baerman 2004), which is striking, given both the morphological differences (stem suppletion vs. largely affixal inflection) and the semantic differences (kinship relation vs. verb subject).

Linguistic Annotations and Knowledge Representation

In recent years linguists have become more interested in data-oriented research. They use corpora and they manage primary data. In their efforts they are helped by modern linguistic tools which promote standardisation and the use of metadata. In this way persistence and interoperability of primary linguistic data is gradually increasing.

As a timely initiative, Linked Open Data is of relevance for different kinds of linguistic online resources, including specialised encyclopedic knowledge banks such as the WALs, endangered-language archives, and federative linguistic online databases such as TypeCraft and the SSWL. Given new web technologies, it has become possible to search for embedded objects and in particular for classes and properties defined in ontologies. Applied to linguistics this means that indexes, in the form of linguistic glosses, become central as links between primary language data and more abstract linguistic knowledge.

Present attempts to make linguistic ontologies operable are still too weak. The online system, TypeCraft for example, provides URI-links between system tags and GOLD to allow the look-up of linguistic notions. Yet, in its present form the relation is not interactive and not informative enough to be useful for the linguistic glossing process. Although development within the Digital Humanities has made linguistic ontologies more framework independent and more comprehensive, ontologies are still not used to their full potential.

In our presentation we will discuss annotation and ontology integration, building on work by Chiarcos (2008). We will describe our own annotation model which consists of relations between morphemes, strings of tags (rather than individual ones) and tag classes, to suggest a design beyond the simple 1-1 mapping from tag to grammatical concept. We are particularly interested in the annotation of multi-lingual data from less-documented languages. We furthermore would like to reflect the incremental character of the linguistic annotation process (Mosel 2006a) by promoting a more dynamic integration of ontological knowledge.

In our presentation we would like to suggest a design which allows us to supplement subclassOf relations with disjointness and n-ary relations, as well as the flagging of certainty.

In order to discuss design questions on a fairly concrete level, we have acquired in-depth hand annotated data of 4 less-documented languages from typecraft.org. These languages feature between 56186 and 3079 annotated morphemes, and we will discuss this data in our presentation.

References

- Chiarcos, C. 2008. An ontology of linguistic annotations. LDV-Forum 2008 – Band 23 (1) – 1-16.
Farrar, S. and D. T. Langendoen. 2003. A linguistic ontology for the Semantic Web. GLOT International 7 (3), 97 - 100.

Numeral classifiers as markers of (in)definiteness – Conditions of grammaticalization and the existence of pre-established categories

The cross-linguistic homogeneity of grammatical categories is taken for granted by some linguists, while many typologists deny the existence of pre-established grammatical categories (Greenberg 1963, Haspelmath 2007). This paper will support the critical stance against cross-linguistic categorial homogeneity by showing that the function of categorial markers is determined by the specific conditions under which grammaticalization processes take place. To show this, it will discuss numeral classifiers as a source of (in)definiteness markers in some East and mainland Southeast Asian languages and it will look at the extent to which they can be compared to definiteness markers based on other sources (demonstratives, possessives, case in DOM).

Numeral classifiers are generally associated with the function of counting and with the lack of obligatory plural marking (Greenberg 1974). In this context, they individuate or atomize a concept. What is less well-known is that they also can express (in)definiteness if they occur in the [Classifier+Noun] construction. This is the case in various Sinitic languages (Cantonese, Wu Chinese, and marginally also in Mandarin Chinese) as well as in various Hmong-Mien languages, among them Hmong and Weining Ahmao. The paper will analyse the specifics of the (in)definiteness functions of classifiers in each of these languages and it will show that there are two factors that determine their (in)definiteness function:

- (i) Numeral classifiers refer to certain properties of the object they mark (e.g. \pm human, \pm one-dimensional, \pm book-like) and thus restrict the search domain of the hearer. The hearer will look for something that is one-dimensional if a classifier for one-dimensional objects is used.
Due to this identificational function, the classifier in [Classifier+N] marks familiarity rather than uniqueness.
- (ii) Each of the languages discussed is characterized by the omission of grammatical markers if they can be inferred from context.
As a consequence, the classifier is not obligatory even if it is a highly grammaticalized
- (in) definiteness marker. Once a referent is firmly established, it will simply be expressed by a bare noun.

These two properties are manifested as follows in individual classifier systems:

- (a) Sinitic: The interpretation of the classifier depends on word order as it is associated with information structure: In preverbal positions (topic), the classifier in [Classifier+Noun] is definite, while it tends to be indefinite in the postverbal position (focus). The definiteness expressed by the classifier is not based on uniqueness but rather on familiarity (on identifiability/familiarity and topic, cf. Lambrecht 1994). If the classifier occurs with a unique concept, it refers to its familiarity in the discourse situation.
- (b) Hmong: The classifier only marks definiteness in terms of familiarity irrespective of word order but its use is still driven by discourse and pragmatic inference.
- (c) Weining Ahmao has developed an inflectional paradigm for classifiers that combines singular/plural, definite/indefinite and size (augmentative, medial, diminutive). In spite of this, the classifier is not fully obligatory. It is obligatory only with foregrounded concepts, while backgrounded referents do not take a classifier.

Optional ergative case marking: what can be expressed by its absence?

Both the presence and the absence of information are sometimes equally useful for communication. A substantive example of this paradox is found in languages where otherwise obligatory grammatical information identifying core arguments (e.g. ergative case) may be 'optionally' absent without any consequences for the grammatical function of NPs in the clause. Far from being communicatively uninformative, the absence of ergative case marking has been linked to a range of different effects on the meaning of a clause within languages exhibiting this variability (McGregor 2009). These include focus alternations (e.g. Tounadre 1995), and the marking of modality (e.g. Hildebrandt 2004) and aspect (e.g. Li (2007)).

The existence of optional ergative case marking (OEM) raises important questions about our understanding of the role of case in language by (i) contesting the theoretical predominance of purely structural and lexically governed cases in mainstream linguistic theory and (ii) challenging preconceived ideas about the relationship between effability, obligatoriness and grammaticality. The factors that condition OEM cross-linguistically indicate that an adequate model of language must take into account subtle yet generalisable semantic and pragmatic conditions on the morphological form of core arguments (McGregor 2009). This clearly indicates that both morphosyntactic *features* (such as case) and *conditions* on those features (in the sense of Corbett 2006, 2012) play an important role in the distribution of case marking.

OEM is attested in many languages of the Himalayas, Australia and Papua New Guinea, yet little is currently known about possible variation in conditions on case-optionality across closely related languages in contact. This research reports on the results of a micro-typology of Indic, Tamangic and Tibetan languages spoken within the Tibetan Plateau Buffer Zone between the more typologically consistent Indospheric and Sinospheric Tibeto-Burman languages of the region (Matisoff 1991, Bickel and Nichols 2003, Hildebrandt 2007). Our approach, which uses data gathered using parallel elicitation and discourse collection methods, permits the exploration of linguistic variability through exploring the consistencies and subtle differences among the languages under investigation.

In this paper we discuss the factors that permit OEM for each language, including the types of features underlying splits in grammatical domains that permit OEM, and the language-specific pragmatic and structural conditions under which ergative case marking is absent. Specifically, we consider the roles that features and conditions play in establishing the distribution of ergative case and consider whether instances of 'optional ergative case' involve an ergative case feature.

Our approach aims to distinguish between (i) arguments that are consistently ergative (i.e. where the role of this case feature value is clear), (ii) arguments where the absence of ergative marking simply indicates the use of a morphologically unmarked case (such as absolutive, which is zero-marked in many of the languages in our survey), and (iii) arguments where the absence of ergative case marker indicates an alternation in the morphosemantic or information-structural properties of the clause, but not the grammatical function of the NP.

We demonstrate that while tense-aspect, focus and volitionality of the subject are clearly important factors in determining the splits in the marking of ergative case in some languages such as Nepali and Lhasa Tibetan, conditions on the distribution of ergative in other languages such as in Nar-Phu (Noonan 2003) and Manange (Hildebrandt 2004) are much less consistent. Rather than being predictable on the basis of a single condition or, indeed, being rigidly fixed, the evidence examined points to an analysis of OEM in which a multitude of conditions on case marking are employed to indicate a meaningful contrast.

On the crosslinguistic status of directional deixis

It is standard practice in reference grammars and dictionaries to use the glosses ‘come’ and ‘go’ in such loose fashion as to suggest some sort of crosslinguistic semantic/pragmatic “equivalence” between the markers at issue. Such congruence is also usually taken for granted by linguists working within the grammaticalization “paradigm”. However, several studies addressing specific languages or groups of languages (e.g. Ricca 1993; Wilkins & Hill 1995; Botne 2005) have shown how problematic it is to discount the actual range of discongruence from one language to the next: for instance, German *kommen* does not exhibit as much deictic strength as Spanish *venir* and Russian *prijti/prixodit’* even less so.

This paper argues that it is empirically justified, on balance, to bestow on directional deixis (DD) as a notional category much the same status as is classically granted to temporal deixis. This is because there is a clear tendency for markers implementing this putative category to display any number of a reasonably well-defined set of properties, most of which are attested across a broad spectrum of languages.

Putative DD systems are structured by a fundamental two-pronged asymmetry: (a) itive markers are deictically weaker than ventive markers, a semantic property with demonstrably transparent morphosyntactic correlates; (b) the intrinsically relevant variable is the deictic status of the goal of the motion event, not that of its source.

It is extremely common for DD markers to belong to closed paradigms, whether they be clitics (e.g. Somali) or affixes (e.g. German, Laal, Mohawk) or whether they consist in segmental or suprasegmental alternations (e.g. Kwaami, Pokot). When DD is lexically implemented, the verbs involved are almost invariably singled out by morphological idiosyncrasies (e.g. suppletion) or distinctive types of syntactic behaviour, proneness to serialization being only the best known of these. Obligatoriness of coding is another recurring property. Thus, failure to specify the deictic status of the goal may be disfavoured in varying degrees for some types of motion episodes, whether these are self-standing (e.g. Lisu, Palauan) or “associated” with an open-ended set of events (e.g. Australian and Chadic languages). Obligatoriness may also be instantiated indirectly: DD is not infrequently handled by portmanteau morphemes co-specifying the status of the motion event with respect to selected topographical coordinates (e.g. Sobei, Anggor) or co-encoding such prototypically grammatical categories as sentence force, tense/aspect, modality, person, diathesis or evidentiality (e.g. Iraqw, Barasano, Karajá).

Redundancy is a fairly common feature of DD systems (e.g. Yavapai, Nahuatl). It sometimes allows for the sort of syntagmatic redundancy typically associated with such functional categories as person, number or tense (e.g. Turkana, Dahalo, Comanche).

The susceptibility of putative DD exponents to grammaticalization processes is endemic across languages. Because they involve a break from the referential domain, pathways that lead from the encoding of DD to that of valency manipulation (e.g. passive voice in Italian and Scottish Gaelic; applicative voice in Krongo) are of special significance, especially those followed by markers “already” belonging to closed paradigms (e.g. Burushaski, Sochiapan Chinantec, Mosetén).

Theme session: Predicate-centered focus from a cross-linguistic perspective

Verb apocope as a marker of predicate backgrounding in Yulu (Central Sudanic, Chad/Sudan)

Provided that they are followed either by a pause or by the polar interrogative *nee* / *-ee*, some Yulu verb forms are articulated without their final vowel (non-intense ə). Contrast 1-2a and 1-2b below:

- | | |
|---|---|
| <p>(1a) mēsó àayā
 chief.DEF 3.come
 ‘The chief came.’</p> | <p>(1b) níinā k-āay
 who? FOC-come
 ‘Who came?’</p> |
| <p>(2a) à-lāayā nèe
 3.FUT-come Q
 ‘Will he come?’</p> | <p>(2b) à-cē làay-ēe
 3.FUT-NEG come-Q
 ‘Won’t he come?’</p> |

This apocope, which is then to be observed in very restricted contexts only (final position or before the polar interrogative), affects systematically – i.e. without any possible choice – the *subject-focalizing* forms (e.g. 1b), the *negative* forms (e.g. 2b), and, in dependent clauses, the *purposive* and the *consecutive* forms. It is clearly related to discourse hierarchy (or information structure) and indicates predicates that are backgrounded because of the assertion (or ‘focus’ in a wide sense) bearing either on another constituent (subject, negation) in the same clause, or on the predicate of the preceding main clause (in the case of purposive and consecutive).

As was pertinently pointed out by Hyman and Watters (1984), many African languages display two types of verbal inflections, one type consisting of usually shorter verb forms that are put ‘out of focus’ in the utterance hierarchy, may or may not be optional (‘pragmatic vs grammatical control’), and are frequently observed, be they free or compulsory, in the following contexts: constituent focalization (term focus), constituent questions (WHquestions), negation, imperative, relative clauses, and consecutive clauses. Furthermore, compatibilities of certain tenses/aspects with forms of either type are sometimes limited.

The aim of the present paper is to give a more detailed account of the situation of Yulu, in relation with similar phenomena in some languages of the same continent. Finally two different questions will be asked: 1. What is the real value of a function the marking of which is limited to the above mentioned contexts? and 2. Are such cases of interaction between verbal morphology and information structure attested outside Africa? If yes, where and how?

Reference

Hyman L.M. & J.R. Watters. 1984. *Auxiliary Focus*. Studies in African Linguistics 15/3, 233-273.

Encoding indefinite NPs in Austronesian languages: kind, specificity, free choice

Elaborating on Haspelmath's (1997) typology of indefinite pronouns, the focus will be on a typology of strategies expressing indefinite NPs in Austronesian languages.

Indefinites have to do with reference to kind and genericity, and with referential status in discourse. Strategies for indefinite reference generally differ along the parameter of specificity, sometimes along the parameter of reference to real world existence. Some languages encode reference to kind and to non-specific indefinite entities with bare nouns, others restrict bare nouns to kind and generics, and use distinct \pm specific indefinite articles for other types of indefinites. Non-assertable existential reference may be another parameter at play with a distinct paradigm of pronouns or determiners used for 'referentially' unknown, unexperienced entities with uncertain existence. In Nêlêmwa (Bril), bare nouns express kind, the specific indef. article is based on 'one', the non-specific indef. article is *xa* (also used for free choice), non-referential indef. are marked by suffix – *xo*. Generally, reference to kind, free choice and non-specific indefinites are distinct from specific indefinites. Non-specific indefinite articles generally correlate with T.A.M (irrealis, imperative, optative, conditional, etc.), or with interrogative, negative or negative existential clauses. While kind is often expressed by bare nouns, reference to kinds or sub-types of some kind, and to free choice is often marked by reduplication or plurality possibly with a non-specific article (Biak: *sup=o sup=o* (lit. land-non.sp land-non.sp) 'different/whatever places' Heuvel). As for indefinite pronouns ('somebody/thing' etc.), they often display a mixed type using (i) either ontological nouns (person, thing, etc.) together with existential construction, non-specific articles, or a classifier (Mokilese: *armaj-men* (lit. person-HUMAN.CL) 'someone', Harrison), or (ii) interrogative WH- pronouns (Amis: *cima a tamdaw* (lit. who LNK person) 'somebody'), possibly combined with disjunctive markers (Maori: *wai rānei* (lit. who or) 'someone' Bauer). Indefinite free choice (F.C.) pronouns and determiners ('any X', 'WH-ever (X)) have scope over a set of atoms/ variables that are expressed by plurality, or combination with universal quantifiers (all, every), distributive markers, reduplication, reduplicated WH- pronouns (Kupang Malay: *kekayaan apa-apa* (lit. riches RED-what) 'riches of any kind' Paauw); inclusiveadditive scalar morphemes may also combine ('too, even, etc.) (Amis: *maṭa-maṭan aca a munu* (lit. RED-what also LNK goods) 'any kind of goods'), as well as disjunctive markers (Tuvaluan: *me se aa te fakalvelave* (lit. or a what the problem) 'whatever problems' Besnier). Non-specific indefinite and F.C. forms only have a potential referent that often triggers epistemic (x *perhaps* y) or irrealis morphemes, possibly combining with WH- pronouns (Amis: *anu i cuwacuwa* (lit. if LOC RED.where) 'wherever, somewhere'). A typology of the strategies for indefinite NPs will be outlined.

References

- Bauer, W. 1997. *Reference grammar of Maori*. Auckland: Reed Publ.
- Besnier, N. 2000. *Tuvaluan. A Polynesian Language of the Central Pacific*. Routledge.
- Bril, I. 2002. *Le nêlêmwa (Nouvelle-Calédonie) : Analyse syntaxique et sémantique*. Paris: Peeters.
- Harrison, S. 1976. *Mokilese Reference Grammar*. Honolulu: The University of Hawaii Press.
- Haspelmath, M. 1997. *Indefinite pronouns*. Oxford University Press.
- Heuvel (van den), W. 2006. *Biak. Description of an Austronesian language from Papua*. Ph.D., Vrije Universiteit, Amsterdam.
- Paauw, S. H., 2009, *The Malay contact varieties of eastern Indonesia: A typological comparison*. Ph.D., State University of New-York at Buffalo.

Gender in Walman

The goal of this paper is to discuss how many genders there are in Walman, a language in the Torricelli family spoken in Papua New Guinea. This is not a straightforward question. There are two uncontroversial genders in Walman, masculine and feminine, and the majority of nouns belong to one of these two genders. But there are two other categories that might or might not be considered genders. One is diminutive, the other is pluralia tantum. In this paper we discuss possible reasons for considering or not considering each of these to be a gender.

The assignment of nouns to masculine and feminine follows a number of principles. First, all nouns denoting inanimate objects are feminine (except for the nouns for 'sun' and 'moon', which might traditionally have been considered animate by the Walman). Second, for humans and higher animals, gender is determined semantically. Only for lower animals are the factors determining the assignment of gender less clear. Masculine and feminine genders are distinguished in Walman only in the third person singular.

Walman also has an inflectional diminutive, which, like masculine and feminine gender, manifests itself not on the noun but on words agreeing with the noun, namely a number of nominal modifiers and both subject and object agreement on verbs. Corbett (2011) argues that the Walman diminutive is a gender (albeit an aberrant one), but the issue of whether it is a gender is partly if not entirely a matter of definition. If one defines gender as a feature associated with nouns as lexical items, such that some nouns are diminutive while other nouns are not, the Walman diminutive is not a gender: in principle any noun in Walman can be associated with diminutive agreement. For example, the noun *pirinyue* 'cockroach' is grammatically feminine and never masculine, but can optionally be associated with diminutive agreement, as in (1) below.

Another category which might be considered a gender in Walman is pluralia tantum. Unlike pluralia tantum in many languages, pluralia tantum in Walman do not bear plural inflection, but always trigger plural agreement, as in (2) below, where *chrikiel* 'net' is associated with plural agreement and can never be associated with feminine or masculine agreement. There are four ways in which pluralia tantum is like a gender in Walman. First, there are a large number of pluralia tantum nouns; in fact, they outnumber masculine nouns considerably. Second, apart from nouns denoting higher animals, which are associated with masculine or feminine agreement depending on their inherent gender, all nouns in Walman must belong to one of three classes of nouns: masculine, feminine, or pluralia tantum. Third, pluralia tantum noun phrases behave like singular masculine and feminine noun phrases in that they are optionally associated with diminutive agreement, something that is not possible with ordinary plural noun phrases (example (1) cannot mean 'I saw a number of small cockroaches'). And fourth, there is an irregular form *ngony* of the word *ngo* 'one' that occurs only with pluralia tantum nouns, as in (4) (where *tokun* 'knot' is pluralia tantum), something we might not expect if pluralia tantum nouns were simply grammatically plural.

- | | | | |
|-----|---|-----|--|
| (1) | Kum m-etere-l pirinyue.
1SG 1SG-see-3SG.DIM cockroach
'I saw a small cockroach.' | (2) | Chrikiel y-o lapo-y.
net 3PL-be big-PL
'The net is large.' |
| (3) | Chrikiel pa<l>ten l-o nyopu-l.
net that<DIM> 3SG.DIM-be good-DIM
'That tiny net is good (useful).' | (4) | N-optu-y tokun ngo-ny!
3SG.MASC-tie-3PL knot one-PL
'Tie one knot!' |

Middle distance agreement in adpositions: a typological niche

Agreement on adpositions is well-known but typologically uncommon, as indicated by Bakker's (2011) study of person marking. The familiar instances typically involve agreement in person (and number), and it is relatively easy to define the syntactic domain of agreement. Thus, in (1) the domain is a prepositional phrase (PP), in (2) it is an NP which has a whole PP as its dependent:

- (1) Welsh:
 llun [ohoni hi]
 photograph of.3SG.F she
 'photograph of her'

- (2) Hindi:
 [us strii kaa] bētaa
 that woman of.M.SG son
 'that woman's son' (McGregor 1995: 9; Spencer and Nikolaeva 2012: 210)

Agreeing adpositions of type (1) are observed in genetically and areally diverse languages, including Breton, Hebrew, Hindi, Savosavo (Papuan), Tehuelche (Chon), Turkish. Less is known about type 2.

We wish to draw attention to a third pattern, where the agreement expresses gender and number, and the controller is outside both the adpositional phrase and the NP but within its immediate clause. We call this phenomenon 'middle-distance agreement' by analogy with long-distance agreement, i.e. agreement outside the clause. The Daghestanian language Archi presents an example of this phenomenon:

- (3) goroχči b-aq'a ha'tar-če-q'a-k eq'en
 rolling.stone(III)[SG.ABS] III.SG-come.PFV river(IV)-SG.OBL-INTER-LAT <III.SG>up.to
 'The rolling stone went up to the river.'

In (3) *ebq'en* governs the lative and heads a phrase 'up to the river', an adjunct of the verb 'come', but agrees with the absolutive 'rolling stone'. The phrase *ha'tarčeq'ak ebq'en* forms a syntactic constituent: nothing can be inserted between the postposition and its governee, and the whole phrase can be fronted. But the controller is external to this constituent.

Other Daghestanian languages present a similar picture: in Dargi the postposition *salaw* governs the genitive but agrees with the absolutive *lbin*:

- (4) qalla sala-w kejž-ib-li ibin ü-di
 house(N).GEN before-M M.sit:PF-PRET-CVB lbin.ABS M.be-PST
 'lbin was sitting in front of the house.'

In Tsakhur the postposition *ab* agrees with the absolutive 'we':

- (5) ši wo-b-nī centr-ē a-b
 1PL.ABS be-HPL-PCL centre-IN inside-HPL
 'We were in the centre.'

Dargi and Tsakhur allow their agreeing postpositions to be used adverbially, i.e. without the governee. Similar behaviour is observed in other Daghestanian languages, such as Bagwalal, Godoberi and

Khvarshi. Adverbial agreement is much more common typologically, unlike middle-distance agreement. In Daghestanian languages, there are normally more agreeing adverbs than agreeing postpositions. Indeed, in Tsakhur the word *sana* 'together' does not agree in its postposition function, but does when being used adverbially.

Archi stands out in that the agreeing adposition does not allow the adverbial usage yet it shows middle-distance agreement and as such violates the typological expectation for the agreement target and controller to make a syntactic constituent. It is not, however, surprising to find this type of agreement in this language family. Daghestanian languages are also famous for long-distance agreement. As with LDA, middle-distance agreement is lexically defined (only some postpositions exhibit it), and they are grounded in the pervasive mechanism which requires agreement with the absolutive, which may explain its infrequency elsewhere.

The emergence of *zeonhang* as a progressive marker in Hong Kong written Chinese - and its typological comparison with Dutch *aan het*-construction

Chinese has long been known as a language having a relatively rich repository of aspect markers. In Hong Kong written Chinese (HKWC), *zoi* 'at' and *zingzoi* 'right at' are the markers that are most commonly used for expressing the progressive aspect. However, it is recently discovered that the lexical verb *zeonhang* '(be) in progress' is becoming more versatile and is increasingly used as a progressive marker, as shown in the following example.

GwokTaai honghung zingsat , ceotsi feigei zoi zeonhang
 Cathay.Pacific airlines confirm accident plane at- PROG

cenggit si, waattai dat bei kaidung .
 clean -time evacuation.slide suddenly PASS activate

'Cathay Pacific Airlines confirms that when the plane which had the accident was being cleaned, the evacuation slides were suddenly activated.'

(*Sharp Daily*, 05/10/2012)

From a synchronic perspective, this paper examines the degree of grammaticalization of *zeonhang* by identifying the perceived preferences and constraints of the marker by native users of HKWC. An acceptability judgment task, based on the model of Flecken (2011) on the grammaticalizing Dutch *aan het*-construction, was administered to 121 participants. Results show that *zeonhang* is anchored mainly in here-and-now contexts and is most compatible with dynamic predicates. Also, the thematic role of the subject is found to be the most determining variable for the adoption of *zeonhang*, in which subject as patient is highly favorable for the marker. This can be attributed to the semantic retention of the lexical origin of *zeonhang*, as well as the topic-prominent property of Chinese. For age-related differences, the middle group (20-30 years old) is discovered to be least likely to apply *zeonhang*.

Though Dutch and Chinese are languages which are typologically distant from each other, a comparison of the results between the Dutch study (Flecken, 2011) and the present one reveals similarities in their contexts for the grammaticalization of progressive markers: In terms of temporal contexts and situation types, the acceptability ranking of *aan het* and *zeonhang* are alike.

Consistent with the findings of previous studies (cf. Bybee, Perkins and Pagliuca, 1994; Comrie, 1976), this paper shows that there are some language-universal criteria for the development of progressive aspect markers in different languages, even if these languages are from distinct language families.

References

- Bybee, Joan, Perkins, Revere, & Pagliuca, William (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Comrie, Bernard (1976). *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Flecken, Monique (2011). What native speaker judgments tell us about the grammaticalization of a progressive aspectual marker in Dutch. *Linguistics*, 49(3), 479-524.

On the development of benefactive *lai* in Cantonese: Implications for the relationship between benefactive and purposive uses of directional verbs

Directional verbs are highly versatile elements that are susceptible to grammaticalization (Svorou 1994). Of these, the directional verb 'come' is observed to have grammaticalized into markers that can develop into tense-aspect markers, as well as purposive markers, among other functions, in various languages (Heine and Kuteva 2002). In Cantonese, the lexical verb *lai* 'come' has likewise grammaticalized into a purposive marker, and it can also convey past and/or emphatic meanings when used as a sentence final particle (SFP) (Cheung 1972; Leung 2005; Yiu 2001). In this paper, we will examine the relationship between purposive and benefactive uses of *lai* in Cantonese, with implications for other Chinese dialects and other languages.

The present study explores the grammaticalization of *lai* from a diachronic point of view using Cantonese data from the 17th to 20th century. These data are obtained from several corpora, among them the *Early Cantonese Tagged Database* and *A Linguistic Corpus of Mid-20th Century Hong Kong Cantonese Movies*. Our findings indicate that the directional verb *lai* underwent two major grammaticalization processes, one pathway yielding past and emphatic markers, and the other pathway yielding a purposive marker, with benefactive uses attested more recently in the 19th and early 20th century.

Regarding the first pathway, our data show that postverbal uses of *lai* as a co-verb meaning 'come' came to be increasingly used in sentence final position from the mid-19th century. This led to its reinterpretation within the temporal domain as a perfective marker, and within the pragmatic domain as an emphatic sentence final particle. The perfective marker was then further grammaticalized to allow it to carry past-tense meanings as well.

Regarding the second pathway, our data reveal that lexical verb *lai* was already used as a purposive marker during the Ming period; it was attested in Cantonese opera lyrics from the 17th century, and this purposive usage survives to present-day Cantonese. Interestingly, benefactive uses of *lai* were attested in 19th century Cantonese (e.g. Bridgman 1841), and this usage survived till the mid-20th century. Of theoretical interest here is the diachronic evidence which suggests that purposive and benefactive markers could emerge from directional verbs independently of each other.

In this paper, we will also compare the usage of *lai* in Cantonese with cognates in other Chinese varieties, including Mandarin, Jin, Hakka and other Cantonese varieties. While the aspectual and purposive functions of *lai* are also popular in other Chinese dialects, the benefactive use of *lai* was attested only in early modern Cantonese. We posit that structural variations (e.g. a strong tendency in Mandarin to prepose modifying elements to preverbal position, whereas Cantonese has a higher tolerance for postverbal elements) help explain this typological asymmetry.

References

- Bridgman, Elijah Coleman. 1841. *Chinese Chrestomathy in the Canton Dialect*. Macao: S. Wells Williams.
- Cheung, Hung-nin Samuel. 1972. *Cantonese as Spoken in Hong Kong*. Hong Kong: Chinese University Press.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of grammaticalization*. New York: Cambridge University Press.
- Leung, Chung-sum. 2005. *A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong*. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.
- Svorou, Soteria. 1994. *The Grammar of Space*. Amsterdam/Philadelphia: John Benjamins.
- Yiu, Yuk-man Carine. 2001. "Cantonese final particles 'LEI', 'ZYU' and 'LAA': An aspectual study." M.Phil thesis, The Hong Kong University of Science and Technology, Hong Kong.

Theme session: Typological hierarchies in synchrony and diachrony

Word Order Change and Hawkins's Prepositional Noun Modifier Hierarchy

This paper examines Hawkins's Prepositional Noun Modifier Hierarchy (PNMH; Hawkins 2004). The PNMH states that in prepositional languages, the longer the modifier, the less likely it is to be pre-nominal. This also leads to an implicational hierarchy of modifiers: if a prepositional language has pre-nominal possessive nouns, then it will also have prenominal adjectives; if it has pre-nominal relative clauses, then all the other modifiers will be pre-nominal. This can be summarized as: Dem > Adj > PossP > Rel (if a prepositional language preposes one of those modifiers, then it preposes the others further up the hierarchy). However, I argue that Hawkins's hierarchy is inaccurate when WALS data on word order is used from Dryer (2011). For example, it incorrectly predicts that prepositional languages with GenN order will have AdjN order (in fact of the 52 languages with prepositions and GenN order, 37 of them across nine different families have NAdj order); and that languages with AdjN, NDem order should have postpositions (whereas in fact 22 out of those 25 languages have prepositions).

I advocate an alternative hierarchy in this paper, the 'Head-Initial Hierarchy': NRel > VO, NAdj > NDem > AdpN, NGen > NProGen > VS. For example if a language has noun-demonstrative word order then it is likely to have noun-adjective word order; and if it has noun-genitive order then it is likely to have noun-demonstrative order. Data from WALS to support these and the other elements of the implicational hierarchy will be given, and argued to hold more strongly than Hawkins's hierarchy (e.g. there are 475 languages with NDem, NAdj compared with 25 with NDem, AdjN; and 282 with NGen, NDem compared with 78 with NGen, DemN; Dryer 2011 Chapters 85, 86, and 87).

While Hawkins (2004) argues that the PNMH reflects the nature of processing (Hawkins 2004), I argue that the Head-Initial Hierarchy reflects two common historical situations: i) word orders changing at different rates, especially in situations of language contact; Greenberg (1969) showed that relative clause-noun orderings, verb-object and adjectivenoun orderings are among the first to change in situations of language contact before noun-demonstrative, noun-adposition and noun-genitive orderings (and these former orderings may be particularly susceptible to syntactic transfer in bilingual acquisition, e.g. Yip and Matthews 2000). ii) SOV languages often acquire SVO word order and other head-initial word orders, much more commonly than the other way around (e.g. Gell-Mann and Ruhlen 2011). Many modern SVO languages come from families which were SOV, and have retained more conservative head-final orderings such as GenN, making the ordering GenN, VO relatively common (122 languages); while the rare type NGen, OV (32 languages) and other violations of the Head-Initial Hierarchy are primarily found in the less common situation of families which were VO becoming OV (e.g. Tigre in the Ethiopian Semitic family, Kairiru and Manam in the Oceanic languages of PNG; Dryer 2011 Chapters 83, 86). These two historical tendencies taken together result in languages tending to have degrees of head-initiality along the implicational hierarchy given. This hierarchy is thus argued here to emerge from directionality of word order change and stability of word orders in language contact, rather than from processing principles.

References

- Dryer, M. 2011. Chapters 82, 83, 85, 86, 87, 88, 90. In Dryer, M. and Haspelmath, M. (eds) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library.
- Gell-Mann, M., & Ruhlen, M. 2011. *The origin and evolution of word order*. PNAS Early Edition, 2011.
- Greenberg, J. 1969. Some methods of dynamic comparison in linguistics. *Substance and structure of language*, 147-203. J. Puhvel. Berkeley: University of California Press.
- Hawkins, J. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Yip, V. and Matthews, S. 2000. Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition* 3 (3), 2000:193-208.

Theme session: Generalized Noun Modifying Clause Constructions

GNMCCs in Bezhta and Hinuq (Nakh-Dagestanian)

A generalized noun modifying clause construction (GNMCC) is a single construction consisting of a head noun and a modifying clause that covers the range of English relative clauses, fact-S constructions, and others. Current interest in the topic was spurred by Matsumoto's (1997)'s treatment of GNMCCs in Japanese, with the three types of English translation equivalents illustrated in examples (1)–(3).

- (1) [gakusei ga kat-ta] hon
student NOM buy-PST book
'the book that the student bought'
- (2) [gakusei ga hon o kat-ta] zizitu
student NOM book ACC buy-PST fact
'the fact that the student bought the book'
- (3) [sakana o yak-u] nioi
fish ACC grill-PRS smell
'the smell of (someone) grilling fish'

The possibility of GNMCCs in Nakh-Daghestanian languages has been noted in earlier work, for instance Daniel and Lander (2010), but this earlier work has restricted itself to presentation of a small number of examples without attempting to assess the full range of the construction. We present extensive data from two languages of the Tsezic branch of the Nakh-Daghestanian language family, Bezhta and Hinuq, in order to demonstrate that these languages show a range of GNMCCs almost comparable to that of Japanese. Illustrative examples from Hinuq are given in (4)–(6).

- (4) [ked-i r-u:-ho goła] xok'o-be
girl-ERG NHPL-make-IPVCVB be.PTCP khinkal-PL
'the khinkal (a kind of dumpling) that the girl made'
- (5) [uži: mecxe b-ik'ekko goła] xabar
boy.ERG money(III) III-steal.IPVCVB be.PTCP story
'the news (or story) that the boy stole the money'
- (6) [de mašina toł-o goła] mecxe
me.ERG car sell-IPVCVB be.PTCP money
'the money from my selling the car'

By working through the questionnaire developed within the Stanford University-based project "Noun-Modifying Constructions in Languages of Eurasia: Reshaping theoretical and geographical boundaries" led by Y. Matsumoto, B. Comrie, and P. Sells, we show that Bezhta and Hinuq do indeed show a variety of GNMCCs covering roughly the range found in Japanese, but with some notable restrictions. In particular, when the complement of a head noun expresses purpose, Bezhta and Hinuq do not allow a GNMCC, but instead require an infinitival complement, as in Hinuq example (7).

- (7) [de idu-do Ø-iλ-a] pikru
me home-DIR I-go-INF thought
'my intention to go home [lit. the thought for me to go home]'

We examine possible explanations for this discrepancy between Japanese and Bezhta/Hinuq, in particular the different nature of clausal complementation in the two sets of languages: Bezhta and Hinuq have well-developed clausal complementation constructions independent of GNMCCs, whereas in Japanese clausal complementation, even when governed by a verb, is largely parasitic on GNMCCs.

References

- Daniel, Michael and Yuri Lander. 2010. *A girl of word, meat of a ram, and a life of longing*. On peculiar cases of relativization in East Caucasian languages. Paper presented at the conference Syntax of the World's Languages IV, Lyon.
- Matsumoto, Yoshiko. 1997. *Noun-modifying constructions in Japanese: a frame semantics approach*. Amsterdam: Benjamins.

Lexical splits and “complete typologies”

A key notion in understanding language is ‘possible word’. While some words (lexemes) are internally homogeneous and externally consistent, others have splits in their internal structure (morphological paradigm) and inconsistencies in their external behavior (syntactic requirements). I first analyse the most straightforward lexemes, in order to establish a point in the theoretical space from which we can calibrate the real examples we find. We can then schematize the interesting phenomena which deviate from this idealization, including suppletion, deponency, syncretism and defectiveness. These phenomena have been centre stage for morphologists over the last decade. I now shift the perspective from the phenomena to the different resulting segments into which lexemes can be ‘split’, in respect to the phenomenon inducing the split. The key point is the dividing line between the two (or more) segments of the lexeme’s paradigm. I set out a typology of possible splits, along four dimensions: (i) splits based on the composition/feature signature of the paradigm versus those based solely on morphological form. Thus a Russian verb has two segments: one with a feature signature requiring person and number and one requiring number and gender. This type of split is to be contrasted with one where the feature signature is the same but the morphological form differs (as when one segment has a stem mutation and the other does not); (ii) motivated (following a boundary motivated from outside the paradigm, such as singular---plural) versus purely morphology--- internal (‘morphomic’); (iii) regular, extending across the lexicon, versus irregular (lexically specified); (iv) externally relevant versus irrelevant: we expect splits to be internal to the lexeme, but some have external relevance (they require different syntactic behaviours). I identify instances of these four dimensions separately: they are orthogonal, and therefore not dependent on each other. Their interaction gives a substantial typology, which can be insightfully represented as a Boolean lattice, with 16 possibilities. Drawing on a range of languages, including Archi, Georgian, Kayardild, Krongo and Sanskrit, I demonstrate that the typology is surprisingly complete. All the 16 possibilities specified by the typology are in fact attested. From the perspective of classical typology, this could be seen as a disappointing outcome: there is no unattested cell whose absence we should justify and attempt to explain. From a canonical perspective, the typology offered a set of possibilities (some of which appeared highly unlikely), and this set indicated the directions in which to look. In a sense, the typology provided the research programme rather than being the result. The fact that a “complete typology” was established is both surprising and significant. Furthermore, since the typology allows for the unexpected patterns of behavior to overlap in particular lexemes, it helps us to recognize some remarkable examples. Such instances show that the notion ‘possible word’ is more challenging than many typologists have realized.

Theme session: typological hierarchies in synchrony and diachrony

The Obligatory Coding Principle (alias Obligatory Case Parameter) in diachronic perspective

In the recent typological literature (Dixon 1994 and others), accusativity / ergativity is defined in terms of $S=A\neq P$ vs. $S=P\neq A$ alignment, but morphological accusativity / ergativity can be viewed as a particular case of a more general principle underlying the organization of verbal valency in languages that have consistent $S=A\neq P$ or $S=P\neq A$ alignment, the Obligatory Coding Principle. According to this principle, regardless of the number of arguments, the only available coding frames are, either (in 'accusative' languages) those including a term with coding properties identical to those of A, or (in 'ergative' languages) those including a term with coding properties identical to those of P. A formal elaboration of this principle can be found in the generative literature under the name of Obligatory Case Parameter.

I would like to present a paper discussing a particular type of diachronic process that may be responsible for the development of coding frames contradicting the Obligatory Coding Principle in languages with (quasi-)obligatory P-like coding: the univerbation of light verb compounds.

Some languages have a very high proportion of predicates expressed by means of light verb compounds whose non-verbal element is a noun encoded like the P argument of typical transitive verbs, and diachronically, there is a general tendency toward univerbation of light verb compounds. When the nominal element of the compound is coded like a patient, this process converts a formally transitive construction $A(X)pV$ (where lower case 'p' symbolizes the P-like coding of a word that does not represent a participant, and (X) refers to possible oblique terms representing additional participants) into $A(X)V$, i.e. a construction with a participant coded like A and no participant coded like P. In languages with obligatory A-like coding, this results in perfectly canonical constructions, whereas in languages with obligatory P-like coding, the same process automatically results in the emergence of constructions violating the Obligatory Coding Principle.

Two opposite tendencies can be observed among languages with obligatory P-like coding: either the verbs resulting from the univerbation of pV compounds tend to maintain the exceptional coding frame $A(X)V$, or they tend to regularize it. The first tendency is predominant in Basque. Basque makes a wide use of light verb compounds whose verbal element is *egin* 'do', and in many cases, the light verb compound is synonymous with a simplex verb whose root coincides with the non-verbal element of a docompound, as in *bultza(tu) / bultza egin* 'push', or *dirdira(tu) / dirdir egin* 'shine'. In most present-day Basque varieties, the predominant tendency is that the simplex verb assigns to its arguments a coding identical to that observed in the light verb construction.

The tendency toward regularization can be illustrated by Andic languages (Nakh-Daghestanian). For example, in Andic languages, the translational equivalent of 'listen to' is either a light verb construction whose etymological meaning is 'fix ear at', with the coding frame <ERG, ALL, abs>, or a simplex verb resulting from the univerbation of this compound. In some of the languages that have a simplex verb resulting from the univerbation of the compound 'fix ear at', this verb maintains the non-canonical coding frame <ERG, ALL>, but in some others, its coding frame has been regularized as <ABS, ALL>.

In my presentation, I would like to discuss a possible correlation with the distinction between strict and loose ergative coding (Harris 1985).

Theme session: 'Typological hierarchies in synchrony and diachrony'

The referential hierarchy: reviewing the evidence in diachronic perspective

The referential hierarchy, 1st person pronouns > 2nd person pronouns > 3rd person pronouns > kin > human > animate > inanimate, has been accounted for in terms of a variety of factors, such as animacy, topicality, definiteness and natural attention flow (Dixon 1979 and 1994, Comrie 1989, DeLancey 1981, Corbett 2000, Song 2001, Croft 2003, among others). These explanations have been proposed based on the synchronic association between individual factors and the presence of particular constructions, independently of the diachronic processes that give rise to these constructions in individual languages.

The paper discusses extensive cross-linguistic evidence about the possible diachronic origins of three major phenomena that have been described in terms of the referential hierarchy, namely alignment splits in case marking, hierarchical alignment, and the presence of singular vs. plural distinctions for different NP types. This evidence poses several challenges both for the explanations that have been proposed for the referential hierarchy on synchronic grounds, and for the very idea of a referential hierarchy, in the sense of a scalar alignment of particular NP types that is relevant for speakers and leads them to use different constructions for these NPs. In particular:

(i) The various constructions involved in alignment splits, hierarchical alignment, and the encoding of singular vs. plural distinctions arise as a result of processes of context-induced reinterpretation of particular source constructions (for example, the reinterpretation of various types of source elements as markers of particular argument roles or plural markers, and the reinterpretation of cislocatives and third person markers as inverse markers). These processes are based on highly specific contextual relationships between the meaning of the source construction and that of the resulting construction, rather than general factors pertaining to different NP types on the hierarchy such as animacy, topicality, definiteness, or natural attention flow.

(ii) The distributional patterns attested for individual constructions also do not appear to originate from these factors. Rather, they reflect the distribution of specific source constructions. When a construction is restricted to particular portions of the referential hierarchy (as is the case with some case or plural markers, and inverse markers), it originates from a construction that is restricted in a similar way. When the distribution of the source construction is unconstrained (as is the case with the constructions that give rise to other case or plural markers), so is the distribution of the resulting construction.

(iii) Different patterns pertaining to the same grammatical domain (for example, different alignment patterns or different types of restrictions in the distribution of singular vs. plural distinctions) originate from different diachronic processes, and the same holds for the various instances of individual patterns in different languages, for example the various instances of hierarchical alignment, or the various cases where a singular vs. plural distinction is limited to human or animate nouns. This suggests that, contrary to the traditional view, the patterns described by the referential hierarchy are not amenable to a unified explanation, and the hierarchy is best regarded as a schema that is general enough to capture the outputs of several independent diachronic processes.

Disentangling the variability of the perfect gram type

Based on Dahl (1985) and Bybee & Dahl (1989) we assume that there are universal “gram types” in the domain of tense and aspect. This paper focuses on one gram type: the perfect (e.g. the English Perfect as in *I have bought a car* and the Indonesian *sudah*). Like Dahl (1985) we assume that perfects exhibit similar distributions in parallel texts which is why a first step in compiling a sample of perfect grams is to extract them from parallel texts by means of collocation measures (here the New Testament is used).

However, at the same time as perfect grams are cross-linguistically similar, they also display considerable cross-linguistic variability which can be roughly ordered into several thematic groups: (i) semantic: perfects can be associated with “iamitive” (forms expressing ‘already’), hodiernal past, evidential, experiential, and resultative; (ii) constructional: synthetic vs. analytic exponence, single marker vs. distributed exponence, auxiliary-based, participlebased, adverb; (iii) combinatorial: special negation strategy, mutual exclusion with some kinds of subordinate clause (e.g., Swahili *-me-* cannot occur in relative clauses), pluperfect (combination with past); (iv) diachronic: there are different grammaticalization sources and the perfect grams are at different stages on the grammaticalization cline (partly reflected by their frequency); and (v) example gram similar: the “Euro”-Perfect is different from the “Sino”-Perfect and from the “Malayo”-perfect; etc.

We compile a database of perfect variability where each factor of variation is treated as a feature of its own. Feature values are either measured in parallel texts or determined manually with reference grammar data. This results in a data table of 100 languages from all continents and 20 features from all five thematic groups outlined above. Unlike in conventional typological databases the default assumption is that many features are correlated since they are all related by being associated with perfect grams. Accordingly we do not use a stratified sample (there is no point in sampling languages without a perfect), but even closely related languages are included since their small differences may be relevant for better understanding the variability. The sample is determined by the availability of the two different kinds of data sources used: electronic parallel texts (N.T.) and reference grammars. After compiling the database we apply a posteriori sampling methods.

Rather than comparing each pair of features individually, statistical methods are used to reduce the dimensionality of variability. The methods used are inspired by dialectometry and register analysis. Our results so far suggest that grams do not fall into neatly delineated subtypes but rather form clusters with graded membership. These clusters correlate to a certain extent with areal distribution and genealogical affiliation (which are no input features for the aggregation analysis).

References

- Bybee, Joan and Östen Dahl. 1989. The Creation of tense and aspect systems in the languages of the world. *Studies in Language* 13.51–103.
Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.

Theme Session: Quantitative Linguistic Typology

Structural stability across methods, language families and geographic areas

Typological features vary in terms of their stability (Wichmann & Holman, 2009; Dediu, 2011), as do meanings in the basic vocabulary (Pagel *et al.*, 2007) and biological genes (Woese *et al.*, 1990). This talk will address the patterning of these differences in stability across language families, methods of estimation, and geographical areas, using advanced quantitative methods.

First, the talk will show that despite the large number of conceptualizations and ways of estimating structural stability, there is a general agreement across very different quantitative methods in what features tend to be stable and which not (Dediu & Cysouw, *in press*). This empirical result is very important and suggests that despite the inherent complexities involved in defining structural stability, there exists a cross-method latent characteristic that ranks features on a stability scale.

Then the focus will move on a particular such method which uses modern Bayesian phylogenetics (Dediu, 2011; Dediu & Levinson, 2012) to estimate the stability of the structural features in WALs across as many language families as possible. To guard against methodological and data coding biases, I used two different software implementations (the off-the-shelf MrBayes 3 widely used in evolutionary biology, and the custom-written BayesLang), two different coding of the WALs data (the original polymorphic and a binary coding) and three genealogical classifications of languages (WALS, the Ethnologue, and Hammarström's 2010 "orthodox" families). This quantitative method found that (i) there are cross-family (universal) tendencies for some structural features to be more stable than others, but that (ii) there are important inter-family differences in their ranking of features' stabilities, and, surprisingly, (iii) that there is also large-scale among-families patterning of structural stability.

The first level (i) reinforces the idea that features are differently affected by the factors driving language change across families and areas in that some features tend to be universally more stable. The second level (ii) simply shows that language family-specific factors have an important role to play. The intermediate level (iii) seems to point to deep genealogical and areal relationships between families, such as the similarity between language families within the Americas, and between those spanning the Americas and North-East Eurasia, possibly reaching to a time-depth of at least 12,000 years ago.

Taken together, these findings suggest that structural stability can be measured across methods, and that it is patterned at three levels: universal tendencies, large-scale among family, and between individual families.

References

- Dediu, D. (2011). A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. *Proceedings of the Royal Society B* **278**:474-479.
- Dediu, D., & Cysouw, M. A. (*in press*). Some structural aspects of language are more stable than others: A comparison of seven methods. *PLoS One*.
- Dediu, D., & Levinson, S. C. (2012). Abstract profiles of structural stability point to universal tendencies, family-specific factors, and ancient connections between languages. *PLoS One*, **7**:e45198.
- Hammarström, Harald. (2010) A full-scale test of the language farming dispersal hypothesis. *Diachronica* **27**:197-213.
- Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, **449**:717–721.
- Wichmann, S. & Holman, E. (2009). Assessing temporal stability for linguistic typological features *LINCOM Europa:Munich*, <http://email.eva.mpg.de/wichmann/WichmannHolmanIniSubmit.pdf>.
- Woese, C.R., Kandler, O. & Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *PNAS* **87**:4576-4579.

Word order correlations in the domain of complex sentences: Syntactic processing and/or grammaticalization?

Word order correlations have been a central topic in linguistic typology for 50 years; but despite intensive research some central questions regarding the nature and analysis of word order correlations are still unresolved today (see Special Issue of *Linguistic Typology* 15, 2011). In this paper we analyze new data from a typological project on clause and constituent order in complex sentences in light of some central issues of the word order debate. Specifically, we investigate the correlation between the position of subordinate clauses and the position of the subordinator (i.e. complementizer, relativizer, subordinate conjunction). While this correlation is only indirectly related to the VO/OV typology, it plays an important role in the theoretical literature on word order universals (e.g. Hawkins 2004; Dryer 2009).

Earlier studies observed that the subordinator often occurs at the borderline between main and subordinate clauses (e.g. Grosu and Thompson 1977); but this has never been systematically investigated. Using data from a stratified sample of 106 languages, we examined the subordinate markers of pre- and postnominal relative clauses, pre- and postverbal complement clauses, and pre- and postposed adverbial clauses. In accordance with previous observations we found that postposed subordinate clauses are commonly marked by an initial subordinator, whereas preposed subordinate clauses typically include a final marker. The correlation is highly significant (Fisher exact $p < 0.01$); but deviant from the general trend, preposed adverbial clauses (notably *if* and *when* clauses) are also often marked by an initial subordinator and some postposed subordinate clauses include a final marker.

In a first step we analyze these data from the perspective of syntactic processing. Specifically, we show that Hawkins' processing theory of order and constituency accounts for nearly 80 percent of our data; but while this theory implies that processing shapes linguistic structure over time, it does not explain how this correlation may have evolved. In fact, as it stands the theory suggests that there is a set of predefined categories, i.e. subordinate clauses, main clauses, and subordinators, and then processing 'decides' as to how these elements are arranged. But this is not a plausible scenario.

In a second step we argue that our data can also be explained by grammaticalization (cf. Givón 1975). Specifically, we show that the position of the subordinator is determined by its position in the diachronic source. For instance, it is well known that the subordinate markers of (some) complement clauses are derived from quotative verbs. In the grammaticalization literature, this development is commonly described as reinforcement; but syntactically it involves the reanalysis of two recursively embedded subordinate clauses in which the higher level clause, i.e. the quotative clause, is reduced to a subordinator, i.e. a complementizer. Since quotative clauses (often) function as some kind of complement, they typically precede the main verb in OV languages and follow it in VO languages so that preverbal quotative verbs develop into final markers of preposed subordinate clauses, whereas postverbal quotative verbs are reanalyzed as initial complementizers of postposed subordinate clauses.

The analysis we propose is parallel to the one that has been proposed for other word order pairs, notably for the correlation between verb&object and auxiliary&verb (Bybee 1988). But unlike auxiliaries, subordinators evolve from a wide range of sources so that the effects of grammaticalization on word order are not immediately recognizable. Our analysis shows that the developments of subordinate markers are extremely complex and diverse; but there are some very common diachronic paths (for which we will provide quantitative data) that do not only account for the above mentioned correlation between the position of subordinate clauses and the position of the subordinator, but also for the 'deviant cases' that are unexplained by current processing theories (e.g.

postposed subordinate clauses with final subordinate markers). Challenging the syntactic processing account, we conclude that the phenomena investigated in this paper are more effectively explained by local diachronic processes than by global processing principles.

Theme session: Typological hierarchies in synchrony and diachrony

Implicational hierarchies and semantic typology: the case of ideophones

Two analytically separate aspects can be distinguished in the study of typological hierarchies: first, the discovery of structure by means of cross-linguistic comparison, and second, the explanation of structure in formal or functional terms, from synchronic or diachronic perspectives. The first aspect highlights *methods*: how do we design cross-linguistic comparison so as to facilitate the discovery of typological hierarchies? The second highlights *mechanisms*: what processes do we posit to account for these hierarchies (Cristofaro 2010)?

This paper considers these matters from the perspective of the emerging typology of ideophones: marked words that depict sensory imagery, found in many of the world's languages (Voeltz and Kilian-Hatz 2001; Dingemanse 2012). Methodologically, ideophones pose an interesting challenge because traditional linguistic practices largely fail to capture them (Blench 2010). How do we typologise ideophones, and what kinds of cross-linguistic data can be feasibly collected? I show how stimulus-based elicitation, sorting tasks, and collocational analysis can help, not just for ideophones but also for other linguistic phenomena not readily found described in grammars or dictionaries. In terms of mechanisms, ideophones are interesting because they all fall within one semantic domain (sensory imagery, broadly construed) and yet in different languages cover widely varying patches of this domain, at least judging from the data currently available. Is this variation patterned, and if so, by what mechanisms did it arise?

Various categorisations of ideophones have been proposed. Kilian-Hatz (1999) organises ideophones according to a hierarchy of the senses, and Akita (2009) ranks ideophone types based on the degree of lexical iconicity they exhibit. This paper unifies these perspectives by proposing an implicational hierarchy of ideophone systems and specifying the mechanisms that shape this hierarchy. At least three factors need to be distinguished: (1) the human sensory system, which sets limits on the possibility space; (2) the nature of sensory input, which provides for frequency effects; and (3) the semiotics of depicting sensory imagery in speech, which makes available different degrees of freedom for different types of form-meaning mappings. This results in the following hierarchy:

I. SOUND < MOVEMENT < VISUAL PATTERNS < OTHER SENSORY PERCEPTIONS < INNER FEELINGS AND COGNITIVE STATES

The hierarchy is implicational in the sense that having ideophones in a domain to the right implies having ideophones in all domains to the left of it. It is consistent with the cross-linguistic data currently available. As more data comes in (and here the methods outlined are crucial), it will likely be refined in two ways. First, it is probably possible to make more fine-grained distinctions; second, we may find that the hierarchy is not a simple linear sequence but a more multi-faceted semantic map. By highlighting the twin aspects of methods and mechanisms, this paper is not only a contribution to the typology of ideophone systems but also a contribution to the study of typological hierarchies and the factors that shape them.

References

Akita, Kimi. 2009. A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Japanese Mimetics. PhD dissertation, Kobe University.

- Blench, Roger. 2010. The sensory world: ideophones in Africa and elsewhere. In *Perception of the Invisible: Religion, Historical Semantics and the Role of Perceptive Verbs*, ed by. Anne Storch, 275–296. Cologne: Köppe.
- Cristofaro, Sonia. 2010. "Semantic maps and mental representation." *Linguistic Discovery* 8: 35–52.
- Dingemanse, Mark. 2012. "Advances in the cross-linguistic study of ideophones." *Language and Linguistics Compass* 6 (10): 654–672. doi:10.1002/lnc3.361.
- Kilian-Hatz, Christa. 1999. Ideophone: Eine typologische Untersuchung unter besonderer Berücksichtigung afrikanischer Sprachen. Habilitationsschrift, Universität zu Köln.
- Voeltz, F. K. Erhard, and Christa Kilian-Hatz, eds. 2001. *Ideophones*. Amsterdam: John Benjamins.

Form and function of a possibly universal interjection for initiating repair

Everyday language use displays rules and regularities well within the purview of typology. We illustrate this point with reference to other-initiated repair, an elaborate machinery for dealing with problems in speaking, hearing and understanding found in every natural language so far investigated (Schegloff, Jefferson, and Sacks 1977; Clark 1996). One cross-linguistically widespread strategy for initiating repair on a previous turn is the use of an interjection like “*huh?*”. Consider the following example from English. The interjection in line 2 initiates repair on the previous turn and elicits a repetition in the following turn.

- | | | |
|-----|---|--------------------|
| 1 | G | 'It's not too bad, |
| 2 → | E | Huh? |
| 3 | G | 'S not too bad, |

Even though repair mechanisms are fundamental to communication everywhere, research so far has focused on English and has not been comparative in scope. We report on a detailed investigation of the interjection strategy as it occurs in video corpora of informal everyday conversation in a diverse sample of 12 languages. Keeping sequential context constant, we examine over 200 tokens of the interjection for onset, pitch, and vowel quality. We combine phonetic measurements with a rating procedure to arrive at replicable judgements of the phonetic qualities of every single token. We find that the phonetic form of the interjection is strikingly similar across languages: a monosyllable with at most some glottal constriction at onset [h, ʔ], featuring an open non-high non-back vowel [a, æ, ε, ɜ, ə], often nasalized, and often produced with rising intonation. Typical tokens are [hɛʔ] in Dutch, [ʔ] in Chintang (Kiranti, Nepal), [ʔ] in Siwu (Kwa, Ghana), [aʔ] in Chaʔpalaa (Barbacoan, Ecuador), and [ʔ] in Lao (Tai, Laos).

We investigate several questions raised by the strong formal and functional similarities of this interjection across languages. Are there reasons to consider this a word at all or is it simply a pre-lexical grunt? Do all languages aim for the same generic form or do we find language-specific targets? We consider the design of the interjection from the perspective of the linguistic systems it interacts with and the interactional environment in which it is found, and conclude that both are key to its form and meaning. Traditionally, the two main reasons for cross-linguistic similarities have been thought to be shared inheritance or contact. In this study we propose another factor: common interactional environments and their potential to exert selective pressure towards convergent evolution. This factor is likely of far wider relevance in cross-linguistic typology than realised.

References

- Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A., Gail Jefferson, and Harvey Sacks. 1977. “The Preference for Self-Correction in the Organization of Repair in Conversation.” *Language* 53 (2): 361–382.

The typology of syllable structure

Phonotactic complexity has been claimed to correlate with geography (Maddieson 2011). This talk examines geographic and phonological correlates of phonotactic structure, based on a large and detailed database of phonological variables.

Currently the major large sample of cross-linguistic information on syllable types is WALS (Dryer & Haspelmath 2011). Only two maps, however, are dedicated to phonotactic questions, one representing 469 languages, and one displaying 486 languages; we use a database with 1600 languages. More importantly, WALS' description of syllable structure lacks granularity, with only three categories used to account for all variation in onset and coda possibilities: a) simple (CV only); b) moderately complex (CCV or CVC); and c) complex (freer than moderately complex).

More detailed divisions are not only possible, but desirable: the distribution of languages allowing (C)VC or (C)VCC syllables is different from that of languages allowing (C)VCCCC syllables, for instance. In particular, combining the two very different 'moderately complex' syllable types in b) collapses the distinction between constraints on the onset and constraints on the coda, which have different distributions.

We draw on a new compilation of sources, the World Phonotactics Database (<http://phonotactics.anu.edu.au>). Using this database we test Maddieson's claims about areal distribution, examine the stability of phonotactic constraints in families and areas, and explore the typological ramifications of syllable structure.

References

- Dryer, Matthew S. & Haspelmath, Martin (eds.). 2011. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Available online at <http://wals.info/>. Accessed on 2013-01-15.
- Maddieson, Ian. 2011. Syllable Structure. In: Dryer, Matthew S. & Haspelmath, Martin (eds). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 12. Available online at <http://wals.info/chapter/12>. Accessed on 2013-01-15.

On the order of demonstrative, numeral, adjective and noun

Cinque (2005) proposes an account of the different frequencies of orders of demonstrative, numeral and adjective with respect to each other and with respect to the noun among the languages of the world in terms of movements starting with a universal underlying order of Dem-Num-Adj-N, based on what is a possible movement and what types of movements are "marked" in his theoretical framework.

This paper presents data based on a different sample, consisting of 404 languages, and proposes an alternative account of the different frequencies, based on a set of surface principles, without appealing to movement or a universal underlying order.

Cinque does not present actual number of languages but characterizes the number of languages of each type informally in terms of expressions like "very many", "many", "few", "very few", and unattested, and his theory characterizes each type in terms of whether the type is unmarked, involves one marked option, two marked options, or is disallowed. The data presented here provide a number of empirical problems for his claims. First, four of the types that his theory disallows (Num-N-Dem-A, Dem-A-Num-N, Num-Dem-A-N, and N-Num-Dem-A) are attested in the present sample. Second, the third most frequent type (after the two most frequent types Dem-Num-Adj-N and its mirror image N-Adj-Num-Dem), namely Num-N-A-Dem, is a type that involves two marked options on Cinque's theory and should therefore be fairly rare on his theory. I argue that, apart from the two most frequent types and the unattested types, his theory does little better than chance in predicting the relative frequency of the other types.

The alternative account I propose involves the following six principles:

1. Iconicity Principle 1: The adjective tends to occur closer to the noun than the demonstrative when they occur on the same side of the noun.
2. Iconicity Principle 2: The adjective tends to occur closer to the noun than the numeral when they occur on the same side of the noun.
3. Iconicity Principle 3: The numeral tends to occur closer to the noun than the demonstrative when they occur on the same side of the noun.
4. Asymmetry Principle: The Iconicity Principles apply more strongly to prenominal modifiers than they do to postnominal modifiers; exceptions to the Iconicity Principles will occur only with postnominal modifiers.
5. Greenberg's Universal 18: When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.
6. Intra-Categorical Harmony: The demonstrative, numeral, and adjective tend to all occur on the same side of the noun.

I show that the relative frequency of the different types in my sample is predicted very well by how many of these six principles the language conforms to: (1) the two most frequent types satisfy all six principles; (2) the five next most frequent satisfy exactly five principles; and (3) only two types satisfying fewer than four principles are attested.

References

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315-333.

Motion, path, and topography in verbal constructions: An Amazonian perspective

The role of non-linguistic factors in shaping typological variation across languages has been only minimally explored, and tends to be notoriously difficult to pin down (see, e.g., Evans 2003, Enfield 2002, Evans & Wilkins 2000, Sapir 1949: 26, Hill 2006: 619). This paper probes the relevance of culture and environment in the domain of verbal lexicalization patterns associated with the encoding of motion, manner, and path, in light of Talmy's (1985, 1991, 2000) familiar typology of 'verb-framed' vs. 'satellite-framed' languages (in which path and motion are encoded in the verb and manner subordinated, or motion and manner are conflated and path expressed via a satellite, respectively). Prior work has pointed out that non-linguistic factors may play a role in the encoding of motion events (see Slobin 2000, 2004, *inter alia*), most notably involving environmental setting and local approaches to orientation and movement within it (Wilkins 2004, Fortescue & Lennert Olsen 1992:215); however, the extent to which these cultural and environmental factors are actually implicated in the cross-linguistic expression of motion events remains little understood.

This paper considers the resources for encoding motion events in languages of the northwest Amazon, paying special attention to the encoding of path, which is arguably particularly likely to reveal culturally and environmentally relevant characteristics. The discussion focuses in particular on Nadahup, Tukanoan, Kakua-Nukak, and Arawak languages of the Vaupés region, which share many common grammatical structures due to pervasive contact (e.g. Aikhenvald 2002, Epps 2007, Gomez-Imbert 1996). In these languages, motion and path are typically conflated in a single root (which combines with manner in serial verb constructions); most notably, however, the expression of path is *also* frequently conflated with information about topographic features of ground, usually relating to waterways – which are of clear cultural and environmental salience in the region (see examples 1-3 below). This same topographic emphasis on water is evident in these languages' resources for encoding spatial/directional information outside the verb; Hup in particular is known to have an extensive set of water-related postpositions (a subset is listed in 4). Moreover, a comparative look at expressions of motion events in languages of the larger northwest Amazonian region indicates that topographic detail, particular relating to water, is in fact a fairly widespread typological feature, even where lexicalization strategies themselves are variable. This discussion contributes to our understanding of the role that cultural and environmental factors may play in shaping typological distributions of linguistic features, and suggests that a further refinement of Talmy's typology of lexicalization patterns should involve the inclusion of a topographic parameter in the linguistic encoding of path.

- (1) Tariana (Arawak; Aikhenvald 1999:57)

<i>lama-pidana</i>	<i>disa</i>	<i>di-nu-pita</i>
burn-rem.pst.rep	3sg.nonfem+ go.upstream	3sg.nonfem-come-rep
'It went upstream, burning.'		

- (2) Hup (Epps fieldnotes)

tih	ʔət-kəd-cɔp-yiʔ-ay-áh
3SG	cry-pass-go.from.waterway-TEL-INCH-DECL
'She hurried crying from the river.'	

- (3) Kakua (Kakua-Nukak; Bolaños p.c. 2010)

<i>kăn</i>	<i>fâ-bũ</i>	<i>ʔa-men-hăh-beh-bip-na-ka</i>
3sg.msc	downriver-loc	3sg.msc-row-go.downriver-go-fut-imperf-ʔpres?
'He will row downriver.'		

- (4) hũyan 'submerged in water' (all other liquids: *g'od-an*)
 pótʔah 'upriver'
 mæ̀tʔah 'downriver'
 déʔah 'at waterway'
 wáʔah 'on other side of waterway'
 deh páʔah 'at edge of waterway'

References

- Aikhenvald, Alexandra. 1999. *Tariana Texts and Cultural Context*. München: Lincom Europa.
- Aikhenvald 2002. *Language Contact in Amazonia*. Oxford and New York: Oxford University Press.
- Enfield, Nicholas. 2002. *Ethnosyntax: Explorations in Grammar and Culture*. Oxford: Oxford University Press.
- Epps, Patience. 2007. The Vaupés melting pot: Tucanoan influence on Hup. In Alexandra Aikhenvald & R.M.W. Dixon (eds.), *Grammars in Contact: A Cross-linguistic Typology*, Explorations in Linguistic Typology 4, 267-289. Oxford: Oxford University Press.
- Evans, Nicholas. 2003. Context, culture, and structuration in the languages of Australia. *Annual Review of Anthropology* 32:13-40.
- Evans, Nicholas and David Wilkins. 2000. In the mind's ear: the semantic extensions of perception verbs in Australian languages. *Language* 76(3): 546-92.
- Fortescue, Michael & Lise Lennert Olsen. 1992. The Acquisition of West Greenlandic. in Dan Isaac Slobin (ed.), *The Crosslinguistic study of language acquisition, vol 3*, 111-221. Routledge Press.
- Gomez-Imbert, Elsa. 1996. When animals become 'rounded' and 'feminine': Conceptual categories and linguistic classification in a multilingual setting. In John Gumperz & Stephen Levinson (eds.), *Rethinking Linguistic Relativity*, 438-469. Cambridge: Cambridge University Press.
- Hill, Jane H. 2006. Writing culture in grammar in the Americanist tradition. In Ameka et al. 2006, 609-628.
- Sapir, Edward. 1949 [1933]. The psychological reality of phonemes. In D. G. Mandelbaum (ed.), *Selected Writings of Edward Sapir in Language, Culture, and Personality*. Berkeley: University of California Press.

Paradigms, diachronic typology, and language classification: Bayesian phylogenetics and pronoun paradigms

There is growing interest in the use of typology in historical linguistics. One particularly promising approach is to investigate relationships between tightly organised subsystems of language, as fractal-like proxies for overall relationships. One such case was the investigation of Northeast Caucasian languages by Cysouw & Forker (2009) who found close parallels between the phylogenetic classification given by a typological classification of the case systems and that given by traditional measures. However, the particular characteristics of Northeast Caucasian case systems prevent this particular application from being generalisable to most other language families. Therefore we focus on a linguistic subsystem – personal pronouns – which is found in all spoken languages.

This paper compares the architecture of attested pronominal systems. We develop a diachronic typology of the stepwise pathways between them, and then harness this to a phylogeny to infer the most parsimonious historical scenarios relating pronominal systems across a series of languages.

To get the right balance between informativeness and comparability, we focus on four (potentially) distinct forms of each personal pronoun: those for the three core grammatical relationships (A, S and O) and for the possessive pronoun (of course many languages will neutralise some of these distinctions). These four values then combine with the system of person/number combinations as well as other categories such as gender to give a paradigm. Within any pair of cells in the paradigm we code a number of relationships ranging from totally distinct (e.g. *we* and *our* in English), total syncretism (e.g. A=S in English *he*, S=O for Nen *bä* 'he/she'; sg=du=pl in English *you*) to formal overlap (e.g. *me* and *my* in English) to formal increment (e.g. Warlpiri ergative *ngajulurlu* from Warlpiri absolutive *ngaju(lu)*). Taken together, these factors then give a vast 'architectural design space' formed by the product of all featural paradigms (in terms of feature combinations) with all formal relationships between all cells within them.

Diachronic pathways between possible paradigms in the design space can then be modelled by assuming they are achieved by (a) adding or subtracting features or combinations thereof (e.g. neutralising inclusive/exclusive, or developing a distinct ergative form) (b) changing the formal relationship between any two cells (e.g. from formal increment to formal overlap, or from distinct to total syncretism). Once the set of possible pathways has been exhaustively elaborated, we can give the evolutionary distance between any two paradigms by calculating the number of transformational steps needed to get from one to another. We can then use a Bayesian phylogenetic methods to infer a set of phylogenies based on the most parsimonious set of changes across the whole population of paradigms.

The method will be tested against existing classifications of one well-known language family (Austronesian) to determine its reliability, and then trialled as a heuristic classification for Australian and Papuan languages (Trans-New Guinea and Morehead-Maró, plus selected outgroups).

References

- Cysouw, Michael & Diana Forker. 2009. Reconstruction of Morphosyntactic Function: Nonspatial Usage of Spatial Case Marking in Tsezic. *Language* 85(3), 588-617.

Multidimensional variation in person agreement: evidence from Alor-Pantar languages

In Siewierska's (1999) influential typology grammatical and anaphoric agreement are separated by an intermediate type, ambiguous agreement. Haspelmath (2012) uses different terms for these types (gramm indexes, cross indexes and pro-indexes) and emphasizes their *sui generis* nature. I use variation in one family, the Alor-Pantar family of about 20 non-Austronesian languages spoken in eastern Indonesia, to make two important points about person agreement. First, the different types represent related phenomena and should not be treated as unrelated or *sui generis*. Even the boundary between anaphoric and ambiguous agreement is not clearcut. Second, while the two extremes of Siewierska's typology serve as a good basis, the ambiguous type, which represents the majority of languages with person agreement, should be articulated further to cover the significant differences found in languages with it.

It is possible for some verbs to belong to the anaphoric type (Haspelmath's pro-indexes) and others to belong to the ambiguous type (Haspelmath's cross indexes). In Adang (West Alor) verbs typically either have a prefix obligatorily or they do not. This is arbitrary (Haan 2001), and what one might expect for ambiguous or grammatical agreement. But there is still a split according to person.

Repetition of the argument is impossible for all pronouns (except third plural) (1a)-(1b), but not for nouns (2a) and the third person plural pronoun (2b). Splits like this are familiar, of course (see Corbett 2006: 108).

- (1) Adang
- a. *nife na-muning*
our.mother 1SG-kiss
'Our mother kisses me.'
- b. **nife nari na-muning*
our.mother 1SG 1SG-kiss
'Our mother kisses me.'
- (2) Adang
- a. *na boi '-ah=am*
1SG pig 3-feed=PFV
'I fed the pigs.'
- b. *na supi '-ah=am*
1SG 3PL 3-feed=PFV
'I fed them.'

However, there is a small subset of Adang verbs (3a)-(3b) for which prefixation is not obligatory, so that the complementary alternation between affix and pronoun typical of pronominal agreement is possible.

- (3) Adang
- a. *in n-eh*
mosquito 1SG-bite
'The mosquito bites me.'

b. *in* *nari* *eh*
mosquito 1SG bite
'The mosquito bites me.'

The examples in (3) look like an anaphoric system, but clearly share the constraint on co-occurrence of prefix and pronoun found in the system illustrated in examples (1) and (2), which looks more like an ambiguous system.

Support for the view that the ambiguous agreement type needs further articulation can also be found in other Alor-Pantar languages. In Teiwa (Pantar) there are striking differences in the extent to which co-referential nominals are allowed. A corpus study revealed that prefixed verbs occurred 40% of the time with the third person plural free pronoun, and rarely with the first person exclusive plural, first person singular and third person singular. For the other person-number combinations, coreference is impossible.

There is a broad continuum of phenomena which one could label agreement. Within each language a variety of types can be observed, but these overlap in the properties they share. The Alor-Pantar languages show that typology must articulate the space between the two extremes.

Subject anaphors and agreement

Most linguists seem to agree that in reflexive and reciprocal constructions of accusative as well as of ergative languages the controller of the reflexive or reciprocal pronoun must be the most prominent argument (i.e. higher on a hierarchy of grammatical roles), whereas the pronoun itself must be a less prominent argument (Chomsky 1981, Dixon 1994). Similarly, it has been claimed that anaphors do not trigger agreement (Woolford 1999). However, Daghestanian languages provide fascinating data that violate such proposed universals regarding the syntactic alignment in reflexive and reciprocal constructions. For instance, in Icaric Dargwa there is a choice between the canonical reflexive construction for transitive verbs (1), whereby the controller is in the ergative case (agent), and the pronoun is in the absolutive case (patient), and the unusual pattern involving a ‘reversal of grammatical roles’ (2).

- (1) *murad-il cinna_ca-w w-aIX:-a=ca-w*
 Murad-ERG REFL-M[ABS] M-feed:IPFV-PROG=COP-M
 ‘Murad is earning his own living.’ (lit. ‘is feeding himself’)
- (2) *murad cinna_cinni w-aIX:-a=ca-w*
 Murad[ABS] REFL.ERG M-feed:IPFV-PROG=COP-M
 ‘Murad is earning his own living.’ or ‘As for Murad, he is earning his own living.’

In my talk, I will first provide a descriptive account of reflexive and reciprocal constructions in a number of Daghestanian languages by considering parameters that determine the ‘reversal of grammatical roles’: (i) the form of the pronouns (simple vs. different types of complex pronouns), (ii) the valency type of the predicate (canonical transitive, affective, extended intransitive), (iii) the grammatical role of the controllers and the pronouns (S, A, P, experiencer, stimulus, non-canonical agent, other), and occasionally (iv) word order. Most of the data has been gathered by the author during fieldwork in Daghestan.

In the second part I will examine previous analyses of the Daghestanian data and of quirky reflexive and reciprocal constructions in other languages (Anagnostopoulou & Everaert 1999, Amiridze 2003). I will show that Yamada’s (2004) proposal to analyze sentences similar to (2) as intransitive cannot be maintained. Instead, it is possible to analyze the ‘reversal of grammatical roles’ building on Ljutikova (1997). The difference between (1) and (2) must be explained through the historical development of the constructions and subtle pragmatic differences. In (2) the antecedent NP behaves similar to left-dislocated NPs (e.g. it bears the unmarked case) and only the reflexive is a true argument of the predicate ‘feed’ (see the translation). This is supported by the fact that in Icaric Dargwa simple reflexive pronouns are also used to establish coreference between clauses (Sumbatova & Mutalov 2003: 167-168).

The talk concludes by proving that anaphors in Daghestanian languages trigger verbal agreement in gender and number (1) and thus contradict the anaphor agreement effect, but the agreement is rather different from the person agreement in familiar European languages.

References

Amiridze, Nino. 2003. *The anaphor agreement effect and Georgian anaphors*. In Claire Beyssade, O. Bonami & P. C. F. Cabredo Hofherr (eds.), *Empirical issues in formal syntax and semantics* 4, 99–114. Paris: Presses de l’université Paris-Sorbonne.

- Anagnostopoulou, Elena & Martin Everaert. 1999. *Towards a more complete typology of anaphoric expressions*. *Linguistic Inquiry* 30. 97–119.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Ljutikova, Ekatarina A. 1997. *Refleksivy i èmfaza*. *Voprosy jazykoznanija* (6). 49–74.
- Sumbatova, Nina R. & Rasul O. Mutalov. 2003. *A grammar of Icari Dargwa*. München: Lincom Europa.
- Woolford, Ellen. 1999. *More on the anaphor agreement effect*. *Linguistic Inquiry* 30. 257–287.
- Yamada, Hisanari. 2004. *On the ergative reciprocal tsotsaz construction in Standard Avar*. Talk at the LENCA2 conference, University of Kazan.

When a language codes affectedness

Background: The features ‘affectedness’ and ‘affect’ are usually linked with grammatical relations and analyzed as semantic property correlated with the objects of transitive verbs (Van Valin 2005: 57) and as potential entailments of the subjects of unaccusative verbs (Perlmutter 1978, Dowty 1991, Levin and Rappaport Hovav 1994). Dixon 2010: 98 considers ‘Affect’ to be a property of certain verbs, e.g. ‘hit’, ‘burn’.

The question and the hypothesis: Mina has two means of marking the last clause of an episode. The question is what the choice of means codes. The hypothesis is that it codes a distinction between the affected subject and all other subjects.

The larger goal: The study demonstrates that the coding of affectedness is different from the interpretation of the properties of an argument as affected, as has been done in much of the literature so far. The analysis herein (based on fieldwork) is quite different from the one in Frajzyngier et al. 2005. The study shows that affectedness in Mina is not a feature shared by objects of transitive verbs and subjects of unaccusative verbs, as often claimed in contemporary literature, but rather is a feature marked by a specific formal means, *m* Verb-*yi*, which codes the following real-world characteristics: changes in the existential status of the subject (e.g. disappearance) (ex. 1); changes in the physical form of the subject (ex. 2); and displacement of the subject (ex. 3a-b). The function coded by the formal means *m* Verb-*yi* is called here ‘affectedness predication’. The demonstration of the existence of this type of predication is a contribution to the typology of semantic features that can be coded by the grammatical system of a language.

The argumentation: The affectedness predication can have only one argument, the subject, as evidenced by the use of subject pronouns (ex. 3b). The subject may be controlling or not (ex. 1-3), hence the feature [control] plays no role in the affectedness predication. The verb in the affectedness predication may be inherently intransitive or transitive, hence affectedness is not a correlate of transitivity (ex. 1-4) and for the same reason it is not a passive construction.

An entity that is part of a larger object undergoing an event in the real world cannot be the subject of the affectedness predication. Thus, the clause ‘he cut off the leg of goat’ (ex. 5a) cannot be followed by an affectedness predication with ‘leg’ as the subject (ex. 5b), but it may be followed by an affectedness predication with ‘goat’ as the subject, viz. ‘the goat is cut’ (ex. 5c).

Inherently transitive verbs that do not imply affectedness cannot occur in affectedness predications. These include the verbs of perception (see, hear, smell), cognitive verbs (know, think, forget), and even physical-contact verbs that do not inherently affect the form of the object (beat, touch). The affectedness predication is not stative, as stativity is a separate function coded by the construction *m* V1-V1.

Implications: The existence of the affectedness function in Mina correlates with the absence of the passive function in the language. The causal interpretation of this correlation may be that the feature [affectedness], in addition to its function with intransitive verbs, subsumes the prototypical semantic characteristics of the subjects of passive clauses in languages that have passive constructions. Mina has separate formal means to topicalize arguments.

(1) *ˈyám mə shibit-i nə lāk-wát zá*
 water REL disappear-AFF PREP pond EE
 ‘the water has disappeared in the pond’ (EE ‘end of event’)

(2) *mə rēb-i ká*
 REL bend-AFF CONC
 ‘he/it is bent’ (CONC ‘speaker’s concern’)

- (3a) *mə* *ɣim* *ɣim* *zá* *cikád* *mə* *ndəv-yi* *zá*
REL listen listen COMP sesame REL fall-STAT EE
'The one who was good at listening said, "A sesame seed fell down."'
- (3b) *sə* *m* *yár-á-y* *nə* *zá*
1SG REL wander-GO-AFF ? EE
'I have come back (without having succeeded)' (in hunting, for example)
- (4) *kədəm* *mə* *káp-i* *zá*
calabash REL break AFF
'the calabash broke'
- (5a) *á* *pəŋ* *ngáz* *tá* *nkwa*
3SG cut leg GEN goat
'he cut off the leg of goat'.
- (5b) **ngáz* *tá* *nkú* *m* *pəŋ-i* *zá/ká*
leg GEN goat REL cut-AFF EE/CONC
for 'the leg of the goat is cut' or any other meaning
- (5c) *nkú* *m* *pəŋ-i* *ká*
goat REL cut-AFF CONC
'the goat has been quartered'

References

- Dixon, R.M.W. 2010. *Basic linguistic theory*. Volume 1. Methodology. Oxford: Oxford University Press.
- Dowty, David. 1991. *Thematic proto-roles and argument selection*. *Language* 67: 547---619.
- Frajzyngier, Zygmunt, Eric Johnston with Adrian Edwards. 2005. *A Grammar of Mina*. Berlin/New York: Mouton de Gruyter.
- Levin, Beth, and Malka Rappaport Hovav. 1994. *Unaccusativity: At the Syntax-Semantics Interface*. Cambridge, MA: MIT Press.
- Perlmutter, David M. 1978. *Impersonal passives and the Unaccusative Hypothesis*. 4th Annual Meeting of the Berkeley Linguistics Society. UC Berkeley, 157–189.
- Van Valin, Robert. 2005. *Exploring the syntax semantics interface*. Cambridge: Cambridge University Press.

Affix pleonasm—towards a typology of a multifaceted phenomenon

Pleonasm is a pervasive phenomenon that occurs at all levels of linguistic analysis and language processing, including discourse, the lexicon, morphology, and syntax. Pleonastic formations are attested in a wealth of languages and, within each language, at several stages of its evolution. Therefore, pleonasm is an ideal candidate both for the investigation of individual languages and for large cross-linguistic studies. While a few case studies on single languages have been published (e.g. Stolz 2010), no work committed to this multifaceted phenomenon from a typological orientation has been produced hitherto (nor do Haspelmath 1993 and Lehmann 2005 constitute any exception, for the former focuses on the subtype of externalized inflection, and the latter is rather concerned with providing a general theoretical frame of pleonasm in language, and inevitably has a quite constrained empirical foundation).

The present paper aims to contribute a first step towards a typology of pleonasm, in terms both of the range of features/meanings/functions which are realized pleonastically, and the rich evidence that the languages of the world provide. The focus is on pleonastic affixation in both inflectional and derivational morphology. The empirical basis is a large body of data drawn from a variety of typologically heterogeneous languages.

The paper attains the following goals: Firstly, it provides a precise definition of affix pleonasm by constraining its conceptual scope based on semantic and functional (i.e. morphosyntactic) motivation; secondly, it distinguishes two main types of pleonastic realization, viz. implicit pleonastic marking, e.g. the addition of the overt feminine marker *-aa* to inherently feminine nouns in Hausa (Newman 1979), vs. explicit pleonastic marking, e.g. the Modern Greek *manulitsa* ‘mommy’, which is formed via diminutivizing *-ul(a) + -its(a)* suffixation of the base *man(a)* ‘mother’; thirdly, it details the realizational scope of affix pleonasm with respect to: (a), the areas of grammar in which it occurs, viz. morphosyntactic features, such as case, number, gender, person, and a wealth of derivational formations, including the agent nouns, abstract nouns, processes of intensification, among many others; (b), the parts of speech which it affects, viz. nouns, adjectives, verbs, adjectives, and conjunctions; and, (c), the position that pleonastic affixes assume, viz. suffixes before interfixes and infixes.

In light of the evidence provided, the paper suggests that neither a purely diachronic nor a purely synchronic perspective helps us properly understand the dimension of affix pleonasm in terms of its motivations and psycholinguistic salience. Rather, a gradual variationist approach is propelled forward, in order to be able to account both for sociolectal and idiolectal variation, and the role that pleonasm plays in language acquisition, intermediate evolutionary stages (e.g. along the path of grammaticalization), and language contact.

References

- Haspelmath, Martin. 1993. The diachronic externalization of inflection. *Linguistics* 31(2). 279–309.
- Lehmann, Christian. 2005. Pleonasm and hypercharacterisation. In Geert E. Booij & Jaap v. Marle (eds), *Yearbook of morphology 2005*, 119–154. Dordrecht: Springer.
- Newman, Paul. 1979. Explaining Hausa feminines. *Studies in African Linguistics* 10(2). 197–226.
- Stolz, Thomas. 2010. Pleonastic morphology dies hard: Change and variation of definiteness inflection in Lithuanian. In Franz Rainer, Wolfgang U. Dressler, Dieter Kastovsky & Hans Christian Luschützky (eds), *Variation and change in morphology. Selected papers from the 13th International Morphology Meeting, Vienna, February 2008*, 217–244. Amsterdam: Benjamins.

A cross-linguistic examination of the Symmetry & Dominance Constraints

It has been suggested that the Symmetry and Dominance (S&D) Conditions, proposed by Battison (1) to describe the restrictions on signs produced with two hands are universal across signed languages (2). Yet, while Battison himself never claimed S&D to be universal, they are often presumed to hold cross-linguistically. The Symmetry Condition stipulates that for signs in which both hands move, the hands should be specified for the same configuration. The Dominance Condition requires that signs in which the hands do not bear the same configuration, the weak hand must bear an unmarked handshape and must remain passive; only the dominant hand can move. To assess whether it is reasonable to assume S&D can be applied to other languages, I evaluate data from a sample of nine sign language dictionaries (Brazilian, Chinese, Dutch, French, Italian, Hong Kong, Nicaraguan, Indian, and Mongolian) and discuss the types of violations exhibited in each language and the implications for S&D. I argue that the typological variation in data can be productively accounted for using Optimality Theory (OT) and building from recent work by Eccarius (3) in this framework.

From 3,454 signs, 100 S&D violations were identified. Violations fell into four main categories: signs with unified hand configuration (where both articulators move but do so together, like the ASL signs LEAD (Fig. 1 a) and TAKE - RIDE (Fig. 1 b), signs in which a marked base handshape was used, signs in which both hands move but the movement is not unified, and signs which violated S&D in some other manner.

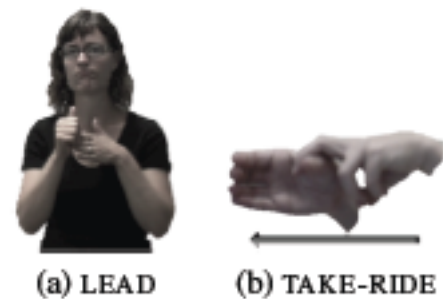


Figure 1: S&D violations in ASL

The data show that each language has signs which violate S&D, but the extent to which this is allowed varies. For example, French, Nicaraguan and Italian SLs had fewer than 2% of S&D violations, while Chinese exhibited more than 10%. Trends emerged in the data with respect to the types of violations each language exhibited. One way to capture these patterns and to predict which types of violations will surface in different languages is to use OT. Prince & Smolensky's (4) model assumes that formational constraints that govern well-formedness can be ranked differently across languages yielding different types of outcomes; in this case a typology of two-handed signs. Eccarius suggested one type of faithfulness constraint necessary to account for sign language data is faithfulness to some real-world entity being depicted with the hands. Such an analysis, given the right constraints, can account for the S&D violations in the present investigation (Fig. 2). This tableau represents a constraint ranking necessary to demonstrate a language which allows the two hands to bear different configurations only if both hands move together. The first and second constraints, IDENT UNIF and IDENT SF reflect faithfulness to real-world referents. Respectively, taking a ride in a vehicle necessitates both entities move together and the H handshape on the dominant hand represents a two-legged entity. The third constraint, a markedness constraint, reflects a dispreference for hands bearing different configurations, but because the other constraints rank above it, it is violated in this instance and others throughout the dataset analyzed here.

Fig. 1b [TAKE-RIDE(C:H)]	IDENTUNIF	IDENTSF	*DIFFHS
a. TAKE-RIDE(C:H)			*
b. TAKE-RIDE(C:C)		*!	
c. TAKE-RIDE(C:H only H moves)	*!		

Figure 2: Tableau demonstrating the ranking IDENTUNIF»IDENTSF»*DIFFHS. Letters in parentheses indicate the handshapes each hand assumes in the production of the sign for the different candidates. IDENTUNIF states that corresponding elements between external referent contrasts and output contrasts should have comparable amounts of unification, IDENTSF states corresponding segments between external referent contrasts and output contrasts should have identical Selected Finger combinations (3), and *DIFFHS states that hands must be specified for the same configuration.

References

- [1] BATTISON, R. Lexical borrowing in American Sign Language. Linstok Press, Silver Spring, 1978.
- [2] CHANNON, R. The symmetry and dominance conditions reconsidered. In Proceedings of Chicago Linguistic Society (2004), vol. 40, pp. 44–57.
- [3] ECCARIUS, P. A constraint-based account of distributional differences in handshapes. In Formational Units in Sign Languages, R. Channon and H. van der Hulst, Eds. De Gruyter, Nijmegen, The Netherlands, 2011, pp. 261–284.
- [4] PRINCE, A., AND SMOLENSKY, P. Optimality Theory: Constraint Interaction in Generative Grammar (ROA Version 8/2002). Rutgers Center for Cognitive Science, Rutgers, NJ, 1993.

Social and Grammatical Complexity: Language Spread and TAM Marking

A naive layman's view would have it that more complex societies have more complex languages. Professional linguists have rejected this view, initially by adopting the dogma that all languages are of equal grammatical complexity, and more recently with proposals by McWhorter, Trudgill and others to the effect that languages of large nation states actually tend to be simpler than their counterparts associated with smaller speech communities. This paper, however, argues that in one domain at least, the original naive layman's view is factually correct.

This paper presents empirical evidence for a correlation between language spread and Tense-Aspect-Mood (TAM) marking. A typological study of 868 languages distinguishes between two grammatical types:

- (1) TAM marking
 - (a) *obligatory*: if all basic declarative affirmative main clauses contain an overt grammatical expression of at least one TAM category
 - (b) *optional*: if basic declarative affirmative main clauses may occur without any overt grammatical expression of any TAM categories

Worldwide, both types enjoy widespread distribution, though with striking areal patterns: a large contiguous area of obligatory TAM marking encompasses East and North Africa plus West, North and South Eurasia, while another large area of optional TAM marking includes Mainland and Insular Southeast Asia plus parts of Western New Guinea.

The occurrence of obligatory TAM marking correlates positively with languages that belong to large genealogical families that have, at some time in the past, undergone spread over extensive land masses. Quantitatively, this correlation can be demonstrated by assigning to each language various indices representing the size of its genus and/or family in accordance with various databases such as WALS or Ethnologue. Qualitatively, this correlation can be observed in specific families or regions characterized by a mix of spread and other languages. For example, within the Niger Congo family, Bantu languages, a prime example of language spread, are overwhelmingly obligatory TAM marking, while non-Bantu languages are majority optional TAM marking. Similarly, in the greater New Guinea area, amongst the Non-Austronesian languages, the one large spread family, namely Trans-New-Guinea, is almost exclusively obligatory TAM marking, while families belonging to other smaller families are majority optional TAM marking.

The correlation between language spread and TAM marking may thus be viewed as an instance in which social and grammatical complexity correlate positively. Although seeming to contradict the abovementioned proposals by McWhorter and Trudgill, there is in fact no conflict, since the correlations apply at different historical stages. For example, complexification and the development of obligatory TAM marking would have occurred at the earliest stages of Indo-European if not before, while simplification takes place much later as English becomes a "Non-Hybrid Conventionalized Second-Language" and its descendant

creole varieties then move from obligatory to optional TAM marking. In conclusion, it is suggested that the correlation between social and grammatical complexity demonstrated here may provide a model for similar developments in the realms of early child language acquisition and in the evolution of human language itself.

Theme session: Typological hierarchies in synchrony and diachrony

Motivated versus unmotivated pathways in the evolution of main clause alignment patterns

Blevins (2004) and Givón (2008) argue that if we can explain synchronic linguistic patterns via their origins, then we require no synchronic explanation (i.e., a putative function that is served, or a markedness hierarchy). This paper will discuss four distinct hierarchical alignment patterns, all of which mark the relative places of certain referent (or NP) types on a referential hierarchy. In these systems, personal indices on the verb cross-reference the higher-ranked of two core arguments. Sometimes the lower-ranked argument occurs marked as an oblique (OBVIATIVE) and sometimes a DIRECT/INVERSE morpheme indicates the role of the verbal index.

We examine the diachronic sources of each of these patterns in four languages: Panare (Cariban, Venezuela), Potosino Huastec (Mayan, Mexico), Arizona Tewa (Tanoan, USA), Reyesano (Tacanan, Bolivia).

Both Panare and Huastec show a situation where loss of a prior third person morpheme creates a single indexation slot, which is then filled solely by the SAP prefix, thereby creating an unmotivated synchronic hierarchy. In Panare (Payne & Payne 2013), distinct person prefixes index first and second person (SAP) A and P, but not third person except when there is no SAP core argument. The 2P prefix occurs when 1A acts on 2P, the 2A prefix and the 1P proclitic when 2A acts on 1P. A prefix γ - 'INVERSE' also occurs when 3A acts on SAP P. To create this alignment, Panare lost a 3P prefix from Proto-Cariban, changed from a neutral 1A2P/2A1P form, and reduced the distribution of $*j$ - 'relator' to create γ - 'inverse'. In Potosino Huastec (Zavala 1994), the pan-Mayan ergative (Set A) and absolutive (Set B) prefixes are augmented by an inverse-like morpheme (*ta-*, possibly a cislocative); the 3B form was already \emptyset - and the 3A form changed from *in-* to \emptyset -, creating a fairly typical-looking hierarchical system. Colonial records show that the addition of *ta-* came first (already in place by 1767), with *in-* '3A' lost since then.

In Reyesano (Guillaume 2011), second position SAP person clitics became verbal prefixes (invariant for role), interacting with a prior 3A person suffix (which somewhat resembles an inverse morpheme). The set of clitics did not have a third person form, so the absence of third person forms in verbal prefixation again appears to be incidental.

In contrast, the Arizona Tewa system appears to come from a functionally motivated source. Arizona Tewa (Zúñiga 2006) develops hierarchical indexation by reanalysis of a passive clause as the obligatory means of indicating an inverse situation, that is, where 3A acts on SAP P. Thus, the person indexes on the verb reflect only the SAP, but they differ for SAP A (< transitive subject) and SAP P (< passive subject) and third person has differential case: 3A is marked as obviative (< agent of passive) or unmarked (< transitive subject) depending on person or topicality of P. This alignment of form and function is motivated by discourse topicality (as predicted by Givón 1994).

Theme session: Linked Data in Linguistic Typology

Fine-grained typological investigation of grammatical constructions using Linked Data

Overview: This talk discusses the use of Linked Data to create a database of grammatical constructions, emphasizing the potential advantages of Linked Data in such a context and clarifying the extent to which the technology is ready to be utilized more generally for typological work. Unlike many prominent projects making use of Linked Data (e.g., Nordhoff 2012), the technology was not chosen specifically to facilitate interoperability with other datasets but, rather, because it was well suited to model grammatical objects of theoretical interest. The details of this work, therefore, should be of value both to those interested in data interoperability and to those engaged in more traditional typological investigation.

A database of templates: Linearization templates—that is, grammatical devices describing unexpected patterns of linear stipulation—do not lend themselves to straightforward typological classification due to their heterogeneous nature, ranging from Semitic CVskeletons, to Athabaskan position-class morphology, and beyond (Good 2011). Describing templates in ways which allow them to be rigorously compared, therefore, requires a flexible and easily extendable database system.

Bickel (2010) deals with comparable concerns for clause-linkage constructions by adopting a “multivariate” approach. This involves the development of an extensible database that allows for the coding of an open-ended range of patterns of variation as they are discovered. Bickel’s specific implementation is effectively based on a model where a construction is conceptualized as a “bag” of feature-value pairs. This approach is appropriate for his dataset but requires refinement to be applied to a database of templatic constructions. This is because, in addition to encoding a template’s holistic properties, it is clearly also desirable to encode fine-grained structural relations among its subcomponents.

The role of Linked Data: Formal syntactic approaches making use of feature structures like HPSG (Sag et al. 2003) provide well-developed solutions for encoding structural relations in the form of nested attribute-value matrices. However, associated implemented systems (e.g., Copestake 2002) are designed to describe grammars of individual languages rather than for cross-linguistic investigation. Linked Data, by contrast, can both readily encode nested attribute-value matrices and allow for data to be straightforwardly embedded within an ontology, facilitating rigorous comparison. It is, therefore, an excellent tool for describing templatic constructions typologically and, indeed, for any type of grammatical construction that can be modeled using sets of nested attribute-value pairings.

The Linked Data database of templatic constructions forming the basis of this talk was created using the readily available Protégé tool and is stored in an XML format which can be queried for typological purposes using existing code libraries. A fortunate byproduct of the use of Linked Data is that the database is immediately available in an open, interoperable format, lowering technical barriers to data sharing. Thus, a technology originally chosen for its expressive power also has important advantages with respect to data reusability. While the skillset required to exploit the relevant technologies is not yet common among typologists, it is not obviously more complex than what is needed to use the popular R statistical programming language, meaning the barriers are not as high as they may otherwise seem.

References

- Bickel, B. 2010. Capturing particulars and universals in clause linkage: A multivariate analysis. In *Clause-hierarchy and clause-linking: The syntax and pragmatics interface*, 51–102. Amsterdam: Benjamins.
- Copestake, A. 2002. *Implementing typed feature structure grammars*. Stanford: CSLI.
- Good, J. 2011. The typology of templates. *Language and Linguistics Compass* 5:731–747.
- Nordhoff, S. 2012. Linked data for linguistic diversity research: Glottolog/Langdoc and ASJP Online. In *Linked data in linguistics: Representing and connecting language data and language metadata*, 191–200. Berlin: Springer.
- Sag, I. A. et al. 2003. *Syntactic Theory: A formal introduction*. Stanford: CSLI.

Theme session: Typological hierarchies in synchrony and diachrony

Diachronic factors and the animacy hierarchy in Old Irish

Old Irish has a set of enclitic pronouns which may optionally attach to person- and numbermarked constituents, i.e. verbs, conjugated prepositions, and possessed NPs:

- | | | | |
|-----|-----------------------|---------------------|-------------------|
| (1) | a. dó(=som) | b. a=macc(=si) | c. biru(=sa) |
| | to.3SG.M./N.(=3SG.M.) | 3SG.F.=son(=3SG.F.) | carry.1SG.(=1SG.) |
| | “to him” | “her son” | “I carry” |

Pronominal objects are also marked on the verb in Old Irish, and these clitics may agree with the object. As noted in Griffith (2008), however, only one such clitic may appear on the verb, and it must agree with the argument highest on the animacy hierarchy: 1st person > 2nd person > 3rd person animate > inanimate:

- | | | |
|-----|----------------------------------|-----------------------------------|
| (2) | a. no=m·bir(=sa/*su) | b. ata·cí(=su/*si) |
| | PVB=1SG.·carry.2SG.(=1SG./*2SG.) | PRV.3SG.F.·see.2SG(=2SG./*3SG.F.) |
| | “you (sg) carry me” | “you (sg) see her” |

The only proposal treating the diachronic development of this hierarchy is Griffith (2011). His claims are plausible on the whole, though they rely on a number of assumptions that are difficult or impossible to prove (e.g. the argument status of the clitics, the chronology of cliticization), and he does not address why only one such clitic appears after the verb. If two clitics were allowed, there would be no hierarchy.

Now, if we follow Schrijver (1997: 18-25) and assume that these clitic pronouns have a deictic origin, an answer to this problem presents itself. It appears that the change in meaning from deictic to pronominal was rather late. Presumably, having two clitics of conflicting deixis attached to the same verbal form was problematic. Only one was allowed. After the rather late transition to pronominal meaning, the restriction remained in place.

It still remains to explain why the one available clitic agreed only with the argument topmost on the animacy hierarchy. We can adopt a suggestion from Griffith (2011): the clitic pronouns serve to indicate topicality. Since 1st persons are more topic-worthy (Wierzbicka 1981), when there was competition between two potentially topical elements, only the highest was permitted. While this explanation appears to replace one hierarchy with another and thus may seem circular, the topic-based account appears to have more explanatory power than the animacy-based one.

In summary, this paper notes a case of the animacy hierarchy in Old Irish and presents language-specific arguments for a diachronic pathway leading to this state of affairs. Due to the deictic origins of the pronouns, only one could appear on a given verb. The inherent topic-worthiness of the arguments then determines which clitic may appear on that verb.

References

- Griffith, Aaron. 2008. “The animacy hierarchy and the distribution of the *notae augentes* in Old Irish.” *Ériu*, 58: 55-75.
- Griffith, Aaron. 2011. “The genesis of the animacy hierarchy in the Old Irish *notae augentes*.” In *Akten der XIII Fachtagung der idg. Gesellschaft*, ed. Krisch and Lindner, 182-191. Wiesbaden.
- Schrijver, Peter. 1997. *Studies in the History of Celtic Pronouns and Particles*. Maynooth.
- Wierzbicka, Anna. 1981. “Case marking and human nature.” *Australian Journal of Linguistics*, 1: 43-80.

Distinguishing between class terms, classifiers and noun classes: the case of Movima (Bolivia)

There seems to exist pervasive confusion surrounding the phenomenon of nominal classification systems found in languages of the world, partly induced by the nature of misnomer of the term “classifier” itself. There are those that lump all the systems together and tend to call them all classifiers (Aikhenvald 2003, 2012) and those that try to keep them apart as separate systems, even if sharing lexical origins (Craig 1986, Grinevald 2000). This paper will be an exercise in clarification of the issues involved, invoking the need to always keep the dynamics of the systems in view, and the evaluation of the productivity and degree of grammaticalization of any system.

Amazonian languages have often been cited as particularly challenging to the typology of noun classification, since their systems seem to cut across the different types identified so far (see Aikhenvald 2012, Derbyshire and Payne 1990, Grinevald 2000). The Movima language (isolate, Bolivia; Grinevald 2002, Haude 2006) is a case in point. In Movima, noun roots (e.g. *ba* ‘fruit’) can be attached to nominal, verbal, and numeral bases in order to specify particular classes of objects that are named, acted upon, or counted. These roots function as class terms (e.g., a plant name combined with *ba* denotes a particular kind of fruit) or classifiers (characterizing entities according to their shape or consistency; e.g., *ba* is used to refer to three-dimensional, fist-sized round objects), and they can serve to create anaphoric reference in discourse in a way reminiscent of noun class systems. Apparently the Movima system has undergone several modifications over time, through which the inventory of bound nominal elements has become more heterogeneous. At one point – as can be seen, from example, from the treatment of early Spanish loans – noun classes were created by simply truncating the last syllable of a noun, a process which today is no longer productive. Discourse data furthermore show that the anaphoric function of the classifying elements is in decay, too, since words that require such an element (e.g. numerals) usually take the default, semantically neutral element *-ra*.

With Movima as an example, the paper will demonstrate among other things the relations between class terms and noun class systems, the need to keep track of levels of lexicalization of items, the evolution from semantically motivated classification to phonological truncation for generating new noun classes, and the consequences of eventual fossilization of the system.

References:

- Aikhenvald, Alexandra Y. 2003. *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2012. *The languages of the Amazon*. Oxford: Oxford University press.
- Craig, Colette. 1986. *Noun classes and categorization*. Amsterdam: John Benjamins.
- Derbyshire, Desmond C. and Doris L. Payne. 1990. *Noun classification systems of Amazonian languages*. In Amazonian Linguistics, Doris L. Payne (ed.), 243-271. Austin: University of Texas Press.
- Grinevald, Colette. 2000. *A morphosyntactic typology of classifiers*. In *Systems of Nominal Classification*, Senft, Gunter (ed.), 50-92.
- Grinevald, Colette. 2002. *Nominal classification in Movima*. In *Current Studies on South American Languages*, Crevels, Mily, Simon van de Kerke, Sergio Meira, and Hein van der Voort (eds.), 215-239. Leiden: CNWS Publications.
- Haude, Katharina. 2006. *A Grammar of Movima*. Doctoral dissertation, Radboud University Nijmegen.

A rare case of Differential Marking on S/A: the case of Coptic

Differential Case Marking (DCM), i.e. the phenomenon whereby some core NPs are marked by case in certain environments but not in others has attracted a great deal of attention in recent years (Bossong 1985, Comrie 1989, Aissen 2003, Malchukov 2007).

In this paper, we present the distribution of DCM in Coptic and we argue that Coptic presents a cross-linguistically rare, if not unique, type of DCM system, in that i) both Subject/Agent and Patient can receive DCM, which alternates with incorporation, and ii) the factors that govern the appearance of DCM are extremely rare cross-linguistically.

Grammatical relations in Coptic show a quite elaborate coding system. There is pragmatically determined word order, with the preverbal position used to accommodate both topic and focus in verbal clauses. S/A can be either incorporated (1, 3) or topicalised (2). When postverbal, S/A is encoded by the NOM marker *nci* (4), which is found only with full lexical NPs as opposed to pronouns. The picture is further complicated when the encoding of P is considered. P arguments can be incorporated, as in (1) and (2). If not incorporated, they receive overt accusative marking. This alternation is strictly regulated by referential factors in the imperfective tenses. In non-imperfective tenses, the factors regulating the selection of incorporation vs. accusative marking is still poorly understood.

In this presentation, we concentrate on the encoding of S/A in Coptic. Based on textual data, we show that the three different strategies employed to encode S/A respond to different information-structural statuses of the S/A referent. While subject fronting is deployed both as a topicalization or focalization strategy, and subject incorporation typically occurs with frame-evoked referents, the most noteworthy strategy of Coptic is the overt encoding of postverbal S/A. It has been argued (Loprieno 2000, Shisha-Halevy 1986) that the main factor determining the choice of this construction with postverbal S/A is information structure.

However, unlike what is commonly found in information-structural based DCM with S/A, DCM is not triggered by focus (Fauconnier 2010, among others), but is rather a grammaticalized “antitopic” marker (Chafe 1994) used to reintroduce in the discourse identifiable S/A that are discourse-old or otherwise presupposed (Prince 1992).

There are further interesting details which set Coptic apart from other languages with DCM on S/A. First, the NOM marker sporadically spreads to non-nominative NPs, such as P or obliques and possessors. Second, though the above description fits all the Coptic dialects, in one dialect, we find a slightly different system where 1st/2nd person independent pronouns can also be marked as nominative unlike the other dialects where marking on postverbal pronouns is not allowed. Interestingly, a different marker is employed, giving thus rise to a tripartite DCM system for S/A.

- (1) Case-marking of lexical NPs -> NOM/ACC (S/A) topicalized**
- (a) *a-f-nouje ebol n-ne-pn(eu)ma n-akatharton*
 PST-3SGM-throw out ACC-ART.PL-spirit of-impure
 'He cast out the impure spirits.'
- (2) Incorporation of lexical A/S**
- (a) *a-ke-aggelos ei ebol*
 PST-another-angel come out
 'Another angel came out'
- (b) *auO ntere-pe-pn(eum)a n-akatharton raht-f e-p-kah*
 and temp-the-spirit of-impure throw-3SGM to-the-ground |
 'And when the impure spirit had thrown him to the ground...'
- (3) Incorporation of lexical P**
- (a) *e-k-e-taie-pek-eiOt mn-tek-maau*
 OPT1-2SGM-OPT2-honour-your-father with-your-mother
 'You shall honour your father and your mother.'
- (4) Incorporation of both A and P**
- (a) *a-SptOre take-mEESe eusoutOn*
 PST-handshake destroy-crowd upright:EPL
 'Handshakes have ruined multitudes who are upright.'
- (5) Postverbal overt nominative**
- (a) *a-f-ei nci-iEsus*
 PST-3SGM(S)-come NOM-Jesus
 'Jesus came'
- (b) *a-f-jnou-f nci-petros*
 PST-3SGM(A)-ask-3SGM(P) NOM-Peter
 'Peter asked him'

References:

- Aissen, J. 2003. "Differential object marking: iconicity vs. economy." *Natural Language and Linguistic Theory* 21, 435-483.
- Bossong, G. 1985. *Differentielle Objektmarkierung in der neuiranischen Sprachen*. Tübingen: Narr.
- Comrie, B. 1989. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Chafe, W. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Fauconnier, Stefanie. 2010. Differential Agent Marking and animacy. *Lingua* 121: 533-547.
- Loprieno, A., 2000. 'From VSO to SVO? Word order and rear extraposition in Coptic,' in R. Sornicola, E. Poppe and A. Shisha-Halevy, *Stability, variation and change in word-order patterns over time*. Amsterdam: Benjamins.
- Malchukov, Andrej. 2007. "Animacy and asymmetries in differential case marking". *Lingua* 118: 203-221.
- Prince, E. 1997. "The ZPG Letter: Subjects, Definiteness and Information Status" In S. Thompson and W. Mann (eds) *Discourse Description Diverse Analyses of a Fundraising Text*: Amsterdam: Benjamins.
- Shisha-Halevy, A. 1986. *Coptic Grammatical Categories: Structural Studies in the Syntax of Shenoutean Sahidic*. Roma: Pontificium Institutum Biblicum.

Towards a typology of associated motion in South American languages and beyond

Grammatical markers of associated motion (AM), a newly recognized typological category (Koch 1984, Wilkins 1991, Guillaume 2006), primarily attach to non-motion verbs and express the fact that the verb action (V) is associated with a backgrounded motion which can be temporally **prior** ('go and V', 'come and V', etc.), **concurrent** ('V while going', 'V while coming', etc.) or **subsequent** ('V and go', 'V and come', etc.). The category of AM contributes to the typology of motion events in recognizing the possibility that in some languages the (translational) motion component of a motion event be expressed by grammatical morphemes rather than lexical verbs (Levinson & Wilkins 2006, Guillaume 2006).

Initially proposed and discussed in the descriptive literature on **Australian** languages (Koch 1984, Tunbridge 1988, Austin 1989, Wilkins 1991, Nordling 2001, Dixon 2002), the category of AM has also been recognized in a number of languages from other parts of the world, especially **South America** (Sakel 2004, Guillaume 2008, Silva 2011, Vuillemet 2012, Rose under review, Fabre under review), **Central America** (Zavala 2000, O'Connor 2007, Caballero 2008, McFarland 2009) and **North America** (Dryer 2007). More marginally, AM systems have also been identified in **Africa** (Bourdin 2005, Voisin to appear) and **Asia** (Jacques to appear).

In this paper, I will be concerned with the expression of AM in South America. I will present the results of an on-going investigation of this category in some **45 neighboring South-American languages** spread over Bolivia, Peru et Western Brazil, and belonging to approximately 20 distinct genetic groupings.

The goals of the study are to answer the following questions:

- **descriptive**: Which of these languages have AM systems, whether recognized as such or discussed under a different terminology? How complex are these systems? Which parameters and semantics features are needed to analyze these systems?
- **typological**: How do these parameters and features distribute across different languages? Are there any typological correlations / implicational dependency between them?
- **genetic & areal**: How do the identified types of AM systems distribute across distinct genetic grouping and geographic regions?

Among the results to be discussed during the talk, one can highlight the following:

- AM is a **widespread** phenomenon in South America, being found in an overwhelming majority of the languages surveyed
- in several of these languages AM is expressed by **unusually complex systems** of 6 or more distinctions (sometimes up to 13 distinctions)
- the distinctions basically operate according to the following 4 parameters: (1) grammatical function of the **moving argument**, (2) **temporal relation** between action and motion, (3) **path** of the motion and (4): **aspectual realization** of the action with respect to the motion
- the manifestation of the parameters and the distinctions correlate with the level of complexity according to the following implication scale: **prior motion** of the subject > **concurrent motion** of the subject > **subsequent motion** of the subject > **motion of the object**
- the most complex AM systems are found in the neighboring **Arawak, Panoan and Tacanan** families along the western margins of the Amazon basin down the eastern foothills of the Andes

References

- Austin, Peter. 1989. Verb compounding in Central Australian languages, in *La Trobe University Working Papers in Linguistics*: 2, 1-31.
- Bourdin, Philippe. 2005. The marking of directional deixis in Somali. In *Studies in African Linguistic Typology*. Voeltz Erhard (ed.). Amsterdam / Philadelphia: John Benjamins. 13-41.
- Caballero, Gabriela. 2008. *Choguita Rarámuri (Tarahumara) Phonology and Morphology*. Ph.D. dissertation. University of California, Berkeley.
- Dixon, R.M.W. 2002. *Australian languages: their nature and development*. Cambridge: CUP.
- Dryer, Matthew S. 2007. Kutenai, Algonquian, and the Pacific Northwest from an areal perspective. In *Proceedings of the Thirty-Eighth Algonquian Conference*, edited by H. C. Wolfart, pp. 155-206. Winnipeg: University of Manitoba.
- Fabre, Alain. Under review. Applicatives and associated motion suffixes in the expression of spatial relations: a view from nivacle (mataguayo family, paraguayan chaco)
- Guillaume, A. 2006. La catégorie du 'mouvement associé' en cavineña : apport à une typologie de l'encodage du mouvement et de la trajectoire, *Bulletin de la Société de Linguistique de Paris*, 101.1: 415-436
- Guillaume, A. 2008. *A Grammar of Cavineña*. Berlin and New York. Mouton de Gruyter.
- Jacques, Guillaume, To appear. Harmonization and disharmonization of inflection. *Linguistic Typology*.
- Koch, Harold. 1984. The category of 'associated motion' in Kaytej. *Languages in Central Australia* 1: 23-34.
- Levinson, S. and D. Wilkins. 2006. The background to the study of the language of space. *Grammars of Space*. S. Levinson & D. Wilkins. Cambridge, CUP: 1-23.
- McFarland, Teresa Ann. 2009. *The phonology and morphology of Filomeno Mata Totonac*. Ph.D. dissertation. University of California, Berkeley.
- Michael, Lev. 2011. Matsiguenga directionals and associated motion: a Kleinian descriptive semantics. Paper given at Syntax and Semantics Circle. University of California, Berkeley, December 2.
- Nordlinger, Rachel. 2001. Wambaya in Motion. In *Forty Years On*, J. Simpson, D. Nash, M. Langhren, P. Austin and B. Alpher (eds.), 401-413. Canberra: Pacific Linguistics.
- O'Connor, Loretta M.. 2007. *Motion, transfer, and transformation: The grammar of change in Lowland Chontal*. Studies in Language Companion Series 95. Amsterdam: John Benjamins.
- Rose, Françoise. Under review. On the grammaticalization of associated motion.
- Sakel, Jeanette. 2004. *A Grammar of Mosetén*. Mouton de Gruyter.
- Silva, Léila de Jesus. 2011. *Morphosyntaxe du rikbaktsa (Amazonie brésilienne)*. Doctoral dissertation. Université Diderot - Paris 7.
- Tunbridge D., 1988, Affixes of motion and direction in Adnyamathanha, in P. Austin (ed.), *Complex Sentences Constructions in Australian Languages*, Amsterdam, John Benjamins, 265-83.
- Voisin, Sylvie. to appear. Expressions de trajectoire dans quelques langues atlantiques (groupe Nord), in *Faits de Langues*.
- Vuillermet, M. 2012. *A grammar of Ese Ejja, a Takanan language of the Bolivian Amazon*. Ph.D. dissertation. Université Lumière Lyon 2.
- Wilkins, D. P. 1991. The Semantics, Pragmatics and Diachronic Development of 'Associated Motion' in Mparntwe Arrernte. *Buffalo Papers in Linguistics* 1: 207-57.
- Zavala, Roberto. 2000. Olutec motion verbs: Grammaticalization under Mayan contact, in Andrew K. Simpson, éd., *Proceedings of the Twenty-sixth Annual Meeting of the Berkeley Linguistics Society*, pp. 139-151.

Passives with agent saliency

The traditional treatment of passive sentences as patient-promotional is thought to be illustrated, for example, by languages with formally passive verbs which, in addition, exhibit affixed patients, and do not, furthermore, have an obligatory expression of the agent. This case seems to be exemplified by Indonesian (example 1). However, this treatment has been questioned on the basis of other languages, which seem to require an agent-demotional interpretation (cf. Shibatani 1985). In fact, both views are ultimately based on the Chomskyan analysis of passive sentences as transformations of active sentences.

However, agents, far from always being transformed subjects, can appear in the morphology even when the meaning is not active: some languages exhibit verbs which, although they have a passive meaning, contain a personal agent affix; an example is Latin *vapulo* “I am thrashed”. And conversely, there are languages in which the verb meaning “to be”, which would seem to rule out any passive form (since it expresses a state or an essence, and, consequently, does not have an active meaning), does, nevertheless, have a passive, and, thus, may suggest one to posit an agent (admittedly an impersonal one). Lithuanian is such a language, as illustrated by example 2.

Moreover, agent markers can serve as morphological elements through which passive verbs are formed: such passive verbs, when analyzed literary, appear as active subject-verb structures. This case is exemplified by such languages as Ainu or Kimbundu, in which the agent markers which make part of the structure of passive verbs are, in fact, first inclusive and third plural pronominal elements respectively (examples 3 and 4).

Some languages exhibit even more explicit agent-promotional structures. A well studied case is represented by anti-passive sentences (example 5, from Warrungu). We find another case in negative passive sentences that express the inability of the agent. An illustration is Japanese (example 6). Another is Hindi (ex. 7a and its neutral counterpart 7b). A third device stressing agent saliency consists of reduplicating the agent, first marked as an affix on the verb, but in addition also marked as an adverbial complement. We find this double agent-marking structure in such Mon-Khmer languages as Semai (example 8), and also in Austronesian languages, like Acehnese (example 9). Some verb-initial languages even offer a more striking, and rare, structure, i.e. a co-reference phenomenon between a pronominal agent and a sentence-final reflexive pronoun, as illustrated by Tagalog (example 10) or Toba Batak (example 11). Whereas there is a generalization that in simple sentences subjects control reflexivization, what we observe in these sentences is a reflexive pronoun which is itself the subject, and which expresses the same participant as the agent.

All these agent saliency phenomena may be opposed to passives with agent occultation, like the one found in such classical Semitic languages as Biblical Hebrew or Koranic Arabic: the latter is illustrated in example 12 (in which, following a rule in this language, there is no number agreement between the verb, *wu'ida* “was promised”, and the sentence-final subject *(a)l-muttaqu:n*, because the word-order is VS: literally, we have “the paradise which god-fearing men was (= “were”) promised).

Thus, agent occultation in passive sentences is, actually, in polar relationship with agent saliency, which can be considered to be the other pole in this continuum.

Examples:

1. Indonesian (S. Wulandari, pers. comm.): *penjual ini di-teriak-i (oleh mahasiswa itu)* (seller this PASS.MARK.-scold.PRET-PAT.MARK.) (by student that) “this seller was scolded (by that student)”.
2. Lithuanian (Eckert 1999: 154): *jõ bũta kareĩvio* (3MASC.SG.GEN) be(PASS.NEUT.PRET.PARTIC.) soldier(SG.GEN.) “he was a soldier”.
3. Ainu (Shibatani 1985: 824): *chip a-nukar* (ship 1INCL-see) “a ship is seen”.

4. Kimbundu (Givón 1981: 182): *nzua a-mu-mono kwa meme* (John 3PL.SUBJ.-3SG.OBJ.-see by me) “John was seen by me”.
5. Warrungu (Tsunoda 1988: 602): *ngaya nyaka-kali-n wurripa-wu katyarra-wu* (1SG.NOM. search-ANT.-NONFUT bee-DAT. opossum-DAT.) “I was looking for bees and opossums”.
6. Japanese (Shibatani 1985: 823) *boku wa nemur-are-nakat-ta* (1SG TOP sleep-PASS.-NEG.-PAST) “I could not sleep”.
7. Hindi (Davison 1982: 158): a. *mujh-se kuch bhī kahā nahī gayā* (1SG.-INSTR. nothing also say.PAST NEG PASS.PAST) “I couldn’t say anything”
b. *maī-ne kuch bhī nahī kahā* (1SG.-ERG. nothing also NEG say.PAST) “I didn’t say anything”.
8. Semai (Diffloth 1974: 132) *tley-?ajeh ?nj- ca: la-?enj* (that-banana 1SG.-eat by-1SG.) “that banana was eaten by me”.
9. Acehnese (Durie 1988: 109): *jih lôn-peu-ingat lé lôn geu-peureksa lé dokto* (3FAM. 1SG.-CAUS.-remember by 1SG. 3POL.-examine by doctor) “he was reminded by me to be examined by the doctor”.
10. Tagalog (Schachter and Otnes 1972: 138): *s-in-ak-tan ko ang sarili-ko* (hurt-PASS.MARK.<->-hurt-DIR 1SG.AG ART self-1SG.) “I hurt myself”.
11. Toba Batak (Keenan 2007: 126): *di-pukkul si Bissar diri-na* (PASS.MARK.-hit ART. Bissar self-his) “Bissar struck himself”.
12. Classical Arabic (Koran, XIII, 35): *mathalu l-ğannati (a)llati: wu’ida (a)l-muttaqu:n* (image.NOM ART-paradise(FEM).GEN REL.PR.FEM.SG promise.PASS.PAST.MASC.SG. ART.-God.fearing.men.NOM.PL) “Such is the paradise which was promised to those who fear God”.

References

- Davison, A. 1982. *On the form and meaning of Hindi passive sentences*. In: *Lingua*, 58, 149-189.
- Diffloth, G. 1974. *Body moves in Semai and in French*. In: *Papers from the 10th Regional Meeting of the Chicago Linguistic Society*. Chicago. 128-138.
- Durie, M. 1988. *The so-called passive of Acehnese*. In: *Language*, 64, 104-113.
- Eckert, R. 1999. *Die Baltischen Sprachen*. In: *Sprachen in Europa*. Innsbruck: Innsbrucker Beiträge zur Kulturwissenschaft, 147-162.
- Givón, T. 1981. *Typology and functional domains*. In: *Studies in Language*, 5, 163-193.
- Keenan, E. 2007. *The syntax of subject-final languages*. In: W. Lehmann (ed.) *Syntactic typology: Studies in the phenomenology of language*. Austin: University of Texas Press, 120-132.
- Schachter, P. and F. T. Otnes. 1972. *Tagalog Reference Grammar*. Berkeley, Los Angeles, London: University of California Press.
- Shibatani, M. 1985. *Passives and related constructions*. In: *Language*, 61, 821-848.
- Tsunoda, T. 1988. *Antipassives in Warrungu and other Australian languages*. In: M. Shibatani (ed.) *Passive and voice*. Amsterdam: John Benjamins, 595-651.

The Basic Word Order Typology: Universality, Genealogy, Areality

This paper will investigate one of typology's most celebrated themes, 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 results from over 3000 languages (combining WALS, Ethnologue and the database of the author). Using orthodox sampling procedures, there are no significant correlations between word order type and population sizes, as has sometimes been suggested using unorthodox sampling procedures (Nettle 1999). We propose that the distribution of word order frequencies can be accounted for by three factors (remaining variance can be accounted for by chance):

Universal:

A consistent frequency distribution with SOV being the most common word order reappears across genealogically and areally stratified subsamples. This effect must thus be accounted for as a universal functional tendency. Contrary to popular belief, SVO is not almost as common as SOV, neither is it consistently the 2nd most common word order across genealogically and areally stratified subsamples.

Genealogical:

Large families exhibit vastly different internal basic word order frequency distributions, which is difficult to account for simply by universal transition probabilities and birth-death effects within a family (as per, e.g., Maslova 2000). It seems necessary, therefore, to invoke family specific biases, which are perhaps transmitted with the remaining (nonbasic-word-order) typological profile of the family.

Areal:

Using novel techniques for inducing areas without starting from a set of predefined areas (Hammarström and Güldemann 2012), we can show that, with respect to basic word order, large (i.e. continent size) areas do exist.

References

- Dryer, M. S. (1992). The greenbergian word order correlations. *Language*, 68(1):81_138.
- 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.
- 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 Wortfolgeregustern*. Hamburg: Helmut Buske.
- Hammarström, H. and Güldemann, T. (2012). Quantifying geographical determinants of large-scale distributions of linguistic features. Paper presented at the Workshop on Quantitative Approaches to Areal Linguistic Typology, Amsterdam, Dec 13, 2012.
- Hawkins, J. A. (1983). *Word order universals*, volume 3 of *Quantitative Analyses of Linguistic Structure*. San Diego: Academic Press.

- Maslova, E. (2000). A dynamic approach to the verification of distributional universals. *Linguistic Typology*, 4(3):307_333.
- Nettle, D. (1999). *Linguistic Diversity*. Oxford University Press.
- Tomlin, R. S. (1986). *Basic word order: functional principles*. London: Croom Helm.

To be or not to be: A typology of existentials

Existential predications (henceforward: existentials) are a type of construction that exhibits a number of interesting properties in crosslinguistic comparison. The most striking property might well be their behavior under negation. Whereas other types of (verbal and non-verbal) predication are negated by extra morphological material (affix, particle or auxiliary) in almost every language (Miestamo 2005), negative existentials do not contain a separable negative morpheme in a large number of languages, but are expressed through a lexicalized form expressing the combined meaning of negation and existence (Dryer 2007: 246). Croft (1991: 18) suggests that this negation pattern (his type B languages) is probably the most frequent type of existential negation in the world's languages. He, however, does not give any actual figures to support this claim.

Apart from this notable negation pattern, existentials show a resemblance in coding strategies to a number of other types of intransitive and/or non-verbal predications in many languages (Hengeveld 1992, Stassen 1997). On the one hand existentials commonly share a construction with locational predicates (in which an additional locational phrase is added), the same strategy is often used for predicate possession in addition. On the other hand the coding strategy is sometimes also shared with nominal predications. A common property that existentials do not readily share with nominal predication, even if the two types of predication make use of parallel constructions, is the occurrence of zero-copulas. Croft (1991: 19) notes that he is not aware of a language that allows existentials with zero-copulas, even though Dryer (2007: 244) provides a counterexample (Tolai), the general tendency seems to hold. Still in other languages, existentials are expressed via regular verbal predicates (Hengeveld 1992: 100).

While the quoted studies of non-verbal/intransitive predications make reference to existentials, this phenomenon is not at the core of them. The earlier studies especially fall short in cases in which existentials are encoded differently from the other types of predication studied. In his study of standard negation Miestamo (2005: 44), who explicitly excludes negative existentials and other types of special negation (unless they use the standard negation pattern of a language), notes that a more comprehensive typology of clausal negation – including these special types – is a very desirable goal for future research.

The aims of this paper are twofold. Firstly, to provide quantitative information on the formal encoding (positive and negative) existentials in general and more specifically to test Croft's (1991) claim on the relative frequency of negation strategies with existentials on the basis of a probability sample of 100+ languages. Secondly, to provide a comparison of the coding strategies for existentials with respect to the coding of other types of predication (nominal, locational, possessive).

References

- Croft, William. 1991. The evolution of negation. *Journal of Linguistics* 27(1), 1–27.
- Dryer, Matthew S. 2007. Clause types. In Timothy Shopen (ed.) *Language Typology and Syntactic Description*. Vol. I: Clause Structure, 224–275. Cambridge: Cambridge University Press, second edn.
- Hengeveld, Kees. 1992. Non-verbal Predication. *Theory, Typology, Diachrony*, 15 of *Functional Grammar Series*. Berlin, New York: Mouton de Gruyter.
- Miestamo, Matti. 2005. Standard Negation. *The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin, New York: Mouton de Gruyter.
- Stassen, Leon. 1997. *Intransitive Predication*. Oxford Studies in Typology and Linguistic Theory. Oxford: Clarendon Press.

The interaction of negation and constituent order in Iquito (Zaparoan)

Negation in Iquito, an endangered Zaparoan language of the Peruvian Amazon, provides an interesting case study for the application of Dryer's (2011) typology of negative morphemes and their positioning with respect to the subject, object, and verb of a clause.

Iquito uses word order to mark the reality status of a clause. Irrealis clauses are characterized by an 'SXV' order, where the subject and verb are separated from each other by an intervening element, whereas realis clauses exhibit 'SVX' order, where the subject and verb must be contiguous. (The intervening element can be an object, determiner, adverb, postpositional phrase, or negation particle, which is why it is labeled 'X' rather than 'O'.) Word order is the sole indicator of a clause's reality status; there is no additional morphological marking associated with this grammatical category.

This word order alternation interacts with negation in an interesting way. In independent or main declarative clauses and finite complement clauses, negation is straight-forwardly marked by a negative particle (*caa*) immediately preceding the subject (which in turn precedes the verb). This strategy is by far the most common strategy for marking negation in Dryer's (2011) survey, occurring in over one-third of the 1326 languages sampled. In Iquito, it occurs in both realis and irrealis constructions. Clausal negation in interrogatives and subordinate clauses, however, is marked by an obligatory negative suffix (*-ji*).

Suffixation is the second most common strategy outlined by Dryer (2011a), occurring in over 200 languages, but only 12 languages employ both of these types.

When we look more closely at the suffixal negation strategy, we see that it can optionally co-occur with the negative particle *caa*, and when it does, the positioning of the negative particle is sensitive to whether the clause is realis or irrealis. In realis clauses, *caa* follows the verb, and in irrealis clauses, *caa* precedes the verb. As a result, we see another word order alternation within Iquito irrealis subordinate clauses: the affirmative order for these clauses is SXV and the negative order is SNegVX. Dryer (2011b) lists Hungarian as the sole instance of a language of this sort; I submit that Iquito should also be included in this list. We also see optional triple negation in these clause types (a strategy evident in only 6 languages), as it is possible for the negative particle to occur both before and after the suffixed verb in irrealis subordinate and interrogative clauses. Although Dryer excludes from his classification instances of double and triple negation that only occur in subordinate clauses, I argue that it is necessary to include this type of clausal negation in a thorough treatment of the Iquito negation system.

By examining the ways in which Iquito clausal negation interacts with realis and irrealis word orders, we are able to see how two fairly common negation strategies co-exist with several rare negation strategies and expand our understanding of the complexity of negation strategies evident within a single language.

References

- Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 143. Available online at <http://wals.info/chapter/143>.
- Dryer, Matthew S. 2011b. Position of Negative Morpheme With Respect to Subject, Object, and Verb. In: Dryer, Matthew S. and Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 144. Available online at <http://wals.info/chapter/144>.

How widespread is transitive encoding?

It is often taken for granted that languages have a large number of transitive verbs, or even that the typical two-argument verb is transitive. But we know that languages differ in the extent to which they make use of transitive encoding (e.g. Hawkins 1986 on English/German contrasts: in German, verbs like ‘help’ and ‘follow’ are not encoded transitively). Typological studies such as Tsunoda (1985) and Malchukov (2005) have tried to formulate generalizations concerning the kinds of verb meanings that tend to be coded non-transitively in different languages. But so far there has been no serious published attempt to quantify transitive encoding: How strongly do languages differ? Is English atypical in the prominence it accords to transitive encoding? There seem to be three major obstacles to such quantification: how to define transitivity crosslinguistically, how to sample verbs, and how to get systematic cross-linguistic data.

In this presentation, we report on a major study of valency patterns in 35 languages from around the world, which allows us to overcome the data obstacle: We brought together a consortium of 35 author teams (experts in their respective languages) to provide a dataset of about 80 verbs with detailed valency information. The individual datasets are comparable because they consist of counterparts to the same set of 80 basic verb meanings. (The aggregated database, called ValPaL [=“Valency Patterns Leipzig”] will be published online, so that our results can be easily verified.)

As for the two first obstacles, we basically follow the path of Greenberg (1963): We define our comparative concepts in a rigorous way, but we do not make an attempt to justify them, being content with some intuition-based decisions. This concerns, in particular, the choice of 80 verb meanings: We tried to include verb meanings of diverse kinds, which seem to us reasonably representative. As argued by Lazard (2005), intuition-based decisions are unavoidable in typology and do not detract from the methodological rigour of the enterprise. Our definition of transitivity follows Lazard (2002) and Haspelmath (2011) in spirit: We start out from the typical transitive verb ‘break’ and define “transitive encoding” as the encoding that is used by this verb. A verb is considered transitive if it contains an A and a P argument, and A and P are defined as the arguments that are coded like the ‘breaker’ and the ‘broken thing’ roles of the ‘break’ verb. We have found that this definition appears to give the same result as the criterion of “major two-argument verb class” (Witzlack-Makarevich 2011), but it is easier to apply.

Our findings are not particularly surprising: Languages differ in the extent to which they use transitive encoding in their ca. 80 sample verbs, but not dramatically: All our 35 languages have the transitive class as their major verb class, so the prominence of transitivity seems to be a robust language universal. What is perhaps most surprising is that English is not extreme in its degree of transitivity prominence.

The typology of comparative constructions revisited

Stassen (1985) was one of the first monographs dealing with a grammatical phenomenon from a world-wide perspective, at least in the modern (post-Greenbergian) era. Stassen examined comparative constructions in a sample of 110 languages and found correlations with word-order and clausecombining features. A quarter of a century later the time seems ripe for a reassessment.

In this paper, we look at a sample of 300 languages from around the world which does not overlap with Stassen's (1985) sample, nor with the larger sample (167 languages) of his 2005 *WALS* chapter. This novel set of data allows us to replicate his findings on comparative constructions and their relationship to clausal word order (cf. Plank 2007 on replication in typology), and our results offer no major surprises. The major types identified by Stassen (especially the locational construction with its subtypes, the exceed construction, and the conjoined construction) recur widely in the languages of our sample, and the geographical trends seen in Stassen (2005) are largely confirmed. Our study is the first replication of a *WALS* chapter, as far as we know, and the fact that our findings do not differ greatly can be seen as support for the sampling approach adopted in *WALS* (if such support is needed).

However, our study goes beyond Stassen's work in a number of important ways. First, our classification is more fine-grained. Thus, we distinguish between two types of exceed constructions, the primary exceed construction ("Pat exceeds Kim in tallness"), where the exceed verb is the main predicate of the construction, and the secondary exceed construction ("Pat is tall exceeding Kim"), where the parameter of comparison is expressed as the main predicate. We also subdivide Stassen's locational type into an ablative type ("tall from Kim"), a locative type ("tall at Kim"), an allative type ("tall to Kim"), and a comitative type ("tall with Kim"). Especially the latter (comitative) type does not fit well into an overall "locational" macrotype. Within the conjoined ("double predication") macrotype, we distinguish between an antonym type ("Pat is tall, Kim is short"), a negative type ("Pat is tall, Kim is not tall"), and an increase type ("Kim is tall, Pat is very tall"). Not surprisingly, most of these subtypes are associated with specific clausal word order patterns. Stassen's "particle comparative" is not defined in such a way that the concept can be applied readily to any language (it involves a particular use of case), so we use a category "other standard marker" instead.

In addition to confirming some of Stassen's findings, we have also identified a number of new universals, e.g. no language lacks both a degree marker and a standard marker ("Pat is tall Kim"), and almost no language lacks a standard marker even when a degree marker is present ("Pat is tall-er Kim"). However, we also find that there is more diversity than is apparent from Stassen's rather simple, lumping typology. Quite a few languages do not easily fit into any of the types.

Garifuna attributive possession in comparative perspective

In this paper, I outline the strategies for expressing attributive possession in Garifuna, an Arawakan language. Furthermore, by comparing the Garifuna data to other languages I will highlight some (areally) unusual features of the Garifuna system. The data underlying the research presented here come from fieldwork currently being carried out by the author in Garifuna communities of Northern Honduras, Central America.

Attributive possession in Garifuna makes a distinction between alienable and inalienable possessed items, the latter being restricted to kinship terms and body parts. Furthermore, some nouns take different classifiers, here referred to as relational classifiers following Lichtenberk (1983) depending on the use for which they are intended by the possessor: either, "X's Y for eating", "X's Y (meat) for eating", "X's Y for drinking", "X's Y for keeping as a pet" or "X's Y (for general possession)". Relational classification is widely attested in Oceanic Languages (Lichtenberk 1985:106) but it is a rare feature in The Americas, even within Arawakan, and is possibly borrowed from Cariban (Aikhenvald 2013:46).

According to Grinevald (2000:81) the Amerindian relational classifiers are emergent, i.e. less grammaticalized than the Oceanic ones, because the choice of relational classifier depends on discourse context rather than being fixed in the grammar. However, since Grinevald offers no documentation for this claim, I will compare Oceanic, Cariban and Garifuna examples. In (1-2) there seems to be no difference between the way relational classifiers are used in Bau and Garifuna respectively.

(1) Garifuna (my field notes)

'n-eygã 'faluma
1SG.POSS-CLF coconut
'my coconut (to eat)'

nu-'niye 'faluma
1SG.POSS-CLF coconut
'my coconut (to drink)'

'n-ani 'faluma
1SG.POSS-CLF coconut
'my coconut (to dispose of as I please)'

(2) Bau (Oceanic, Austronesian; (Pawley 1973:168)

na ke-na maqo
ART POSS-his mango
'his mango for eating (i.e. green mango)'

na me-na maqo
ART POSS-his mango
'his mango for sucking (i.e. ripe, juicy mango)'

na no-na maqo
ART POSS-his mango
'his mango (as property, e.g., which he is selling)'

Fijian, another Oceanic language, has relational classifiers for food and drink which are only used if these items are intended by the possessor to be consumed immediately; in other contexts a classifier

for general possession is used (Dixon 1989:136); thus, Fijian also seems to work in essentially the same way as Garifuna.

References:

- Aikhenvald, Alexandra Y. 2013 Possession and Ownership: a Cross-linguistic Perspective. *In* Possession and Ownership. Alexandra Y. Aikhenvald and R. M. W. Dixon, eds. Oxford University Press, USA.
- Dixon, R. M. W. 1989 A Grammar of Boumaa Fijian. 1st edition. University Of Chicago Press.
- Grinevald, Colette 2000 Systems of Nominal Classification. Gunter Senft, ed. Language, Culture, and Cognition, 4. Cambridge ; New York: Cambridge University Press.
- Lichtenberk, Frantisek 1983 Relational Classifiers. *Lingua* 60: 147–176. 1985 Austronesian Linguistics at the 15th Pacific Science Congress. Andrew Pawley and Lois Carrington, eds. Pacific Linguistics, no. 88. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.
- Pawley, Andrew 1973 Some Problems in Proto Oceanic Grammar. *Oceanic Linguistics* 12: 103–188.

Valence in a typological and theoretical perspective

We juxtapose valence-related construction type inventories of the Kwa language Ga (spoken in Ghana) and the Germanic language Norwegian, with three aims: (i) to display how pervasively different these inventories are; (ii) to identify their main differentiating factors; and (iii) to illustrate a methodology for conducting (i) and (ii).

The methodology resides in the ‘Construction Labeling’ system, a notation system for verb constructions and verb valence, proposed in Hellan and Dakubu (2010) – see http://www.typecraft.org/w/images/d/db/1_Introlabels_SLAVOB-final.pdf, so far used in establishing fairly large-scale construction inventories for a few languages from Germanic, Niger-Congo and Ethio-Semitic. The system is based on a cross-linguistically grounded repertoire of *properties of linguistic constructions*, such as, e.g., ‘has Valence Frame X’, ‘has Aspect Y’, ‘has a Subject with properties Z’, ‘expresses situation type S’, etc. Each such property is packaged in the notational code as an atomic *element*, construction types are represented through combinations – called *templates* – of such elements, and lists of templates constitute *c(onstruction)-profiles* of a language. Below are two examples of the code applied to Ga constructions, (a) a ditransitive construction and (b) a serial verb construction, both with standard morphological glossing. In the former case, the element **v** indicates that the construction is headed by a verb, **ditr** indicates that the argument frame is syntactically ditransitive, **suAg** means that the subject has the semantic role of ‘agent, and so on, **COMMUNICATION**

finally indicating the situation type expressed. In the latter case both verbs occur with an expressed object; their subjects are identical, and likewise their aspects, expressed in the code element **svSuAspIDALL**.

a. v-ditr-suAg_iobTrgt_obThmover-COMMUNICATION

E-f	mi	nine
3S.AOR-throw	1S	hand
V	Pron	N

‘She waved to me; invited me.’

b. svSuAspIDALL-v1tr-v2tr

Á-gbele	gb	á-ha	bo
3.PRF-open	road	3.PRF-give	2S
V	N	V	Pron

‘You have been granted permission.’

C-profiles of the two languages are to be found on the following sites:

http://www.typecraft.org/w/images/a/a0/2_Ga_appendix_SLAVOB-final.pdf,

http://www.typecraft.org/w/images/b/bd/3_Norwegian_Appendix_plus_3_SLAVOB-final.pdf

As can be verified, the number of shared templates constitutes less than 10% of either of the profiles. The typological interest lies in identifying elements characteristic of those templates which are specific to either language, and in turn to their language types. The methodology is innovative in enabling such an investigation in a more efficient way than has been so far possible. The methodology offers a specification space within which ‘alternation’-based approaches to valence can be grounded (cf. Levin (1993) and its computational extension *VerbNet* as regards single-language investigations, and the *Leipzig Valency Classes Project* as regards cross-linguistic investigations), but allows in principle for contrastive valence studies *not* based on frame alternations.

References

- Hellan, L. and M. E. K. Dakubu. 2010: *Identifying verb constructions cross linguistically*. Studies in the Languages of the Volta Basin 6. 3. Accra.
- Leipzig Valency Classes Project: <http://www.eva.mpg.de/lingua/valency/index.php>
- Levin, B. 1993. *English Verb Classes and Alternations*, University of Chicago Press, Chicago IL
- VerbNet: <http://verbs.colorado.edu/~mpalmer/projects/verbnet/downloads.html>

Morphosyntactic coding of proper names and its implications for the Empathy Hierarchy

The Empathy Hierarchy (EH) is one of the most important generalizations in linguistic typology employed for the description and explanation of the typological distribution of different morphosyntactic phenomena in the domain of case marking and agreement in the languages of the world. The different names assigned to the EH in the literature reflects different functional interpretation: 'lexical hierarchy' (Silverstein 1976), 'Nominal Hierarchy' (Dixon 1994) 'animacy hierarchy' (Comrie 1981) 'empathy hierarchy' (Kuno & Kaburaki 1977; DeLancey 1981), 'hierarchy of reference' (Zwicky 1977), and 'prominence hierarchy' (Aissen 1999). The EH is a scale of different classes of referential expressions (1/2 > 3 > **proper names/ kin terms** > human > non-human > inanimate common nouns; cf. Dixon 1979) stating that the speaker is more likely to take over the perspective of a referent that is higher on the EH. Proper names are claimed to occupy an intermediate place between personal pronouns and common nouns. Despite the large body of research on the EH since its first extensive formulation in Silverstein (1976), it is astonishing to discover that there is almost no empirical evidence for this claim. Silverstein (1976), for instance, discusses a couple of Australian languages with split ergativity marking, but does not give a single example to demonstrate the position of proper names on the EH. The same lack of evidence can be found in Blake (1994) and almost all publication that deal with the EH in one way or other; even the articles in the recently published *Oxford Handbook of Case* (Malchukov & Spencer (eds.) 2009) ignore case marking of proper names entirely. Some of the examples given in Dixon (1994) could be considered as evidence, but are often inconclusive or even contradictory to this hypothesis.

The goal of our proposed talk is to give an answer to the question whether the morphosyntactic coding of proper names in the languages of the world confirms or falsifies their hypothesized position within the EH.

In the first part of our talk we will give a very brief evaluation on the state of the art of research on the EH and the kind of data that were presented in the literature in favor of the EH in general and with regard to proper names in particular. We will further discuss the question what kind of data are in principle required in order to prove (or disprove) that proper names are correctly positioned between pronouns and nouns in the EH. Some remarks on language sampling and methods of analysis will follow: in order to find the data that we need to answer the research question, we compiled a probability sample selecting languages that have an ergative split marking system and languages that have a hierarchical marking system. For each language of the sample we looked whether proper names pattern with personal pronouns, or with common nouns (or sub-categories of them), or with neither in coding the core grammatical relations.

In the second part we will present the results together with consequences that have to be drawn for the form of the EH. Up to now, our findings do not allow drawing firm conclusions. There are very few clear data in favor of the hypothesis, and there are also some data that thoroughly contradict the hypothesis. Most data are simply inconclusive. A clearer picture will be given in our talk which will be based then on a much broader empirical basis.

References

- Aissen, Judith. 1999. Agent Focus and Inverse in Tzotzil. *Language* 75:451-485.
Blake, Berry J. 1994. *Case*. Cambridge: Cambridge University Press.
Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

- DeLancey, Scott. 1981. An Interpretation of Split Ergativity and Related Patterns. *Language* 57:626-657.
- Dixon, R.M.W. 1979. Ergativity. In: *Language* 55:59-138.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Kuno, Susumo & E. Kaburaki 1977. Empathy and Syntax. In: *Linguistic Inquiry* 8:627-72.
- Malchukov, Andrej & Andrew Spencer (eds.) 2009. *The Oxford Handbook of Case*. Oxford: Oxford University Press.
- Silverstein, Michael 1976. Hierarchy of Features and Ergativity. In: R.M.W. Dixon (ed.) *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112-71.
- Zwicky, Arnold 1977. Hierarchies of Person. In: *Papers from the Thirteenth Regional Meeting of the Chicago Linguistic Society*, pp. 712-33, Chicago: Chicago Linguistic Society.

When is there agreement? Typologizing restrictions on agreement

In some languages, bound person forms on the verb present difficulties when one attempts to classify a language as either showing agreement or not for the purposes of typological investigations.

The first type of problems arises from the fact that in some languages bound pronominals can be analyzed as pronouns (and thus arguments) and not as agreement markers. According to an influential line of research, pronouns are distinguished from agreement markers on the basis of the co-occurrence restriction: (i) if the co-occurrence of two argument expressions is possible, then one of them is an argument and the other one is an agreement marker (grammatical agreement); (ii) if the co-occurrence is impossible, the pronominal markers are considered to be arguments themselves and thus pronouns (pronominal agreement) (see Bresnan & Mchombo 1987, Siewierska 1999, Bickel & Nichols 2007). Notably, this co-occurrence restriction is not general but often concerns a specific phrase-structural position reserved for true arguments. This diagnostics is orthogonal to the question whether an NP occurs at all in the clause, as in most languages, NPs are optional in all positions, regardless of whether the language has grammatical or pronominal agreement.

The second type of problems concerns the instances of restricted (or “optional”) agreement, illustrated with Mixtec subject agreement in (1). Whereas in certain contexts the bound pronominal markers are obligatory (1a, 1c), they are banned in other contexts (1b, 1d). The distribution of the agreement markers in this and similar cases has been accounted for in terms of language-specific constraints mostly formulated with reference to phrase-structural position, intonation or information structure.

Though this phenomenon is quite well-spread in the languages of the world, we are not aware of any attempt to typologize such constraints on agreement. In this study we will consider an areally-balanced sample of 50 languages with restricted agreement. For every language we investigate the restrictions on agreement and develop a typology of structural positions necessary to represent the observed variation.

Examples

- (1) Mixtec (Chalcatongo) (Macaulay 1996:139ff.)
- | | |
|--|--|
| a. <i>ni-žee=rí</i>
COMP-eat=1sS/A
'I ate (it).' | b. <i>rùḡù ni-žee</i>
I COMP-eat
'I'm the one who ate (it).' |
| c. <i>rùḡù ni-žee=rí</i>
I COMP-eat=1sS/A
'As for me, I ate (it).' | d. <i>*ni-žee(=rí) rùḡù</i>
COMP-eat=1sS/A I
'I ate (it).' |

References

- Bickel, Balthasar & Johanna Nichols. 2007 Inflectional morphology. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 3, Cambridge: Cambridge University Press 2nd edn.
- Bresnan, Joan & Sam A. Mchombo. 1987 Topic, pronoun, and agreement in Chichewa. *Language* 63
- Macaulay, Monica. 1996 *A Grammar of Chalcatongo Mixtec*, vol. 127 University of California Publications in Linguistics. Berkeley: University of California Press.
- Siewierska, Anna. 1999. From anaphoric pronoun to grammatical agreement marker: why objects don't make it. *Folia Linguistica* 33(2). 226–251.

Theme session: Typological hierarchies in synchrony and diachrony

The decay of direct/inverse systems

Two general diachronic causes can be proposed to explain the relative rarity of prototypical direct/inverse systems in the world's languages.

First, the pathways of grammaticalization leading to the creation of direct/inverse systems might be complex and unusual, in which case few such systems would be created to begin with.

Second, these systems could be unstable, and hence subject to decay, optionally giving rise to a new type of agreement. Since most languages with direct/inverse systems are either isolates (Mapudungu, Movima, Kutenai) or very small families (Sahaptian, Shastan, Kiowa-Tanoan), the diachronic stability of these systems is difficult to assess. Among larger families, in which diachronic hypotheses are easier to evaluate, only two have direct/inverse systems: Algic and Sino-Tibetan.

In the first one, direct/inverse systems appear to be very stable, as even very innovative languages, such as Arapaho, preserve them fully. In Sino-Tibetan, on the other hand, there is no consensus as to whether the direct/inverse systems observed in some branches are ancient (as proposed by DeLancey (1981)) or innovative. Even in the subgroups where prototypical direct/inverse systems are attested, as in Rgyalrongic, not all languages share this feature. Sino-Tibetan seems thus to be the ideal testground for studying both the development and the dissolution of direct/inverse systems. The presentation will accordingly focus on the issue of decay of direct/inverse systems in two subgroups of Sino-Tibetan: Rgyalrongic and Kiranti.

First, based on first-hand data on two Rgyalrongic languages (Japhug and Resnyeske), we show how a formerly pristine direct/inverse system, attested in the central Rgyalrong languages (see Sun and Shidanluo (2002), Jacques (2010), Gong (to appear)) has changed into opaque and partly hierarchical systems in the neighbouring Lavrung and Rtau languages. The most unexpected outcome of this research is the discovery that when the direct/inverse contrast is lost in 3>3 forms, it is always the inverse form that is preserved and the direct one that is lost (Jacques (2012)).

Second, we show that some Kiranti languages have apparently undergone a similar pathway of evolution as Rtau and Lavrung, and propose some elements of reconstruction of proto-Kiranti morphology.

By presenting some possible pathways of decay for direct/inverse systems this study shows that opaque hierarchical systems may in some cases have evolved from prototypical direct/inverse systems which have lost their synchronic motivation due to a combination of phonetic change and analogy.

References

- DeLancey, Scott. 1981. *The category of direction in Tibeto-Burman*. *Linguistics of the Tibeto-Burman Area* 6.1:83–101.
- Gong, Xun. to appear. *Personal agreement system of Zbu rGyalrong (Ngyaltsu variety)*. *Transactions of the Philological Society*.
- Jacques, Guillaume. 2010. *The Inverse in Japhug Rgyalrong*. *Language and Linguistics* 11.1:127–157.
- Jacques, Guillaume. 2012. *Agreement morphology: the case of Rgyalrongic and Kiranti*. *Language and Linguistics* 13.1:83–116.
- Sun, Jackson T.-S., and Shidanluo. 2002. *Caodeng Jiarongyu yu rentong dengdi xiangguan de yufa xianxiang 草登嘉戎語與「認同等第」相關的語法現象 (Empathy Hierarchy in Caodeng rGyalrong grammar)*. *Language and Linguistics* 3.1:79–99.

Theme session: Predicate-centered focus types

'Do'-periphrasis as a cross-linguistic predicate focus strategy

Insertion of an auxiliary equivalent to English *do* into a sentence is a cross-linguistically common strategy for the expression of discourse functions such as predicate focus and topicalization. The sentences in (1) exemplify this. They are marked and contrast with non-periphrastic declaratives that do not highlight the predicate or parts thereof. Based on a sample of 200 languages I will show that this is in fact one of the major functional types of 'do'-periphrasis. It is argued that by virtue of their schematicity 'do'-auxiliaries lend themselves to pragmatic purposes and make this a likely strategy independent of genetic affiliation.

- (1) a. English (Indo-European):
Watch a film he did.
- b. Gude (Afro-Asiatic): [HOSKISON 1975: 228-229]
bələnə nə sətə ci John ada tə bwaya.
kill SUBJUNCTIVE thing CONTINUOUS John do OBJECT leopard
'John is KILLING a leopard now.'
- c. Fon (Niger-Congo): [LEFEBVRE 1991: 40-41]
asɔ_ sɔ yi axi-mɛ! wɛ! kɔ_ ku ɖɛ.
crab take go market-LOCATIVE PROGRESSIVE Koku do
'It is bringing a crab to the market that Koku is doing.'
- d. Korean (Isolate): [HAGSTROM 1995: 32-33]
Chelswu-ka chayk-ul ilkki-nun ha-ess-ta.
Ch.-NOMINATIVE book-ACCUSATIVE read-TOPIC do-PAST-DECLARATIVE
'Read the book, Chelswu did.'

Languages with rigid word order often use 'do'-periphrasis to mark non-canonical clause types that display a deviant or irregular word order. The strategy maintains a close approximation of the regular word order, i.e. it upholds the relative order of verb and object. Functionally such clause types are strongly discourse dependent. If the change of canonical word order makes periphrasis obligatory, the resulting periphrasis appears grammatically conditioned, i.e. retaining canonical word order as its chief motivation. 'Do'-periphrasis, however, likewise occurs in languages with relatively free word order. Here the same form-function-relations apply. This suggests a crosslinguistic tendency to associate the aforementioned pragmatic functions directly with 'do'-periphrasis, where degrees of optionality indicate different stages of grammaticalization. That is to say that language A employs periphrasis optionally in contexts that are functionally similar to the contexts that make periphrasis obligatory in language B.

References

- Hagstrom, P. 1995. *Negation, focus and do-support in Korean*. Ms., MIT. (available from <http://www.bu.edu/linguistics/UG/hagstrom/papers/NegFocDo.pdf> (accessed January 2013))
- Hoskison, J.T. 1975. *Focus and Topic in Gude*. In: Herbert, R.K. (ed.) *Proceedings of the Sixth Conference on African Linguistics*, Ohio State University, Columbus, April 12-13, 1975. Columbus: Ohio State University, 227-233.
- Lefebvre, C. 1991. *Take serial constructions in Fon*. In: Lefebvre, C. (ed.) *Serial verbs: grammatical, comparative and cognitive approaches*. Amsterdam: John Benjamins, 37-78.

Functional Explanations for Referential Hierarchy Effects on Grammar

The growing documentation and analysis of American indigenous and other understudied languages has revealed several unique grammatical systems based on referential hierarchies, some of which overtly express event direction, triggering a recent surge in typological work on the topic (Bickel 2008a, Richards and Malchukov (eds) 2008, Zavala 2007, Zúñiga 2006, 2008). However, it is still unclear whether hierarchical systems should be treated as an alignment type in its own right (Nichols 1992, Siewierska 2005, Zúñiga 2006), viewed in terms of voice (Givón 1994, Shibatani 2006), or analyzed based on the properties of individual systems (Bickel 2008a). In this paper I examine referential hierarchy effects on grammatical marking in 40 languages. My aim is to show that all hierarchical systems can be explained in terms of subjectivity, politeness, and topicality, and that in different languages these functions are fulfilled in structurally distinct ways conditioned by genetic inheritance and contact-induced change, as proposed by Bickel (2008a), consequently supporting the idiosyncratic approach.

Previous studies have focused either on a specific subcategory of hierarchical systems, i.e. inversion (Klaiman 1992, Zavala 2007, Zúñiga 2006, 2008) or obviation (Aissen 1997, Dryer 1992), on a particular structural correlate (Bickel 2008b), on the origin of hierarchy effects (Mithun 2010, in press), or on individual systems. This work attempts to consolidate and expand these studies by examining a large number of languages. The following parameters are analyzed for each language: (a) domain (i.e. involving speech-act participants or not), (b) locus of marking, (c) type and presence of person marking, (d) presence of event direction marking, (e) alignment type, (f) presence of obviation and specific obviation triggers, and (g) rankings in individual hierarchies.

The results reveal vast formal variability among the languages studied and, therefore, a difficulty for categorizing. Nevertheless, strong similarities are apparent within language families (e.g. Algonquian, Mixe-Zoquean, Sahaptian) and in linguistic areas (e.g. California), thus corroborating Bickel's (2008a) claim that genetic and areal reasons rather than universals account for structural patterns in referential hierarchies. Most languages exhibit ergative or mixed alignment, although hierarchy effects are irrelevant in intransitive clauses. Inverse languages often use different markers for A and O, in addition to marking event direction, but not in all scenarios. Languages labeled as hierarchical (but not inverse), generally leave third persons unmarked and do not show obviation, but they present mechanisms similar to inverse marking (i.e. passives) where inverse would be expected (e.g. Yana, Yurok), hence equally marking event direction. Whereas rankings generally follow the animacy hierarchy (Silverstein 1976), the ranking of speech-act participants exhibits variability. Systems where first persons always surface can be related to subjectivity (Scheibman 2002), while higher ranking second persons can be associated with politeness (Mithun 2008). All other rankings are based on topicality with higher ranked participants being more topical.

Overall, the analysis of referential hierarchy effects in 40 languages provides evidence of great structural variability linked to genetic and areal sources therefore favoring an idiosyncratic structural and a functional approach and arguing against hierarchical systems as a separate alignment type.

Transitive directionals in Mon – form, function, and implications for linguistic typology

Directionals are used to indicate absolute or relative directions in a verb complex in Mon. The directionals form a closed set of verbal morphemes, consisting of ‘movement away from origo’, ‘movement towards origo’, ‘movement up’, ‘movement down’, ‘movement in’, ‘movement out’ and ‘movement back to point of origin’. All directionals appear in two forms, basic/intransitive and causative/transitive. The causative/transitive forms are in either morphological causatives or suppletive forms. The choice of the form of the directional employed depends on the movement or affectedness of the participants of an expression. If the S/A argument is described as moving by the main verbal predicate, the basic form of the directional is used, as in (1). If the P (or T) argument is set in motion, the causative form of the directional is obligatorily used, as in example (2). In transitive expressions, the basic form is used if the A rather than the P argument is set in motion, or if the setting in motion of P is backgrounded. As seen in examples (3) and (4), the same main verb may combine with either the basic or the causative directional. In the former case, it is the movement of the A argument that is important, while in the latter it is the movement of P. In ditransitive expressions, the causative directional refers to the movement of the T, never the G argument. The main trigger for the choice of the directional is apparently the “affectedness of the O argument” (Hopper & Thompson 1980). This systematic distinction between basic and causative directionals, which is rare not only in Southeast Asian languages, but also globally, allows a distinction in the degree of (semantic) transitivity of an event based on the linguistic expression. It can be shown, for example, that morphological causatives in Mon have a higher degree of transitivity than periphrastic causatives, as only the former trigger the causative directionals. The findings of this study are therefore relevant in a broader typological context, both regarding transitivity parameters and causative constructions.

The present study investigates the different uses of causative/transitive directionals in Mon and the functional differences between the basic and causative forms. Dealing with a typologically rare phenomenon, this study adds to our understanding of complex verbal predicates and transitivity not only in the Southeast Asia context, but also crosslinguistically. The study is based on original data collected in Thailand and Myanmar from different varieties of Mon, supplemented by published texts such as journal articles and short stories.

Examples

(1)

kon.ηàc kwac cao ʔa phèə.
 child walk return go school.
 ‘The child walked back to school.’

(2)

*ràə kok phyao na hɔəʔ. (*kok cao ʔa)*
 friend call CAUS.return CAUS.go house
 ‘The friend brought (her) back home.’

(3)

dəh pək ʔa ka.
 3 follow go car
 ‘He is driving the car.’

(4)

dəh pək na klèə.
 3 follow CAUS.go cow
 ‘He is driving the cattle.’

References:

- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 251-99.
- Jenny, Mathias. 2005. *The verb system of Mon*. Zurich: ASAS.
- Kittilä, Seppo. 2002. *Transitivity: towards a comprehensive typology*. Turku: University of Turku.
- Næss, Ashild. 2007. *Prototypical transitivity*. Amsterdam: John Benjamins.

Theme session: Lexical typology of qualitative concepts

Adjectives describing surface texture: towards lexical typology

This paper deals with adjectives describing surface texture ('slippery', 'smooth', 'level', 'rough', etc.). The language sample comprises Russian, English, Chinese, Spanish, Korean, and a set of the Uralic languages (Finnish, Estonian, Erzya, Mari, Komi, Udmurt, Hungarian, Khanty, Nenets, Selkup). Their in-depth study was aimed at exploring the dependence between the genetic proximity of languages and the similarity of their lexical systems in the domain concerned.

Adjectives referring to **absence of roughness** are basically opposed by the way a surface is perceived, which may be visual (the prototype is a level field) or tactile (the prototype is a stone slipping out of one's hands). The latter comprises slippery surfaces and also smooth surfaces, like a well shaven wooden board. Languages adopt different strategies here (Russian *skol'zkiĭ* 'slippery', *gladkiĭ* 'smooth', *rovnyiĭ* 'level' vs. Erzya *nolaža* 'slippery', *valan'a* 'smooth, level' vs. Shuryshkary Khanty *wǎł'ǫk* 'smooth, slippery', *pajli* 'level'). However, no system opposition of lexeme 'slippery, level' vs. lexeme 'smooth' has been attested, neither is there a system with one lexeme dominant over all the frames.

Slippery surfaces are further divided into those one walks on and those of the objects dropping out of one's hands. Among smooth surfaces, the surface of body parts is sometimes categorized as a special frame. Besides, adjectives meaning 'smooth' often have a secondary visual feature, conveying the idea of shining or glittering (which presents interest for cognitive studies, cf. [Viberg 1984]). The subdomain of level surfaces opposes artifacts (sometimes differentiated by their horizontal vs. vertical orientation) and landscapes (vast areas, roads, intentionally levelled places, water surfaces).

Adjectives denoting **roughness** also distinguish between surfaces perceived visually vs. by touch. The former subdomain includes many items with a narrow meaning ('hilly', 'potholed', etc.), but also tends to specify a broader class of wrinkled surfaces. As regards the subdomain of tactile perception, what is consistently brought out is the frame of surfaces with regular rigid roughness perceived by touch, further opposed with the size of roughness, the rigidity/flexibility of an object, and the effect on a contacted object (Udmurt *tšogyr'es* 'rough and scratching').

The **metaphoric uses** of the surface texture lexemes show typologically consistent patterns: (1) 'slippery' → unsteadiness; (2) 'smooth' → absence of defects or difficulties; (3) 'level' → regularity, uniformity; (4) 'rough' → defects, difficulties.

Along with the typological data, our research provides more general **theoretical implications**. Firstly, the two antonymic semantic zones (roughness vs. absence of roughness) are structured according to different patterns. This contributes to the study of the asymmetry shown by antonyms in their semantics and combinability (see [Apresjan 1995]), which has not been systematically investigated from a typological perspective.

Secondly, our study has proved the benefits of including genetically close languages into a lexical typology research. As argued in [Kibrik 1998] with respect to grammatical typology, studying closely related languages shows many subtle typological distinctions. As follows from our data obtained from the Uralic languages, the same holds true for lexical typology. Moreover, working with the Uralic material has enabled us to establish the general structure of the domain in question, as well as the basic polysemy patterns – which all have proved to be present outside the Uralic family.

References

- Apresjan Ju. D. Leksicheskaĭa semantika: sinonimicheskie sredstva jazyka [Lexical semantics: synonymic tools of language]. Moscow, 1995.
- Kibrik A. E. Does intragenetic typology make sense? // Boeder W. et al. (eds.) Sprache in Raum und Zeit: Beiträge zur empirischen Sprachwissenschaft. Tübingen, 1998, pp. 61-68.
- Viberg A. The verbs of perception: a typological study // Butterworth B. et al. (eds.) Explanations for language universals. Berlin, 1984, pp. 123-162.

The distinction between Exclusive and Shared Possession in Kupsapiny

This study shows that there is a dimension to the distinction of possession made by a language similar to but different from the one in terms of inalienability (Nichols 1988, Chappell & McGregor 1996) by demonstrating that the distinction made with the absence or presence of a clitic in Kupsapiny, the Southern Nilotic language of Uganda, is based on whether the possessum is possessed only by the possessor or is shared by (an)other possessor(s). It also shows that although this distinction is semantically different from the inalienability distinction made in various languages, it shares properties in common with it, and that these two distinctions could be subsumed under a more general possessive distinction.

A number of Nilotic languages have been reported to make the distinction between inalienable and alienable possession morphologically: the juxtaposition of possessum and possessor nouns for inalienable possession vs. the use of the possessive suffix on the possessum noun for alienable possession (e.g. Tucker & Bryan 1966). However, Kupsapiny does not make the inalienability distinction. It does not use the juxtaposition of possessum and possessor nouns to express possession, though it uses the possessive suffix for possession. The distinction that this language makes is between exclusive and shared possession. On the other hand, according to Heine (1997), there are cross-linguistic variations as to exactly what entities count as inalienably or alienably possessed. He argues that inalienability is a morphosyntactic entity, which is difficult to define semantically, and is characterized in terms of such properties as (i) unmarkedness, (ii) less heavy morphological marking, (iii) a historically older construction, (iv) head-marking morphology, and (v) a closed category of possessums. The question is whether the exclusive vs. shared possession distinction in Kupsapiny, which appears to overlap with the inalienability distinction, has any similarity with it.

When the possessor is human, and is expressed with a full noun phrase, Kupsapiny makes a distinction between exclusive and shared possession with the use of the possessive suffix *-ap* alone and the use of the enclitic *=mpo* ‘*lit. also, additionally, even*’ in addition to the possessive suffix *-ap*, as in (1a) and (1b), respectively, which can both be used as an answer to the question ‘What is that?’.

- (1) a. *oteliit-ap* *ceepet* ‘the hotel that Ceepet alone owns’
 b. *oteliit-ap=mpo* *ceepet* ‘the hotel that Ceepet shares with someone (for example, the hotel that Ceepet and someone own; the hotel where Ceepet is staying/working)’

There are also possessums for which the exclusive possession construction has to be used because they are always exclusively possessed, independent of context. Such possessums include body parts (e.g. ‘eye’, ‘nail’), the exclusively possessed kin in Kupsapiny speakers’ culture (‘wife’), exclusively possessed artifacts (e.g. ‘eyeglasses’, ‘grave’), and actions that are not conducted with another person (e.g. ‘sneeze’, ‘yawn’). Thus, although this distinction overlaps with the inalienability distinction, these distinctions are semantically different.

Nevertheless, the distinction between exclusive and shared possession shares properties in common with that between inalienable and alienable possession. First, exclusive possession is less marked than shared possession. Second, the possession suffix *-ap* is shorter than the shared possession marker complex *-ap=mpo*. Third, unlike the exclusive possession construction with the suffix *-ap*, which also exists in other Nilotic languages, the shared possession construction has not reportedly been found in other Nilotic languages, and could hypothetically be a new addition to Kupsapiny. Fourth, the possessive suffix attaches to possessum nouns (head nouns), rather than possessor nouns (dependent nouns). Finally, although, unlike inalienably possessed nouns, exclusively possessed nouns do not form a more closed category than nouns for shared possessums, possessums

that are only exclusively possessable form a closed category. Therefore, despite their semantic difference, the distinction between exclusive and shared possession and that between inalienable and alienable possession have properties in common, and could be subsumed under a superordinate possessive distinction.

In sum, Kupsapiny makes the distinction between exclusive and shared possession with the absence and presence of the clitic =mpo. Although this distinction is semantically different from the inalienability distinction, it shares several properties in common with it.

Theme session: Generalized Noun Modifying Clause Constructions

Between correlatives, internally headed clauses and cleft: relative clauses in Mano

In this paper I will discuss specific typological properties of relative clauses in Mano, Southeastern Mande. First, relative clauses in Mano belong to the rare typological type discussed in [Nikitina 2012], that is, "a non-reduction relative clause appearing before a pronominal head, inside the main clause". Such a type "is denied, either implicitly or explicitly, in theoretical work on non-reduction relativization and, more specifically, on constructions with correlative clauses" [ibid.].

mī nófé lé i nàà à ká kō bà
 person any REL 2SG.DIPFV love:IPFV 3SG.NSBJ with 1PL.NSBJ in

'Take any of us whom you like!'

As typical for other Mande languages, relative clauses may as well be used in the topicalized sentence-initial position.

Mēī là lóó lé à dō ɔ̄ lē
 Mei 3SG.POSS trade REL 3SG.DIPFV>3SG stop:IPFV TOP 3SG.EXI

dà là lòò ká.
 fall on trade:IZF with

'Mey's trade that he implements is exploitative'.

However, there are several features that distinguish Mano relative clauses from those in other Southeastern Mande.

1. Morphology

As in Kla-Dan and Dan-Gweetaa, such clauses are marked by a post-nominal relativizer and an optional clause-final marker related to a topic marker [Vydrin 2008]. Only in Mano, at least in the Kpeinson dialect, the relative marker grammaticalizes into a floating high tone which is added to a preceding word.

gbá gbùò lé m̄ gè... --> gbá gbùó
 dog big REL 1SG.PRET>3SG see

'The big dog that I saw...'

2. Syntax

2.1. Resumptive pronoun

The resumptive pronoun in the main clause which immediately follows the relative clause is optional; in the following example relative clauses approach to the internally-headed relativization strategy which, being less autonomous, may be the initial point of the grammaticalization.

l̄ē sè í ī sèkēè ʔō gbū m̀ lé
 3SG.EXI good 2SG.CONJ 2SG.NSBJ gratitude take.off gather on REL

è lō à nō -ɔ̀ ī l̄ē ɲwéì.
 3SG.DIPFV go:IPFV 3SG.NSBJ give-GER 2SG.NSBJ for behind

'You should thank your brother for the help he will bring!'

2.2. Combinability

Like in Dan, as opposed to other Southwestern Mande languages, there are no syntactic restrictions on the position of the relative clause: they are attested not only in front of postpositional phrases (like in Wan and Kla-Dan), but also in front of verb phrases (like in Tura and Dan-Gweetaa), which shows a higher degree of grammaticalization.

2.3. Cleft

The cleft construction is formed exactly the same way as relative clause. It may be interpreted as a relative clause acquiring autonomy in the sentence-initial position.

À légbú lé mā s̀l̀òʔō ā.
 3SG.NSBJ rest REL 1SG.PRET>3SG get TOP

'What I got was only the remains!'

2.4. Appositive function

Relative clauses can also be used postpositively, in an appositive function, which is also the case of the further development of the autonomy of the relative clause.

ŃŃ nàà í ʔū kpàà lé wèŋ d̀ò
 1SG.IPFV love:IPFV 2SG.CONJ rice cook REL salt INDEF

wáá m̀ ɔ̀ .
 NEG.COP>3SG on TOP

'I want you to cook the rice without salt (so that there is no salt)!'

We can conclude that relative clause in Mano shares typological properties of rare relativization strategies of Southwest Mande, that is, the clause-internal correlative. In addition to that, it has several peculiar usages covering a large segment of the grammaticalization path.

Some abbreviations

DIPFV - dependent imperfective, GER - gerund, INDEF - indefinite article, IPFV - imperfective, IZF - izafet, NSBJ - non-subjective, POSS - possessive, PRET - preterit, REL - relativizer, TOP - topicalizer.

References

- Lipták, Anikó. 2009. The landscape of correlatives: An empirical and analytical survey. In: Lipták, A. (Ed.) *Correlatives Cross-Linguistically*. John Benjamins, Amsterdam, pp. 1-46.
- Nikitina, Tatiana. 2012. Clause-internal correlatives in Southeastern Mande: A case for the propagation of typological rara *Lingua*, Volume 122, Issue 4, pp. 319–334.
- Vydrin, Valentin. 2008. Strategies of relativization in the Mande languages: the case of Dan-Gweetaa and Bamana. *The African Collection - 2007*. St. Petersburg: Nauka, pp. 320-330.

Perfect polysemy in Enets

The paper deals with the Enets (Samoyedic, Uralic; North of Siberia) verbal TAM marker *-bi-* which shows a cross-linguistically non-typical polysemy.

In the majority of its uses, this marker, that we label Perfect, has a more or less crosslinguistically expected polysemy. It can be used indeed as a perfect (to express a past event with a current relevance, uncluding a remaining result), as an inferential (to express an event inferred by the speaker from its results), or as a mirative (to express an event that the speaker perceives as unpredicted or surprising). Along with actual mirative uses, it also marks mirative clauses in a narrative, mainly foreground events that are pragmatically unpredicted. However, there is also one more narrative use of this marker that is not in lines with the standard perfect-mirative polysemy. *-bi-* is also used in introductory clauses like (1).

- (1) ηoꞑ nꞑ dꞑiri-bi
 one woman live(ipfv)-PRF.3SG.S
 ‘There lived a woman’.

We aim to present the data on this phenomenon in order to discuss it in the crosslinguistic context looking for any parallels in genetically or areally related and unrelated languages. Through such context we will also answer the following questions. First, are the uses like (1) really exceptional for a perfect-mirative or do they represent one of cross-linguistically valid possibilities? Second, are the uses like (1) recently acquired by a perfect-mirative or are they relics of a gram that the perfect-mirative comes from?

Use of evidential markers in declaratives and interrogatives

Languages display obvious differences in how evidential markers are used in declaratives and interrogatives. First of all, languages may be divided into two based on whether or not evidentials may appear in both declaratives and interrogatives. Languages in which evidential markers are confined to declaratives constitute the first type (English, Finnish Estonian). The type in which evidentials appear only in interrogatives is not attested in the languages I have data for. The second main type is illustrated by languages in which evidentials are attested in both declaratives and interrogatives. This type can further be subdivided according to the nature of evidential markers in the examined constructions. First, there are languages, where there are no obvious differences in the use of evidential in declaratives and interrogatives (Kathmandu Newari, Duna). The second subtype is attested in languages, in which evidentials may appear in both constructions, but with manifest functional differences. For example, in Wutun, ego-evidentials may appear in both constructions, but in declaratives they appear with first person and in questions with second and third person. Third, there are languages like Chechen, in which the occurrence of evidential markers as such is not sensitive to the clause type, but the number of evidentials that may occur in interrogatives is lower and the nature of the markers may also be different. Finally, there are languages in which the form of the evidential markers is determined by the construction they appear in (Foe, Guambiano).

As the typology proposed above shows, the use of evidential markers in declaratives and interrogatives is clearly asymmetric; a higher number of evidential markers may occur in declaratives. The most important reason for this is probably found in the status of information source in the discussed clause types. The source of information is less relevant in interrogatives, because the focus is more clearly on the contents, and it is not important how the addressee has acquired the information. Moreover, we typically do not have access to the addressee's source of information, which renders it less natural to include an evidential marker into an interrogative. In declaratives, we do have a source of our own for the information, which we may (need to) specify. It is also worth noting that languages with and without obligatory evidentiality behave differently; evidential markers are more common in interrogatives in languages where evidentiality is an obligatory category. The status of information source is clearly different in the two language types, only languages with a highly grammaticalized evidentiality system may have evidentials also in interrogatives.

In my paper, the typology proposed above is discussed in light of cross-linguistic data. Moreover, I will also discuss the rationale behind the attested types in light of the semantics of declaratives and interrogatives. This also includes a discussion of the semantic nature of evidentials that may occur in interrogatives as well.

Theme session: Typological hierarchies in synchrony and diachrony

Person-Number Agreement in Mande: Synchronic Hierarchies and Their Sources

It is a textbook knowledge that typological hierarchies condition different morphosyntactic processes, such as case marking and agreement. For example, in Swahili the marker of object agreement is optional when the object NP is inanimate, but it is obligatory when the object is animate – cf. Morimoto (2002).

In this talk I am going to present some data on person-number agreement in Mande language family. Agreement in Mande is conditioned by several hierarchies including the number hierarchy and the referential hierarchy. The question is whether the two hierarchies can have the same synchronic (or diachronic) explanation. It seems that the answer is no. While there is no evidence that the number split originates in any special asymmetric construction and one has to account for it in terms of markedness or frequency; it is quite clear that referential split originates in relative clause constructions (at least in some languages) giving a good diachronic explanation of the hierarchy.

To my knowledge, Mande languages have never been discussed by typologists with relation to agreement. The fact is that the best documented Mande languages simply don't have it – cf. Creissels (1983) on Mandinka, Creissels (2009) on Kita Maninka, Dumestre (2003) on Bambara. The three languages just mentioned belong to the same branch of Manding languages within Mande family. However, other branches including Southern, South-Western and Eastern Mande have developed systems of person-number agreement (Konoshenko in print).

In Kpelle (South-Western Mande) the analytic predicative marker (auxiliary) agrees with the subject in person and number (personal data):

(1) *Pepɛ ɛ pɑ*
 Pepe 3SG.PM come\L
 'Pepe came'.

(2) *Nɛɑɑ dɑ pɑ*
 DEF\woman.PL 3PL.PM come\L
 'The women came'.

In the corresponding negative construction agreement is ungrammatical in singular (the predicative marker appears in a default unconjugated form) though obligatory in plural:

(3) *Pepɛ hvɛ pɑ*
 Pepe NEG come\L
 'Pepe didn't come'.

(4) *Nɛɑɑ dɪhvɛ pɑ*
 DEF\woman.PL 3PL.NEG come\L
 'The women didn't come'.

The number hierarchy PL > SG (and Positive > Negative construction) are in play here. I know of no morphosyntactic sources of these hierarchies in Kpelle, so they can be described by referring to markedness or frequency (Greenberg 1966; Croft 2003; Haspelmath 2006).

In Dan (South Mande) a NP controls agreement whenever it is definite (Vydrine, Kességbeu 2008: 70-71; p.c.):

(5) *Tɛɛ yaɑdŋ*
 basin put
 'Put a basin!'

(6) *Tɛɛ ɓɑ ɑ yaɑdŋ*
 basin DEF 3SG put
 'Put the basin!'

In (6) the definite article 𞐃𞐃 originates from a demonstrative adverb ‘there’ and the definite NP can be “unfolded” into a correlative clause meaning something like “the basin that is there”. It has internal head, and it is referred to by a pronoun 𞐃 in the main clause (cf. Comrie (1989) on the typology of relative clauses, also Nikitina (2012) on correlatives in Mande).

(7) [Tɔ̀ɔ̀ ɛ̀ ɔ̀] ɛ̀ yàɔ̀ɔ̀
 basin REL.3SG there 3SG put
 ‘Put the basin (that is there)!’

Thus the two hierarchies mentioned above appear to have different origins in Mande languages: while the number hierarchy is (probably) discourse-based, the referential hierarchy is syntax-based inheriting the properties of a special syntactic construction – a correlative clause.

Abbreviations

DEF – definite; NEG – negative marker; PL – plural; PM – predicative marker; REL – relative clause marker; SG – singular.

References

- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*, 2nd edition. The University of Chicago Press, Chicago.
- Creissels, Denis. 1983. *Eléments de grammaire de la langue mandinka*. Grenoble: Université des Langues et Lettres., Creissels, Denis. 2009. *Le malinké de Kita*. Cologne: Rüdiger Köppe. Rüdiger Köppe Verlag.
- Croft, William. 2003. *Typology and universals*. 2nd edition. Cambridge: Cambridge University Press.
- Dumestre, Gérard. 2003. *Grammaire fondamentale du bambara*. Paris : Karthala.
- Greenberg, Joseph. 1966. *Language universals, with special reference to feature hierarchies*. (Janua Linguarum, Series Minor, 59.) The Hague: Mouton.
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42, 25–70.
- Konoshenko, Maria. in print. Lichno-chislovoe markirovanie v jazyke kpelle: k tipologii soglasovania po litsu i chislu. *Voprosy jazykoznanija 1-2013*. [Konoshenko, Maria. in print. Person-number marking in Kpelle: on the typology of person-number agreement. *Voprosy jazykoznanija 1-2013*]
- Morimoto, Yukiko. 2002. Prominence mismatches and differential object marking in Bantu. In: M. Butt & T.H. King (eds.) *Proceedings of the LFG-02 Conference*. Stanford, CA: CSLI Publications.
- Nikitina, Tatiana. 2012. Clause-internal correlatives in Southeastern Mande: A case for the propagation of typological rara. *Lingua* 122: 319-334.
- Vydrine, Valentin, Mongnan Alphonse Kessébeu. 2008. *Dictionnaire Dan – Français (dan de l’Est) avec une esquisse de grammaire du dan de l’Est et un index français-dan*. St Pétersbourg: Nestor-Istoria.

Theme session: Lexical Typology of qualitative concepts

The linguistics of temperature: a lexical-typological study of the temperature domain

The aim of this talk is to show how languages vary in their categorization of temperature and how this cross-linguistic variation is constrained. It summarizes the results of a collaborative project, involving 35 researchers and covering more than 50 genetically, areally and typologically diverse languages, the data for which were elicited according to standardized guidelines, in most cases complemented by corpus searches. Both the guidelines and the corpus searches aim at systematically collecting and testing the attributive and predicative uses of temperature terms in various typical temperature-related situations, including applicability to different kinds of entities.

One of the main results of the project has been the identification of the crosslinguistic patterns in the linguistic temperature systems. Languages cut up the temperature domain among their expressions according to three main dimensions: TEMPERATURE VALUES (e.g., the distinction between warming and cooling temperatures, or the distinction between excessive heat and pleasant warmth), FRAMES OF TEMPERATURE EVALUATION (TACTILE, *The stones are cold*; AMBIENT, *It is cold here*; and PERSONAL-FEELING, *I am cold*), and ENTITIES whose “temperature” is evaluated.

A striking fact about the temperature systems across languages is their internal heterogeneity in that their different parts behave differently. Personal-feeling temperatures are often singled out by languages (in lexical choice and/or morphosyntactic patterns, in the reduced system of temperature value oppositions, etc), whereas the linguistic encoding of ambient temperature may share properties with those of either tactile or personal-feeling temperature. The motivation for this lies in the conceptual and perceptual affinities of ambient temperature with both other frames of temperature evaluation. On the one hand, ambient and personal-feeling temperature are rooted in the same type of experience, thermal comfort, whereas tactile temperature relates to evaluation of the temperature of other entities, based on perception received by the skin. On the other hand, tactile and ambient temperatures are about temperatures that can be verified from “outside”, whereas personal-feeling temperature is about a subjective “inner” experience of a living being.

Earlier cross-linguistic research on temperature, largely inspired by the mainstream research on colour, has suggested that all languages would possess basic temperature terms, ranging between two (‘cold’ vs. ‘hot’) and four and following a certain hierarchy (Sutrop 1998, Plank 2003). The cross-linguistically recurrent heterogeneity in the organization of the linguistic temperature systems makes the notion of „basic terms” either irrelevant or uninteresting for the study of many such systems (cf. Lucy 1994). Also, both the recurrent asymmetries in the organization of the linguistic temperature systems and the close interaction between lexicon and grammar in the encoding of temperature justify a more integrated approach to these phenomena than what has been the norm in typology, with its strict distinction between lexical and grammatical typology. A constructional-typological approach (cf. Koch 2012) considers cross-linguistic patterns defined both by lexical and grammatical information and investigates the “division of labour” between the lexicon and the grammar and their “semiotic ecology” (Evans 2011: 508), i.e. whether semantic choices made in the lexicon affect those in the grammar and vice versa.

References

- Evans, N. 2011. Semantic typology. In Song, J. J. (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: OUP, 504 – 533.
- Koch, P. 2012. Location, existence, and possession: A constructional-typological exploration. In Koptjevskaja-Tamm, M. & M. Vanhove (eds.), *New directions of lexical typology. Special issue of Linguistics*, 50, 3: 533-604
- Lucy, J. 1994. The role of semantic value in lexical comparison: motion and position roots in Yucatec Maya. *Linguistics* 32: 623-656
- Plank, F. 2003. *Temperature Talk: The Basics*. A talk presented at the Workshop on Lexical Typology at the ALT conference in Cagliari, Sept. 2003.
- Sutrop, U. 1998. Basic temperature terms and subjective temperature scale. *Lexicology*, 4: 60–104.

Theme session: Generalized Noun Modifying Clause Constructions

Noun (phrase) modifying clause constructions in Turkish and some other Turkic languages

This talk proposes an account for a difference between two types of Altaic languages: One type allows noun—complement clause constructions with very loose semantic and formal connections between the external noun and the complement clause. We refer to such constructions as ‘the smell of roasting meat’ structures. The other type doesn’t allow such constructions.

The talk is concerned with some formal factors which influence this dichotomy. Turkic languages which allow ‘the smell of roasting meat’ also have other properties which set them apart from Turkish, which doesn’t allow this type of construction. We refer to those languages as Turkic 1. Considering three types of complex constructions (1. those without external nouns, 2. those with external nouns corresponding to relative clauses, and 3. those with external nouns corresponding to noun—complement clause constructions), Turkic 1 languages have the *same* range of nominalization morphology for *all three types* of constructions. In Turkish and most dialects of Azerbaijani, languages which we dub Turkic 2, the nominalization morphology is only *partially* similar across these three types of embedded clauses. The “indicative nominalization” morphology is found in all three types; however, “subjunctive nominalization” occurs only with noun—clausal complement constructions and in embedded clauses without external nominal head, but not with relative clauses. Furthermore, in Turkic 2, subject relative clauses require a special marker.

Summary of facts: A: The “subjunctive nominalization” morphology doesn’t show up in Turkic 2 relative clauses; **B:** the special morpheme for subject relative clauses doesn’t show up in other subordinate clauses. In Turkic 1, no such distinctions among nominalization morphemes, determined by distinct syntactic structures, are found. We propose that languages such as Turkic 2 with fine differences among nominalization morphemes also don’t permit constructions such as ‘the smell of meat roasting’. In Turkic 2, the syntax and external morphology of noun-complement constructions show that these are phrasal *compounds* in Turkic 2, with a close relation between a compound head and its complement, and in relative clauses, which aren’t compounds in Turkic 2, the relativization target, corresponding to the external noun, determines the shape of the clause’s predicate. Sakha, as a representative of Turkic 1, does not show these special properties of noun—complement clause constructions and of relative clauses in either morphology or syntax.

We hypothesize that the possibility of exhibiting constructions such as ‘the smell of meat roasting’ depends on how close a relationship an embedded clause has with an external noun *in general, elsewhere in the language*. In Turkic 2, which has a tight subcategorization-like relation between the external noun and the embedded clause in all externally headed constructions in terms of semantics and morphology, utterances such as ‘the smell of meat roasting’ are not possible. In contrast, because these tight relationships between external noun and clause don’t exist in Sakha/Turkic 1 or in Standard Mongolian, such languages do allow utterances such as ‘the smell of meat roasting’, ‘the sound of wind blowing’ etc.

Theme session: Lexical typology of qualitative concepts

'Wet' and 'dry': a cross-linguistic study

The talk deals with adjectives referring to the presence or absence of moisture on the surface or in layers of objects.

Unlike most other groups of qualities, these adjectives show rather different compatibility potential in regard to the nouns they determine. Generally, each of the synonyms covering a selected semantic field (e.g. speed, length, etc.) tends to have a restricted range of subjects it can be combined with, and the same noun rarely may appear in the context of two or more such adjectives with alike meaning. That is, criteria and restrictions of use for every item in the group of synonyms is preset by the semantic class of a noun, cf. the Spanish counterparts of 'sharp': *afilado* for tools with blade (*cuchillo afilado* 'sharp knife') and *puntiagudo* for pointed tools (*lanza puntiaguda* 'sharp spear').

Yet, adjectives of the domain 'wet' seem to deviate from this apparently standard scheme. For this domain, we encounter a significant number of intersecting contexts, cf. in English *wet/damp/moist floor*, *wet/damp/moist towel*, *wet/damp/moist hair*, etc. This property holds for many other languages as well. Dictionaries seem to give a plausible explanation for this fact, differentiating in various languages the adjectives under discussion mainly in terms of degree of humidity they suppose, i.e., for instance, in German *feucht* 'damp, moist' is explained as *ein wenig nass* 'a little wet', in Spanish *húmedo* as *ligeramente mojado* 'slightly wet', etc.

Having compared up to 4 lexical items for excessive wetness in 11 languages (Russian, German, Spanish, English, Chinese, Arabic, Hebrew, Hausa, Swahili, Turkish, Khanty), we can state that the peculiarities of their use cannot be accounted for by the degree of humidity taken alone. Rather, the latter gets a specific interpretation depending on the noun modified. Thus, if applied to 'hands', the adjective 'wet' (i.e. referring to a higher degree of moisture) denotes the situation of washed hands, while 'damp/moist' (i.e. a lower degree of moisture) implies the idea of sweating, cf. German *nasse* vs. *feuchte Hände*, Russian *mokrye* vs. *vlažnye ruki*, etc. When speaking about firewood, 'wet' gives the rain as a cause, while 'damp/moist' stands for a freshly felled tree, cf. Khanty *jinki* vs. *juχ*. For a forest, higher intensity supposes rather a temporary state, while lower denotes a permanent feature (cf. Spanish *el bosque mojado* 'forest wet after the rain' vs. *húmedo* 'rainforest').

Apart from degree which is understood in quite different ways, those languages that have more than two terms for the 'wet' domain reveal several additional oppositions, cf., among others, the speaker's attitude to the situation (e.g., English *moist* VS. *damp soil*, for the Russian cf. Apresyan (ed.) 2004, Tolstaya 2005).

The 'dry' domain in our sample is poorer than that of 'wet' and contains two lexical items at most. The opposition between them concerns the origin of the lacking liquid – it may be either inherent ('dry brook') or due to some external effect ('dry clothes'), cf. Arabic *jaaff-un* vs. *yaabis-un*.

References

- Apresyan Yu.D. (ed.) 2004. *Novyj ob"jasnitel'nyj slovar' sinonimov russkogo jazyka*. Moscow.
 Tolstaya S. M. 2008. *Prostranstvo slova. Leksičeskaja semantika v obščeslavjanskoj perspective*. Moscow.

Schematizing cross-linguistic variation: the case of reciprocal constructions

The reliance on in-depth analyses of grammatical categories in individual languages becomes challenging once integrated into a cross-linguistic (typological) perspective. The detailed and carefully selected data sets from individual languages must form the empirical basis of cross-linguistic comparison and resulting generalizations. Yet, by its very nature, the typological perspective forces the linguist to abstract away from many idiosyncrasies and differences that subsist between languages, and to adopt a terminology that is capable of capturing the similarities in form and meaning.

I here present a method of schematization of concrete data sets (see below) taken from different languages, which allows one to conduct cross-linguistic comparisons, however without losing the language-specific perspective out of sight. The merit of this method, besides resulting in rather concrete, easily interpretable visualizations (schematic diagrams), lies in the fact that comprehensive data sets on grammatical features from different languages can be concisely summarized and juxtaposed to one another, and moreover exploited in quite different ways, depending on the specific formal/semantic parameters chosen for cross-linguistic comparison.

For illustration, I turn to a well-studied grammatical phenomenon, the reciprocal construction, which has been comprehensively studied over past decades in both the descriptive and typological literature (e.g., Nedjalkov 2007; König & Gast 2008; Evans et al. 2011; Maslova & Nedjalkov 2005; König & Kokutani 2006; Evans 2008). In part, I will draw on data sets sampled from primary sources (e.g., descriptive grammars, dictionaries) on a selected number of Australian and Papuan languages. In addition, I will include data presented elsewhere in the literature on reciprocals (cf. the above-mentioned references) to demonstrate that the method presented is applicable and amenable to quite different data resources.

In presenting the data on reciprocals, I will assume both a cross-regional and regional perspective. The cross-regional comparison focuses on the fact that languages tend to rely on quite different structural means for the expression of reciprocal 'each other' meanings (e.g., verb-marking vs. argument-marking strategies, cf. König & Kokutani 2006), but that such different types of structural encoding and their contexts of occurrence (e.g., preferably with non-symmetrical word forms) are yet recurrent across languages. The regional comparison in turn illustrates that both cognate and analogous instances of grammaticalization need consideration, but that cognate marking patterns often require rather different synchronic assessments from one another. For example, variants of the verbal suffix *-nyji are widely distributed among non-Pama-Nyungan languages of Northern Australia (cf. Alpher, Evans & Harvey 2003), yet their productivity and functional status as dedicated markers of reciprocity differs quite strikingly from one language to another.

I propose that the method of schematization presented here can contribute to cross-linguistic generalizations, because it allows one to visualize in single diagrams comprehensive data sets from individual languages, while at the same time enabling linguists to integrate these into the broader perspective of cross-linguistic comparison. In this context, the term 'family resemblance' in reference to analogous or cognate instances of grammaticalization (e.g., Evans & Levinson 2009) also receives a very concrete visual interpretation.

Qualitative concepts ‘sharp’ and ‘blunt’

The semantic fields ‘sharp’ and ‘blunt’ are of interest for lexical typology since (a) they are frequently used physical qualitatives rich in polysemy (cf. Koch & Marzo 2007: 278); (b) they demonstrate a non-trivial combination of tactile and visual perceptual fields, which varies from language to language (‘sharp knife’ vs. ‘sharp nose’).

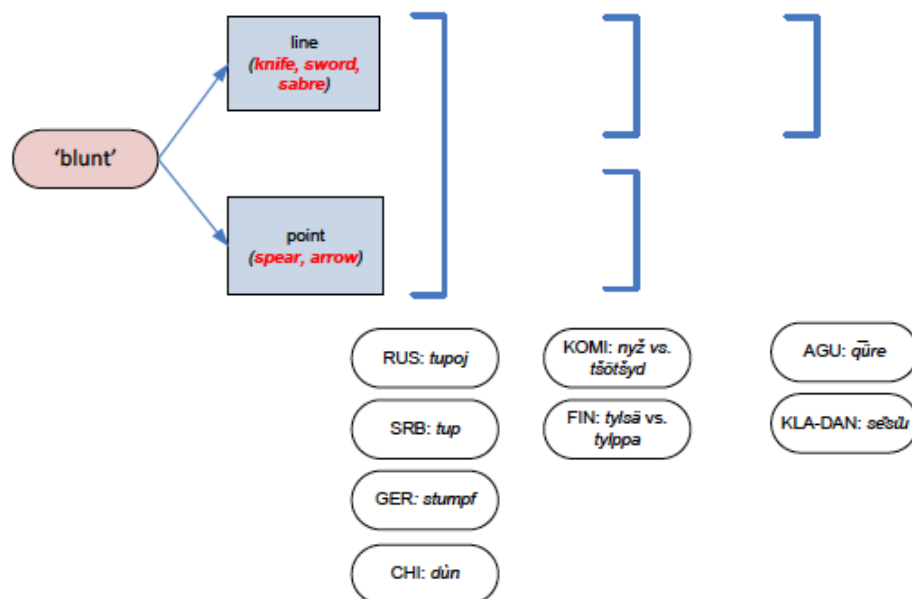
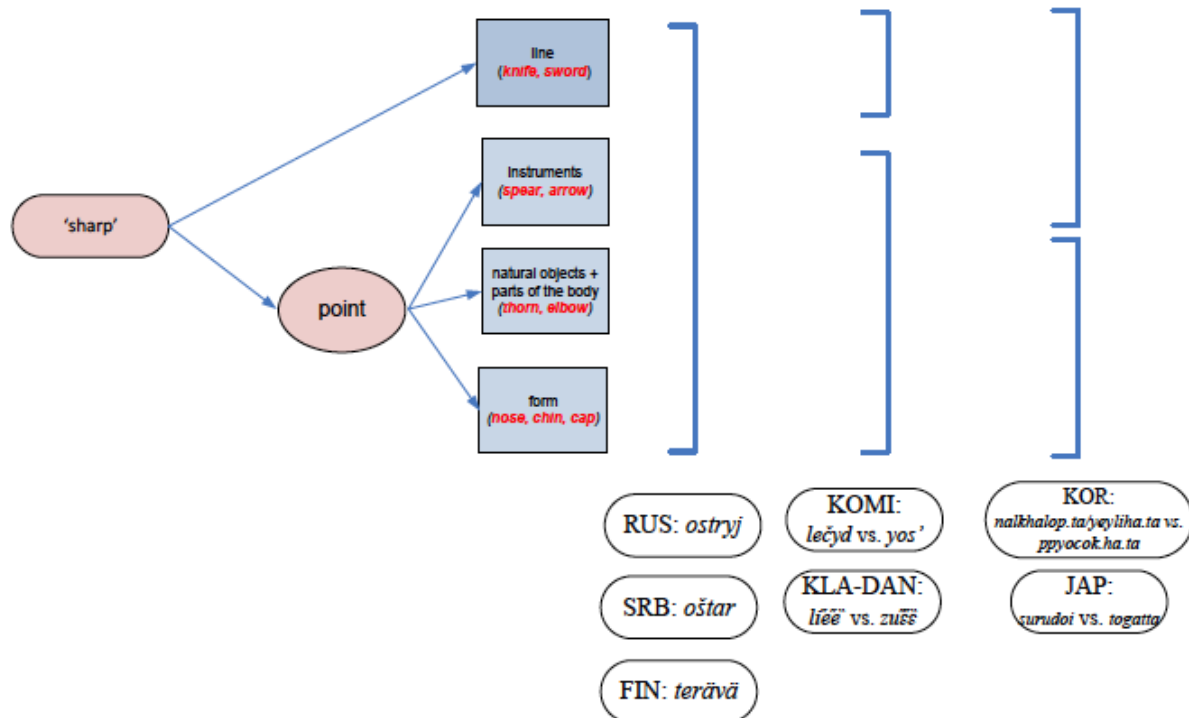
Our primary sample consisted of 15 languages: Russian, Serbian, French, Italian, English, German, Welsh, Finnish, Komi-Zyrian (Permic branch of the Uralic languages), Chinese, Japanese, Korean, Aghul (Lezgian branch of the Northeast Caucasian languages), Malay (Austronesian), Basque and Kla-Dan (South branch of Mande languages). The findings were further tested on five more languages: Hindi, Welsh, Spanish, Armenian and Russian sign languages. Our data sources are corpora and questionnaires filled in by native speakers, as well as previous research as presented for example in ([Fritz 1995], [Fritz 2005: 118-130]).

Our first result is the demonstration of an asymmetrical lexicalization for an antonymous relationship: the field ‘blunt’ is secondary to the field ‘sharp’ and is much less developed, i.e. this field is regularly covered with less lexemes, and these usually have less figurative meanings.

Three parameters were found that describe the differences between lexemes within each of the fields:

1. The type of instrument: that with a cutting edge (‘knife’, ‘saw’, ‘sword’) vs. that with a piercing point (‘needle’, ‘spear’, ‘arrow’). This parameter is relevant for both fields and explains the existence of different lexemes for the description of two types of instruments in Italian, French, Chinese, Kla-Dan, Komi-Zyrian and Finnish.
2. Perceptual field: tactile vs. visual perception, in other words the description of an object according to its function (‘sharp knife’ – a knife that cuts well) or to its form (‘sharp nose’ – a nose with a certain form). This opposition is lexicalized in German, Japanese and Korean.
3. Type of sharpness: functional vs. nonfunctional. The first is attested in instruments that are intentionally sharpened to improve their functional potential (‘arrow’, ‘saw’, ‘knife’). The second is attested in objects (primarily natural), whose sharpness is not functionally useful for humans and can rather hurt them (‘thorn’, ‘heel’, ‘horns’). Specialized adjectives to describe objects of the last type are found in many languages, i.e. Basque and Italian. This parameter is irrelevant for ‘blunt’.

The following “lexical map” shows possible combinations of meanings within one lexical item:



Different classes of direct meanings develop different metaphors that recur from language to language. Thus, adjectives for instruments with a cutting edge describe clear lines, borders and metonymically images, adjectives for pointed instruments describe good hearing or sight and adjectives for sharp natural objects denote a durable but not very intensive pain (cf. Reznikova et al. 2012).

References:

- Fritz, Gerd. 2005. Einführung in die historische Semantik. Tübingen: Max Niemeyer Verlag.
- Fritz; Gerd. 1995. Metonymische Muster und Metaphernfamilien. Bemerkungen zur Struktur und Geschichte der Verwendungsweisen von „scharf“. // Hindelang, G./Rolf, E./Zillig, W. (Hg.): Der Gebrauch der Sprache. Festschrift für Franz Hundsnurscher zum 60. Geburtstag. Münster: LIT Verlag. p. 77-107.
- Koch, Peter, Marzo, Daniela. 2007. A two-dimensional approach to the study of motivation in lexical typology and its first application to French high-frequency vocabulary. // *Studies in Language* 31:2, p. 259–291.
- Reznikova, T.; Rakhilina, E.; Bonch-Osmolovskaya, A. 2012. Towards a typology of pain predicates // *Linguistics*. Volume 50, Issue 3, pp. 421–465.

Local Cases in EIA Languages

The paper describes the local cases of some of the eastern Indo-Aryan languages using cognitive framework. The languages under observation are Angika, Bhojpuri, Maithili, Magahi (Group A), Asamiya, Bangla and Oriya (Group B) spoken in India. I have divided the languages into two groups for the convenience of description. The paper is divided into four parts. First part 'Introduction' introduces the languages and gives a basic idea about the local cases. The second part, 'Static' describes the static cases of these languages. The languages of Group A have two locative markers showing two way distinction; /me/ (/mē/) and /p̄r/. Both are postpositions which follow the noun (LM) in relation to which the other object's (TR) position is marked. Broadly, /me/ marker can be said to be used for the sense of inclosure. To mark the periphery the marker /p̄r/ is used but I have tried to prove that both the markers are of two different levels. Group B languages have one marker which can be called general spatial term (GST) (Levinson 2003, Feist 2008). It shows location of TR in context to LM but the marker neither states the position nor the direction of the TR. It is expressed only through the context.

In the last part of the second section I try to prove that in Group A languages /me/ is the original locative marker which was used in every context of location. I have given arguments in the favour of the argument that /p̄r/ is a recent development in these languages. The occurrence of /p̄r/, the grammaticalised version of /up̄r/ in these languages is redefining the meaning of /me/ in these languages. The occurrence of /p̄r/ has redefined the meaning of /me/ by narrowing down its meaning.

Group A languages do not use static marker with animate objects whereas Group B languages can do so. Unidimensional spatial case systems tend to be organized according to a tripartite distinction between location, destination of movement, and source of movement (Creissels 2009). The languages under observation too mark three spaces which can be stated as static location, starting point and path. The third part 'Dynamic', describes the two dynamic relationships (starting point and path) between the LM and the TR. Among all these seven languages, only Oriya perceives path differently. In other languages it is marked by instrumental and (or) ablative cases.

In the last section 'Conclusion', I have compared and contrasted the static and dynamic cases of both the groups. The use of the verb generally decides whether the TR is static or dynamic. But it was interesting to find that sometimes when the TR is just a patient then depending on the context the TR can either be marked by the static marker or the dynamic marker. Replacing one by the other does not make the utterance infelicitous but they are semantically different and contextually bound.

References

- Creissels, Denis. 2009. *Spatial Cases*. In: The Oxford Handbook of Case. Edited by Andrej Malchukov and Andrew Spencer, 609-625. New York: Oxford University Press.
- Feist, M. I. 2008. *Space between languages*. In: Cognitive Science, 32 (7), 1177-1199.
- Levinson, Stephen C. 2003. *Space in Language and Cognition: Explorations in Cognitive Diversity* Cambridge: Cambridge University Press.

Verb Innovation in Palestinian Arabic

This talk sheds light on the strategies speakers of Palestinian Arabic (PA) use in verb innovation. By verb innovation I refer to denominative verbs (e.g. *tmarkaz* 'become central', derived from *markaz* 'center') and verbs that are formed based on loan words (e.g. *tnarfaz* 'become nervous' derived from the English adjective *nervous*). The verbal system of PA consists of nine templatic configurations. Each template indicates the phonological shape of the verb, i.e. its vowels, its prosodic structure and its affixes (if any). The phonological shape of a verb is essential for determining the shape of other forms in the inflectional paradigm (Bat-El 1989, Aronoff 1994). A verb, which does not conform to one of the existing templates cannot enter the verbal system. This paper examines the criteria that play a role in the selection of a prosodic template for new verbs. I argue that the process of choosing a template involves the interaction of both thematic-syntactic and morpho-phonological criteria.

Morpho-phonological criteria are responsible for favoring two templates, *CaCCaC* and *tCaCCaC*, over others due to their prosodic structure. Both of them have four slots for consonants, where the medial two are usually occupied by a geminate, therefore allowing forms with more than three consonants (e.g. *talfan* 'phone', derived from *telephone*). When the base contains three consonants these templates are selected due to paradigm uniformity (Steriade 1988) within the derivational system; speakers aim for the same prosodic shape for new verbs. In case of three stem consonants, one mora slot remains empty. There are two main strategies for filling this slot: either by gemination of the second consonant or by insertion of a glide. For example, *šarraj* 'charge' is derived from the English verb *charge*, where the second consonant /r/ fills the empty mora slot. In contrast, the derived verb of *kolon* 'cologne' is *kalyan* 'use cologn' and not **kallan*. In this case, the /l/ consonant does not fill the mora slot, but the glide consonant /y/ is inserted. The data I have examined demonstrates a use of these two strategies although the former is more common. I will account for the constraints responsible for using the two strategies.

Thematic-syntactic criteria concern the syntactic valence of verbs and their thematic grids. It is commonly assumed that different thematic realizations of the same concept (e.g. unaccusative, reflexive) are derived from the same basic entry via thematic valence changing operations. The division of labor between *CaCCaC* and *tCaCCaC* is based on the status of new verbs in the lexicon. Verbs that are basic entries are formed in *CaCCaC*. These are mainly transitive verbs, e.g. *_akkas* 'put an X on somebody' but also other non-transitive basic entries, e.g. *barrak* 'apply brakes'. *tCaCCaC*, in contrast, is selected for verbs derived by thematic operations that reduce the syntactic valence of a verb (e.g. reflexivization and reciprocalization). Such verbs are formed in *tCaCCaC* even if they have no transitive counterpart in *CaCCaC*. For example, the reflexive verb *tšaxlal* 'upgrade oneself' is derived from the Hebrew verb *hištaxlel* 'become upgraded' and has no transitive alternate **šaxlal*. These data support the analysis of the division of labor between the two templates based on 'base vs. derived form' criteria.

The analysis reveals the interaction between morpho-phonological and thematic considerations, thereby supporting the interface between morpho-phonology and the lexicon. It also supports the view of the lexicon as an active component in the generation of words (Aronoff 1976). The choice of a prosodic template in coining new verbs provides evidence for knowledge of constraints that are taken into consideration in word formation processes.

Theme session: Generalized Noun Modifying Clause Constructions

East Caucasian relativization: descriptive categories vs comparative concepts

East Caucasian relativization is hardly subject to any syntactic constraint. Grammars have to provide long lists of syntactic positions which are all perfectly relativizable. Moreover, in some cases it is problematic to ‘reconstruct’ a finite clause corresponding to the relative one. In the following example, the noun ‘place’ cannot be inserted into the original clause to yield a grammatically and pragmatically well-formed utterance:

- (1) Aghul (Maisak 2008):
 jak: ug.a-je niᶒ ‘a smell of burning meat’
 meat burn.IPF-PART2 smell

Other syntactic constraints seem to be at loss; cf. (2) where relativized is the position in an embedded relative clause and thus violates an island constraint.

- (1) Tanty Dargwa (elicited)
 [dam čᶒ-ib-se kᶒata b-ibšᶒ-ib xᶒunul] simi r-ačᶒ-ib
 I:DAT give.PF-PRET-ATR cat N-flee.PF-PRET women bile F-come.PF-PRET
 ‘The woman such as that the cat that she brought to me ran away (she) became angry with me’.

Such odd syntactic behavior is sometimes considered an indication that what we deal with are not relative clauses at all, but, to use one of the approaches (see, e.g., van Breugel 2010), a more general syntactic phenomenon defined as nominal modification by a verbal constituent. The problem of the relative clause definition is used as one example in Haspelmath’s (2011) distinction between comparative concepts and descriptive categories. From the point of view of this distinction, whether we count examples above as relative clauses would probably be a purely definitional issue. Note that from the language-internal perspective (descriptive category), the constructions in question are invariably considered as relative clauses (cf. the references below).

We believe, however, that this question may (or must) be settled on empirical grounds. In the vast majority of text occurrences such clauses do not violate any constraints, looking like well-behaving relative clauses. Central arguments are relativized much more frequently than peripheral ones, in full conformity with Keenan-Comrie’s predictions. The use of resumptive pronouns, widely attested in some East Caucasian languages, is again best compatible with viewing them as relative clauses. It would be hardly feasible to posit a separate typological category (comparative concepts) basing on peripheral – even though fascinating – uses of a construction. What we deal with in East Caucasian is not a phenomenon distinct from relativization but its extension: examples such as (1) and (2) are deviations from more ‘natural’ cases of relativization rather than a special syntactic pattern and East Caucasian relativization as a whole (descriptive category) shows to many empirical parallels with the typology of relativization (comparative concept).

References

- Anderson, G. 2006. *Auxiliary Verb Constructions*. Oxford: Oxford University Press
 Heine, B. 1993. *Auxiliaries, Cognitive Forces, and Grammaticalization*. New York: Oxford University Press
 Kuteva, T. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press

Theme session: Generalized Noun Modifying Clause Constructions

On Sino-Tibetan clausal noun modifying constructions as noun compound constructions

Using naturally occurring data, this paper briefly discusses the history of the development of clausal noun modifying constructions in Sino-Tibetan languages, showing the development of a construction with a nominalized clause as the modifier in many ST languages, resulting in a nominal(modifier)-nominal(modified) structure, similar to noun compounds, though in the case of the clausal noun modifying constructions either element could stand alone as a referring expression to refer to the same referent. The paper then focuses on the relevant structure in Mandarin Chinese, showing that there is no grammatical restriction on the interpretation of the semantic relationship between the two elements (modifier and modified) in the construction. That is, unlike in Indo-European relative clause constructions, there is no “gapped” argument in the nominalized clause, and modified element does not have to be understood as an argument of the modifying clause, so the interpretation of the referent of the nominalized clause and its relationship to the head of the construction (if there is one) is left completely to inference from context and real world knowledge. It is argued that a constructionist approach, accepting the construction as is, where the construction has a meaning greater than the sum of the parts, and not talking about it using terminology from old transformational generative grammar such as “relativisation on the subject” or “relativisation on the object”, is more appropriate, particularly as the construction does not necessarily relate clearly to other forms, such as main clauses, in the way imagined in transformationalist approaches.

The lack of grammatical constraints on the interpretation of the relationship between the two elements in the construction is similar to the pattern found in noun compounds in many languages, so seeing the structure as a noun compound might be the explanation for the lack of grammatical constraints on the interpretation, though suggestions about the nature of grammatical constraints in Chinese in general point to it being a wider phenomenon in the language.

A Typology of Metathesis in Formosan Languages

As a common phonological process in language, metathesis has been described in several Formosan languages, including Atayal (Li 1977, 1980, Chen 2011), Bunun (Li 1977, Lin 2008, de Busser 2009), Thao (Blust 2003), Tsou (Tung 1964, Li 1977), and Seediq (Tsukida 2009, Lee 2010, Lee 2012). Despite being a common phonological process, metathesis in these languages appears to be less studied or understood from both typological and theoretical perspectives.

The aim of this paper is thus two-fold: first, metathesis found in Formosan languages is studied from a typological perspective, based on the typology by Blevins & Garrett (1998, 2004), in order to determine to what type the process found in a given language belongs. The cases of metathesis occurred in other languages as discussed in the literature are compared typologically with those found in Formosan languages in order to obtain a generalized view. Secondly, based on the observations above, this paper attempts to investigate whether the process in a given language is phonologically-driven and/or morphologically-conditioned, and most importantly, what linguistic factors motivate the process in these languages.

A preliminary observation reveals that four types of metathesis are found in Formosan languages: perceptual metathesis (Atayal, Seediq), coarticulatory metathesis (Thao, Atayal), vowel metathesis (Bunun, Tsou), and complementary metathesis (Tsou). Coarticulatory and perceptual metathesis are phonologically-driven by phonotactic and syllable constraints. Vowel and complementary metathesis, though involving hetero-morphemic boundaries, are also motivated by prosodic requirements.

This paper not only provides a typological approach on metathesis in Formosan languages but also tries to explain the reason why metathesis occurs in certain contexts in a given language, as well as how these different types of metathesis can fit into current theoretical frameworks (cf. Hume 1998, 2001, 2004).

References

- Blevins, J. and Garrett, A. 1998. The origins of consonant-vowel metathesis. *Language* 74.3: 508-555.
- Blevins, J. and Garrett, A. 2004. The evolution of metathesis. *Phonetically-based Phonology*, ed. by Bruce Hayes, Robert Kirchner, and Donca Steriade, 117-156. Cambridge: Cambridge University Press.
- Blust, R. 2003. *Thao Dictionary*. Language and Linguistics Monograph Series, No. A5. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
- Chen, Yin-ling. 2011. *Issues in the Phonology of Ilan Atayal*. Hsinchu: National Tsing Hua University PhD dissertation.
- De Busser, R. 2009. *Towards a Grammar of Takivatan Bunun: Selected Topics*. Victoria: La Trobe University Ph.D. dissertation.
- Hume, E. 1998. Metathesis in phonological theory: the case of Leti. *Lingua* 104: 147-186.
- Hume, E. 2001. Metathesis: Formal and functional considerations. *Surface Syllable Structure and Segment Sequencing*, ed. by Elizabeth Hume, Norval Smith & Jeroen van de Weijer, 1-25. Leiden: HIL.
- Hume, E. 2004. The indeterminacy/attestation model of metathesis. *Language* 80.2: 203-237.
- Lee, Amy P. 2010. Phonology in Truku Seediq. *Taiwan Journal of Indigenous Studies*. 3.3: 123-168.
- Lee, Amy P. 2012. Contact-induced sub-dialect in Toda Seediq. Paper presented at the 12th International Conference on Austronesian Linguistics, July 2-6, Udayana University, Bali, Indonesia.
- Li, Paul J. 1977. Morphophonemic Alternations in Formosan Languages. *BIHP* 48.3: 375-413.
- Li, Paul J. 1980. The phonological rules of Atayal dialects. *BIHP* 51.2: 349 - 405.

- Lin, Hsiu-hsu. 2008. Morphologically-conditioned Metathesis in Bunun. Taipei: National Science Council Report NSC 97-2410-H-324-016. [In Chinese]
- Tung, T. 1964. *A Descriptive Study of the Tsou Language, Formosa*. Taipei: Institute of History of Philology, Academia Sinica Special Publication, No. 48.
- Tsukida, N. 2009. *The Grammar of the Seediq Language (Taiwan)*. Ph.D. dissertation, Tokyo: University of Tokyo. [In Japanese]

The typology of demonstratives clarified: verbal demonstratives in Ju

The present paper is concerned with the rare category of verbal demonstratives, its exact definition, and the place it deserves in the typology of demonstratives.

In his typology of demonstratives in the world's languages, Diessel (1999) distinguishes four syntactic categories of demonstratives (pronominal, adnominal, adverbial and identificational demonstratives), none of which has verbal or predicative status. Dixon's (2003) typology of demonstratives, relatively close to Diessel's, does acknowledge the existence of a separate category of verbal demonstratives, but does not propose a clear and convincing definition of this category. His typologically extremely rare "verbal demonstrative" category indeed covers two different types of verbs expressing some degree of deixis: an action verb expressing manner deixis ('do thus, do like this') in Boumaa Fijian and Dyrbal, illustrated in (1) below, and the exophoric proximal and distal deictics in Ju'l'hoan (Ju, formerly known as Northern Khoisan), illustrated in (2).

(1) 'Do thus/like this' (abbreviations and glosses are the author's):

a. Boumaa Fijian
 [o 'ae]_S ['enei_i tuu gaa 'enei]_{PREDICATE}
 ART 3SG do.like.this ASPECT just do.like.this
 'He did just like this.' [narrator mimes a spearing action] (Dixon 2003:72)

b. Dyrbal
 bala_O baja! ŋaja_A [yalama-n baja-n]_{PREDICATE}
 there:N chew:1MP 1SG do.like.this-NON.FUT chew-NON-FUT
 'Chew it [the spear grass]! I'm chewing (it) like this.' (Dixon 2003:102)

(2) Exophoric deictics (Ju'l'hoan):

a. jù hè
 person PROX
 'This is a person.'

b. ntòh tò'à
 orange DIST
 'That is an orange.' (Dickens 2005:49)

The present paper, on the basis of a thorough analysis of the available Ju data, in particular Dickens' (2005) description of Ju'l'hoan, aims at 1) showing that the unusual category of verbal demonstratives does exist, and defining its properties, 2) showing that verbal demonstratives are so far only attested in Ju languages (cf. Ju'l'hoan in (2)), while the Boumaa Fijian and Dyrbal verbs illustrated in (1) do not qualify as demonstratives. I further show, based on the predictions of a slightly updated version of Stassen's (1997) typology of intransitive predicates, that both the existence of verbal demonstratives and their rarity are typologically expected. Finally, I propose a modified version of the typology of demonstratives, which, at last, includes a well-defined category of verbal demonstratives, and also propose ways to identify verbal demonstratives in other languages, potentially including well-described languages where they might have gone unnoticed.

References

- Dickens, Patrick J. 2005. *A Concise Grammar of Ju'l'hoan*, Quellen zur Khoisan-Forschung vol.17, Cologne, Rüdiger Köppe.
- Diessel, Holger. 1999. *Demonstratives: form, function and grammaticalization*. Amsterdam, Philadelphia: John Benjamins.
- Dixon, Robert. 2003. Demonstratives: a cross-linguistic typology. *Studies in Language* 27(1).61-112.
- Stassen, Leon. 1997. *Intransitive Predication*. Oxford: Calendron Press.

Distance-Marking Correspondence

This paper posits a rule of correlation between word order and morphology, i.e. Distance-Marking Correspondence, which states that everything else being equal, the further away a dependent is from its head, the more it needs an overt grammatical marker which indicates the semantic relationship between the two units. In what follows, the relevant head and its dependent are underlined and the relevant marker is in bold.

- (1) a. He drives slow(-ly).
 b. He drives his car slow?(-ly).
 c. He drives his car into the garage slow*(-ly).

The obligation for the adverbial marker *-ly* increases with the distance between *slow* and *drives*. Mandarin has the similar cases:

- (2) a. *Ta zai tushuguan renzhen(-de) zhao ziliao.* (Mandarin)
 he in library cautious-ly search data
 ‘He is cautiously searching data in the library.’
 b. *Ta renzhen*(-de) zai tushuguan zhao ziliao.*

In (2), when the adjective *renzhen* ‘cautious’, which serving as an adverbial, is adjacent to the verb, the adverbial marker *-de* is optional, otherwise it is obligatory.

- (3) a. to climb (up) the mountain. b. to climb steadily *(up) the mountain.
 (4) a. women jinhou duoduo yong dianhua lianxi (Mandarin)
 we future more via telephone contact
 ‘Let’s contact more via telephone in the future.’
 b. women jinhou *(yong) dianhua duoduo lianxi
 we future via telephone more contact

In (3-4), the prepositions are optional when the relevant dependents are adjacent to the head verb, otherwise they are obligatory.

- (5) a. to provide somebody with something. b. to provide something for somebody.
 (6) a. to load the hay onto the truck. b. to load the truck with the hay.
 (7) a. I took three years of Chinese. b. I took Chinese for three years.

In (5-7), the dependents adjacent to the verbs do not need a preposition, while the dependents separated from the verbs must take a preposition.

- (8) a. He was my lover *(**for**) 20 years. b. He was 20 years my lover.

Above, *for* can be dropped when *20 years* is adjacent to the verb *was*. Below, the drop of complementizer *that* is also related to the distance between *Mary will win* and the verb *believes*.

- (9) a. John believes (that) Mary will win.
 b. John believes wholeheartedly *?(that) Mary will win.
 c. *(That) Mary will win, John believes wholeheartedly.

The rule of correspondence is also applicable to NP internal structure.

- (10) a. *zhege jianzhuwu shijie youming*. (Mandarin)
 this building world famous
 'This building is world-famous.'
 b. *zhege jianzhuwu zai shijie shang hen youming*.
 this building at world LOC very famous
 'This building is very famous in the world.'

Above, a circumposition *zai...shang* 'in' is added as *shijie* 'world' is separated by *hen* 'very'. In the phenomenon observed above, several factors may be involved, however, the distance between the head and its dependents is clearly an important one among them. More data from several different languages will be provided in the paper. The relevant overt markers included agreement, case markers, applicative markers, etc.

Theme session: Lexical typology of qualitative concepts

‘Direct’ and its synonyms: a case study on the grammaticalization in adjectives

In the recent decades the process of grammaticalization has been attracting much attention in typological work. One of the fundamental issues discussed within this field is what the correlation between the source (lexical) and derived (grammatical) semantics is. Yet, a substantial gap in describing the sources of grammaticalization is that for many researchers attributives are put in the shade of substantives and predicatives (in Heine & Kuteva 2002, for instance, only two adjectives are discussed in this perspective). However, attributive words are frequently subject to grammaticalization and this evolutionary process is still to be studied in more detail.

This research focuses on diachronic processes in attributives that mean ‘direct’/‘straight’ and is based on Russian, Polish, English, German, Yiddish, Italian, Lithuanian, Finnish, Chinese and Japanese. Generally these lexemes are transformed into particles (cf. in this respect the latest debate on grammaticalization and pragmaticalization, see, e.g. Traugott/Dasher 2002, Diewald 2006). These particles fall into two large semantic classes:

- A. Contrastive focus particles, which indicate an entity selected from an associative set.
 - (1) German: *Gerade du solltest das besser wissen!* ‘It is you who (lit. directly you) should have known it better!’
 - (2) Finnish: *Äidin kertoma oli **suoraan** se mitä olin kuvitellutkin.* ‘My mother’s story was exactly (lit. directly) about what I’d expected’
- B. Modal particles, which introduce speaker’s attitude. One of the examples here is emphatic similative. In this case the speaker describes a given situation P1 as similar to a prototypical situation P2, while P2 is rated so highly on his evaluative scale that it was not expected to occur.
 - (3) Lithuanian: *Jis **tiesiog** didvyris.* ‘He is a real (lit. directly) hero’
 - (4) Russian: *Stol **prjamo** lomilsja ot edy.* ‘The table was (lit. directly) groaning with food’

Each of the particles under discussion expresses its own set of meanings (cf. Cienki 1998 on some cross-linguistic differences within one subfield of the semantic domain under study). Most importantly, these meanings correlate with the original semantics of their lexical source, i.e. the difference in semantics of the source adjectives/adverbs influences the meaning of the derived particles. For example, our analysis has shown that the semantics of the emphatic similative is developed from the lexical meaning ‘immediate, having no intermediate states’ (*direct flight*), which accounts for the fact that the German *gerade* that does not have this meaning (in contrast to *direkt*), cannot be interpreted in such a way.

Thus, the typological and diachronic approach to the semantics of particles provides new opportunities for their further research. A detailed description of the source of grammatical meaning enables a researcher both to see the full semantic potential of these lexical items and to come up with a sound comparative analysis of particles in different languages.

References

- Cienki, A. 1998. STRAIGHT: *An image schema and its metaphorical extensions*. In *Cognitive Linguistics* 9-2, 107–149.
- Diewald, G. 2006. *Discourse particles and modal particles as grammatical elements*. In K. Fischer (ed.), *Approaches to discourse particles*. Amsterdam: Elsevier, 403-425.
- Heine B. & T. Kuteva 2002. *World Lexicon of Grammaticalization*. Cambridge.
- Traugott E.C. & R.B. Dasher, 2002. *Regularity in semantic change*. Cambridge.

Or and *or.be* alternative questions - An areal typology of the languages in China

The present work proposes an *or/or.be* typology in alternative questions. In languages like English and German, disjunctions used in declarative sentences and alternative questions demonstrate no difference, cf. English *or* and German *oder*. In Standard Chinese, however, disjunction *huozhe* 'or' is used in the declarative sentences and *haishi* (*or.be*) 'or' is used alternative questions.

In a survey of 138 languages in China, 32 languages are found to have a division in *or* and *or.be* in declaratives and alternative questions (*or.be* languages), 25 languages demonstrate no such division (*or* languages), and 81 languages are unknown or irrelevant. *or/or.be* merits a typology because some parameters are correlated with this distinction (at least) in many languages in China. For instance, the position of adpositions and the position of disjunctions are two parameters of this kind. In English, a pause can only happen after the first token, but before the disjunction (*A, disjunction B*), i.e. it is of disjunction-preposed type; whereas languages like Naxi (Tibeto-Burman, Sino-Tibetan) are of disjunction-postposed type (*A disjunction, B*). The following is the attested number of languages of the two parameters in *or*- and *or.be*-languages, respectively.

Table 1. The positions of disjunctions and adpositions in 25 *or*-languages

Disjunctions	Adpositions	Adpositions	Disjunctions
preposed 21	Pr 15, Po 6	Pr 15	preposed 15, postposed 0
postposed 4	Pr 0, Po 4	Po 10	preposed 6, postposed 4

Table 2. The positions of disjunctions and adpositions in 32 *or.be*-languages

Disjunctions	Adpositions	Adpositions	Disjunctions
preposed 29	Pr 22, Po 7	Pr 23	preposed 22, postposed 0, pre/postposed 1
postposed 2	Pr 0, Po 2	Po 9	preposed 7, postposed 2
pre/postposed 1	Pr 1, Po 0		

It can be seen that *or.be*-languages are more frequently to take prepositions than *or*-languages. More generalizations drawn from the tables above:

- (i) disjunction-preposed languages are more commonly to take prepositions than postpositions;
- (ii) disjunction-postposed languages are postpositional;
- (iii) most prepositional languages are of disjunction-preposed.

More parameters are examined in this work, in terms of the *or/or.be* typology.

The mapping of space onto the domain of benefaction: Beneficiaries that are not Recipients and their sources

The polysemy of Recipient and Beneficiary is cross-linguistically frequent, and is easily explained semantically (Zúñiga, Kittilä 2010:18-19); it often involves polysemy with allative and purpose (ib. 23-24; Schmidtke-Bode 2010). Diachronically, the following development is assumed:

- (1.) allative > beneficiary > recipient > purpose
(Rice, Kabata 2010; Heine, Claudi, Hünemeyer 1991).

Thus, allative provides the spatial source for Beneficiary and Recipient markers, and benefaction is conceptualized as a directional process through the mapping of space onto its domain.

However, Beneficiary markers display patterns of polysemy that exclude Recipient, notably with Cause/Reason, as in the case of Finnish *vuoksi*:

- (2.) *Henkilo opettel-i suome-a yksilo-n vuoksi.*
person.nom learn-3sg.pst Finnish-part individual-gen for
“A person learnt Finnish for an individual.”

- (3.) *Jaatelo sul-i sahkokatko-n vuoksi.*
ice.cream.nom melt-3sg.pst power.failure-gen for
“The ice cream melted because of the power failure.”

(Zúñiga, Kittilä 2010:22-23).

Moreover, the polysemy of Beneficiary and Purpose does not necessarily include Recipient, as shown in Georgian (M. Topadze, p.c.):

- (4.) *ertjeradi gamoq'eneb-is-tvis*
single usage-gen-for
“for a single usage”

- (5.) *es bavshv-is-tvis viq'ide*
this child-gen-for I-bought
“I bought it for the child.”

Notably, Recipient is most often not included in the polysemy of Beneficiary, Cause/Reason, and Purpose. A readily available example is provided by English *for*, German *für*, another example is Turkish *için* (V. Tören, p.c.):

- (6.) *Söylemek için geldim*
say-inf for come-pst-1sg
“I came in order to say...”

- (7.) *Bayram olduğu için toplar atıldı*
holyday be-pst.3sg for cannon-pl employ-pst-3sg
“Because of the holyday, cannons were shot.”

- (8.) *sizin için bir kitap getirdim*
2pl-gen for one book bring-pst-1sg
“I took a book for you.”

Diachronically, markers that display the above polysemy do not originate from allatives, but involve relations that are located elsewhere in the domain of space, such as locative or path: English *for* derives from Proto-Germanic **fura*, ‘before’ (the directional meaning is secondary); the local meaning of Finnish *vuoksi* is ‘through’. In this paper, I describe this pattern of polysemy cross-linguistically, provide the spatial sources for Beneficiary markers that do not extend to Recipient, and show how semantic extension proceeds. I highlight the difference between a directional and a non-directional conceptualization of benefaction. (I concentrate on the mapping of space onto the domain of benefaction, and do not consider markers that derive from serial verbs, such as those discussed in Creissels 2010, which display a different pattern of polysemy.)

References

- Creissels, Denis 2010. Benefactive applicative periphrases: A typological approach. In F. Zúñiga and S. Kittilä 2010 (eds.), 29-70.
- Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer. 1991. *Grammaticalization. A Conceptual Framework*. Chicago: University of Chicago Press.
- Rice, Sally and Kaori Kabata. 2007. Crosslinguistic grammaticalization patterns of the allative. *Linguistic Typology* 11: 451–514.
- Schmidtke-Bode, Karsten. 2010. The role of benefactives and related notions in the typology of purpose clauses. In F. Zúñiga and S. Kittilä 2010 (eds.), 121-146.
- Zúñiga, Fernando and Seppo Kittilä (eds.). 2010. *Benefactives and Malefactives. Case Studies and Typological Perspectives*. Amsterdam/Philadelphia: Benjamins.

Consonant harmony as spreading: an evaluation

A relatively large number of the world's spoken languages display some form of the typological property known as vowel harmony in which vowels in neighboring syllables are required to agree in one or more properties, such as vowel height, rounding, or tongue root position. A rather smaller number display a pattern of consonant harmony in which non-adjacent consonants are either required to or tend to share certain of the place, phonatory or other attributes of nearby consonants. Ohala (1990) claimed that vowel harmony, at least in its origin, is a product of vowel-to-vowel assimilation across intervening consonants. Later Gafos (1999) essentially argued that consonant harmony may similarly be assimilatory in origin. For this to be the case, the segments that intervene between affected consonants — typically vowels — must be capable of transmitting the harmonizing property. For some properties, such as nasality or liprounding, such 'spreading' is non-problematic as these can be properties of either consonant or vowels. An alternative view, e.g. in Hansson (2010), is that consonant harmony (although this term is more narrowly defined in Hansson's usage) is a correspondence or copying process, not an assimilatory effect. That is, the harmonizing property involves the independent repetition of an articulatory gesture (or gestures) rather than being the result of the anticipatory or perseverative prolongation of a single gesture. In this paper a range of attested varieties of consonant harmony will be evaluated in terms of how plausibly an assimilatory component might be involved in their origin. The analysis indicates that consonant harmony patterns vary along a scale of their likelihood to be explicable as assimilatory in nature. Nasal consonant harmony most likely is always (at least originally) triggered by nasal coupling across intervening vowels, although it may be generalized to apply where non-adjacent segments are affected, as in Sundanese (Cohn 1989). Processes such as sibilant harmony — where typically alveolar and palatoalveolar fricatives are not permitted to co-occur — may have an assimilatory component, as suggested by Whalen et al (2011) in relation to Tahltan. This idea is supported by a limited acoustic study reported here, showing that vowel formants in /a/ surrounded by /s/ differ significantly from the formants of this vowel surrounded by /ʃ/. This suggests that some aspects of the different tongue configurations in /s/ and /ʃ/ can be transmitted through a vowel. However, consonant harmony involving certain phonatory and laryngeal features, such as voicing (given that vowels are prototypically already voiced) or ejective production, which cannot be a property of vowels, does not plausibly involve assimilatory transmission of the harmonizing property. The typology of consonant harmony should therefore be accounted for in terms of an interplay of the effects of historical assimilatory processes and other, more cognitively

Theme session: Typological hierarchies in synchrony and diachrony

The Alignment System Changes in Tāti Language Group

The morphosyntax associated with past transitive verbs in most Iranian languages differs from that associated with others. Iranian languages weren't always like this. The Old Iranian period, spoken around three thousand years ago, had a unified accusative alignment in all tenses; however, all the languages attested from the Middle Iranian period (4-3 century BC. to 8-9 century AD.) onwards are characterized by tense-sensitive alignment. The historical evidence shows that all modern Iranian languages must have passed through the tense-sensitive alignment in which the verbal agreement was with an O-past and the case system was an ergative one, while it was nominative-accusative in all other environments. The ergative alignment in Middle Iranian has been preserved in modern Tāti, one of the Iranian languages spoken in north-west of Iran. The different alignments in Tāti dialects are defined using two parameters: 1) The case marking of core arguments, 2) The formal means of cross-referencing core arguments outside of the NPs. Agreement with core arguments in some Tāti dialects is via clitics on other constituents. The alignment in Tāti is tense-sensitive. In present stem verb sentences, all Tāti dialects follow the nominative-accusative case marking which marks the subject of intransitive and transitive verbs with a direct case marker (the morphologically unmarked case), and both determine agreement on the verb; in present tense, object of transitive verbs is marked with an oblique case marker, -e, and plays no role in person agreement with the verb. e.g.

1. Ahmad hasan-e mivine. (Tākestāni dialect of Tāti)
P.N. P.N.-Acc. see.Pres:3sg
'Ahmad sees Hasan.'

In clauses headed by verb forms built with the past stem of transitive verbs, the situation differs in that some Tāti dialects have retained the ergative-absolutive alignment system of the Middle Iranian. In these dialects, subject of transitive verb is marked with the distinct oblique marker, -e, while the object and subject of intransitive verbs are marked with the direct case marker, and both determine agreement on the verb, e.g.

2. Ahmad-e Hasan buind. (Eshtehārdi dialect of Tāti)
P.N.-ERG. P.N. see.Pst:3sg
'Ahmad saw Hasan.'

The aim of this paper is to show changes in alignment system of past-stem transitive sentences in the other Tāti dialects, e.g.

3. Ahmad Hasan-eš bəkəšt. (Tākestāni dialect of Tāti)
P.N. P.N.-Clt:3sg kill.Pst
'Ahmad killed Hasan.'

In past-stem transitive sentences of these dialects, both subject and direct object are in direct case but there is a clitic attached to the object which cross-references to subject. The loss of ergative marker on the subject can be explained by economic motivation. Since the clitic shows the role of subject, there is no need for the overt ergative marker. This loss can also be attributed to contact with dominant Persian language which follows nominative-accusative in all tenses. The other important change is in the verb which remains the same with all persons and shows agreement with neither the subject nor the object.

Theme session: Lexical typology of qualitative concepts

The language of perception across cultures

How are the senses structured by the languages we speak, the cultures we inhabit? To what extent is the encoding of perceptual experiences in languages a matter of how the mind/brain is “wired-up” and to what extent is it a question of local cultural preoccupation? The “Language of Perception” project tests the hypothesis that some perceptual domains may be more “ineffable” – i.e. difficult or impossible to put into words – than others. While cognitive scientists have assumed that proximate senses (olfaction, taste, touch) are more ineffable than distal senses (vision, hearing), anthropologists have illustrated the exquisite variation and elaboration the senses achieve in different cultural milieus. The project is designed to test whether the proximate senses are universally ineffable – suggesting an architectural constraint on cognition – or whether they are just accidentally so in Indo-European languages, so expanding the role of cultural interests and preoccupations.

To address this question, a standardized set of stimuli of color patches, geometric shapes, simple sounds, tactile textures, smells and tastes have been used to elicit descriptions from speakers of over twenty languages—including three sign languages. The languages are typologically, genetically and geographically diverse, representing a wide-range of cultures. The communities sampled vary in subsistence modes (hunter-gatherer to industrial), ecological zones (rainforest jungle to desert), dwelling types (rural and urban), and various other parameters. We examine how codable the different sensory modalities are by comparing how consistent speakers are in how they describe the materials in each modality, by examining the types of responses they produce, and elucidating how the terms produced carve up each sensorial domain. Our analyses suggest that the codability of the senses is highly variable across languages. Moreover, we have identified exquisite elaboration in the olfactory domain in some cultural settings, contrary to some contemporary predictions within the cognitive sciences. Lexical elaboration of olfaction is particularly intriguing because, aside from the challenge they present to notions of ineffability and the limits of language, they also pose new problems for the lexical typology of qualitative concepts and the study of semantics. Taken together, the results of the language of perception project suggest that differential codability may be, at least partly, the result of cultural preoccupation. This shows that the senses are not just physiological phenomena but are constructed through linguistic, cultural and social practices.

Theme session: Predicate-centered focus from a cross-linguistic perspective

Predicate focus and realis mood in Tundra Yukaghir and beyond

The purpose of this talk is to investigate the cross-linguistic validity of the category of predicate focus against the background of a moderate version of contextualism, the view according to which much of what we are used to think of as linguistic meaning is a product of pragmatic inference. Even if we assume that a certain meaning must be present in a certain context, such as focus in an answer to a question, it still does not follow that a linguistic category of focus must be involved: languages may underspecify the relevant meaning and leave it to the hearers' inferential abilities to reconstruct it. Alternatively, the meaning in question may happen to coincide, fully or in part, with another meaning. (Matić & Wedgwood 2013). This seems to be relatively frequent with the categories of predicate focus and realis mood.

In this paper, we provide evidence for the expression of predicate focus via realis mood markers. The pragmatic principles on which this semantic affinity is based are explained through a detailed analysis of the Tundra Yukaghir (TY) proclitic $m\bar{\alpha}(r)=$. This verbal particle has been described in the literature as a marker of declarative illocutionary force, positive polarity, and/or predicate focus. We show that these approaches are inadequate and propose a modal analysis of $m\bar{\alpha}(r)=$. In our analysis, $m\bar{\alpha}(r)=$ is a realis marker, a quantifier which existentially binds propositions and provides for their existential anchoring. This is clearly shown by the incompatibility of $m\bar{\alpha}(r)=$ with non-realis contexts, directives and hypotheticals, and its obligatoriness whenever a realis meaning is intended. Nominal foci carry their own existential import and are marked with a special set of morphemes, so that they are exempt from obligatory marking with $m\bar{\alpha}(r)=$; if there is no nominal focus, $m\bar{\alpha}(r)=$ is attached to the verb by default. We show that this analysis is supported by a number of syntactic and semantic arguments.

We claim that the connection of realis mood, which in effect marks the proposition as true, and predicate focus, which asserts the existence of a state of affairs with respect to an entity, is not accidental. We adduce evidence from other languages, notably Somali and a couple of other Cushitic languages, Basque, Hungarian and Ancient Egyptian, which indicates that the formal affinity between realis and predicate focus is a cross-linguistically well-attested pattern, such that realis markers can regularly occur in contexts in which the discourse requires focus on the predicate. Thus, the category of predicate focus is, in some languages at least, based on an inference of the existence of a state of affairs derived from the use of the realis mood, and not a discrete encoded meaning in its own right.

References

Matić, Dejan & Daniel Wedgwood. 2013. The meanings of focus. The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49.3.

Theme session: Generalized Noun Modifying Clause Constructions

Noun Modifying Clause Constructions in Besemah Malay

Traditional treatments of Malay-Indonesian relative clauses describe the constructions as (a) headed by the relativizer *yang*, (b) relativizing only the subject position, and (c) optionally occurring without a head noun (i.e., headless relative clauses). Although some of these properties have been called into question (see Ewing & Cumming 1998, Cole & Hermon 2005, Tjung 2006, Englebretson 2008), analogous constructions headed by *ndik* and *nde* in Besemah Malay, a little-known language of southwest Sumatra, significantly diverge from this characterization.

First, *ndik* and *nde* constructions express an array of meanings that are not typically expressed by *yang*-constructions, including possession and purpose. Second, while a majority of the cases of relative clauses are headless in (colloquial) Indonesian, there are no instances of *ndik* and *nde* with an external head noun. Finally, *ndik* and *nde* constructions do not show strict syntactic restrictions on the relativization of subjects as in many Austronesian languages (Keenan & Comrie 1977). In example (1), the referent is not the unrealized patient argument (i.e., the people that are pushed), but the third person non-subject agent *-nye* (i.e., the person pushing everyone aside). What is more, *ndik* and *nde* constructions occasionally surface with all arguments realized, taking a different semantic interpretation than would be expected in a relative clause construction, expressing manner in (2) and time in (3).

- (1) *Ui [nik Ø di-pinggir-kah-nye gale] tadi,*
 oh NIK UV-edge-APPL-3 all earlier
 ‘Oh, **(the one)** [that (everyone) was pushed to the side **by him**] earlier,’
- (2) *Aku nik ridat [nik die galak be-antuk-an tu=lah.]*
 1SG NIK angry NIK 3 want RECP-nudge-RECP that=LAH
 ‘I am the one who is angry **(at how)** [he is always running into people.] ’
- (3) *[nik kite main di Sumur.]*
 NIK IPL.INCL play LOC S.
 ‘**(the time)** [we were playing in Sumur.] ’

Based on a corpus of spontaneous speech of Besemah Malay, this study shows that *ndik* and *nde* constructions are best analyzed as Generalized Noun Modifying Clause Constructions (GNMCC; Matsumoto 1997, Comrie 1998). This study has further implications for the typology of GNMCC. First, it extends the current areal and genealogical distribution of GNMCC to an Austronesian language of insular Southeast Asia. Second, this study demonstrates how strict syntactic constraints on voice (or focus) for relativization in some Austronesian languages are in fact pragmatically conditioned in *ndik* and *nde* constructions in Besemah Malay.

References

- Cole, Peter & Gabriella Hermon (2005) Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics* 14(1): 59{88.
- Comrie, Bernard (1998) Rethinking the Typology of Relative Clauses. *Language Design* 1: 59{85.
- Englebretson, Robert (2008) From subordinate clause to noun-phrase. *Crosslinguistic studies of clause combining: the multifunctionality of conjunctions* .

- Ewing, Michael C. & Susanna Cumming (1998) Relative Clauses in Indonesian Discourse: Face to Face and Cyberspace Interaction. In Papers from the 15th annual meeting of the Southeast Asian Linguistics Society, Shobhana L. Chelliah & Willem J. de Reuse, eds., Tempe, AZ: Program for Southeast Asian Studies, Arizona State University, 79-96.
- Keenan, Edward L. & Bernard Comrie (1977) Noun phrase accessibility and universal grammar. *Linguistic inquiry* 8(1): 63-99.
- Matsumoto, Yoshiko (1997) Noun-modifying constructions in Japanese: A frame-semantic approach, vol. 35. John Benjamins Publishing Company.
- Tjung, Yassir (2006) The formation of relative clauses in Jakarta Indonesian: A subject-object asymmetry. Ph.D. thesis, University of Delaware.

Possessor Dissent: Oblique possessive constructions in Bilinarra

Bilinarra is an Australian language belonging to the Eastern Ngumpin subgroup of the Ngumpin-Yapa family (Pama-Nyungan), which includes Warlpiri and Gurindji (McConvell & Laughren, 2004).

Bilinarra, like most Australian languages, has two distinct possessive constructions, one which encodes alienable possession (1), and another which encodes inalienable possession (2). These constructions differ in terms of grammatical relations: in (1), the possessed NP is cross-referenced with the bound object/oblique pronoun, and the possessor is marked as a dative modifier. In the inalienable construction in (2), on the other hand, both nominals appear in the unmarked (accusative) case and it is the possessor which is cross-referenced.

(1) nyawa=ma=∅_a [nyununy gurrurij=ma_b]_a?
 this=TOP=3MIN.O 2MIN.DAT car=TOP
 Is that your car?

(2) nyawa=ma=nggu_b nyundu_b mila=ma_a?
 this=TOP=2MIN.O 2MIN eye=TOP
 Is that your eye?

Bilinarra in fact also has a third possessive construction, which has not previously been identified in the Australianist literature, and which is the focus of this paper. This *oblique possessive construction* encodes alienable possession yet appears to mix the grammatical features of the two possessive constructions described above: cross-referencing is with the possessor (as in (2)), but the possessor appears marked with the dative case (as in (1)). An example is given in (3).

(3) nyawa=ma=nggu_b nyununy_b gurrurij_a=ma?
 this=TOP=2MIN.O 2MIN.DAT car=TOP
 Is that a car of yours?

Oblique possessive constructions in Bilinarra are used to highlight the affectedness of the possessor in alienable constructions. Syntactically the possessor and possessum remain one NP yet the possessor modifier is cross-referenced, rather than the nominal head (cf. (1)). The construction is enabled by the fact that object/oblique bound pronouns are used to mark adjuncts in benefactive, malefactive and animate goal constructions. Thus, these bound pronouns are available to be reanalysed as a marker of 'affected participant' and extended into alienable possession constructions where the speaker wants to highlight the effect on the possessor.

In this paper we discuss the properties of the oblique possessive construction in Bilinarra and surrounding languages within the context of benefactives/malefactives and possessive constructions cross-linguistically (e.g. Lichtenberk 2002, Zúñiga & Kittilä 2010, Rapold 2010). This work has implications for the typology of Australian languages, for which this construction type has previously not been discussed (cf. Dixon 2002:394ff), as well as for our understanding of dative constructions more broadly (cf. Bosse et. al 2012).

References

- Bosse, Solveig, Benjamin Bruening & Masahiro Yamada. Affected experiencers. *NLLT* 30(4): 1185-1230.
 Dixon, R. M. W. 2002. *Australian languages: their nature and development*. Cambridge: CUP.
 Lichtenberk, Frantisek. 2002. The possessive-benefactive connection. *Oceanic Linguistics* 41(2): 439-474.

- McConvell, Patrick & Mary Laughren. 2004. Ngumpin-Yapa Languages. In Harold Koch & Claire Bowern (eds.), *Australian Languages: Reconstruction and Subgrouping*. Amsterdam: Benjamins. 151-77.
- Meakins, Felicity & Rachel Nordlinger. to appear. *A Grammar of Bilinarra, an Australian Aboriginal Language of the Victoria River District (NT)*. Berlin: Mouton de Gruyter.
- Rapold, Christian J. 2010. Beneficiary and other roles of the dative in Tashelhiyt. In Fernando Zúñiga & Seppo Kittilä (eds), 351-376.
- Zúñiga, Fernando & Seppo Kittilä (eds). 2010. *Benefactives and malefactives: Typological perspectives and case studies*. Amsterdam: Benjamins.

Motion constructions in contact languages: cross-linguistic evidence from APiCS

Data on a wide variety of contact languages from the Atlas of Pidgin and Creole Language Structures (APiCS, Michaelis et al. 2013) show that intransitive and transitive motion constructions often show identical marking of goal (motion---to) and source (motion---from). This is rather different from the situation in the European lexifiers. In French, for instance, the marking of motion---to (*dans* 'in, into') is different from motion---from (*de* 'from') in both intransitive and transitive constructions:

- | | |
|---|---|
| <p>(1) a. intransitive motion-to
 <i>Je vais dans la cuisine.</i>
 'I go into the kitchen.'</p> | <p>b. intransitive motion-from
 <i>Je sors de la cuisine.</i>
 'I go out of the kitchen.'</p> |
| <p>(2) a. transitive motion-to
 <i>Je pousse le carton dans la cuisine.</i>
 'I push the box into the kitchen.'</p> | <p>b. transitive motion-from
 <i>Je tire le carton de la cuisine.</i>
 'I pull the box out of the kitchen.'</p> |

But in a number of contact languages in APiCS the four situations corresponding to (1a-b) and (2a-b) are marked in exactly the same way irrespectively of orientation. In Seychelles Creole (Michaelis & Rosalie 2013), for instance, all four constructions are marked by the preposition *dan*:

- | | |
|---|--|
| <p>(3) a. <i>mon al dan bwa</i>
 1SG go in forest
 'I go into the forest.'</p> | <p>b. <i>mon sorti dan bwa</i>
 1SG come.from in forest
 'I come out of the forest.'</p> |
| <p>(4) a. <i>Marcel ti pus Peter dan trou</i>
 Marcel PAST push Peter in hole
 'Marcel pushed Peter into the hole'.</p> | <p>b. <i>Marcel ti redi Peter dan trou</i>
 Marcel PAST pull Peter in hole
 'Marcel pulled Peter out of the hole.'</p> |

Here the hearer must rely on the semantics of the verb to infer the correct interpretation. Other contact languages represented in APiCS that show the same or very similar polysemous marking patterns are, for instance, the French-based Caribbean creoles, Early Sranan, African contact languages (Krio, Sango, Lingala, Fanakalo), the two other French-based Indian Ocean creoles (Mauritian and Reunion Creole), several Chabacano varieties of the Philippines, as well as Tok Pisin and Bislama in the Pacific. Such cases of source-goal nondistinctness have occasionally been discussed for other languages (Lehmann 1992, Wälchli & Zúñiga 2006), but their widespread occurrence in creole and other contact languages is a new finding. After presenting data from a range of languages and a world map of 76 contact languages, we will argue tentatively that these marking patterns are not due to simplification strategies during the process of pidginization or creolization, but to substrate influence: The speakers of the various substrate languages (slaves or indentured labourers) have retained these patterns from their native languages in the developing contact languages.

References

- Lehmann, Christian. 1992. Yukatekische lokale Relatoren in typologischer Perspektive. ZPSK 45(6). 626–641.
- Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath, & Magnus Huber (eds.) 2013. Atlas of Pidgin and Creole Language Structures. Oxford: OUP, to appear.
- Wälchli, Bernhard & Fernando Zúñiga. 2006. Source- Goal (in)difference and the typology of motion events. STUF 59(3). 284–303.

The marking of nominal participants under negation

Negatives that show structural differences with respect to their affirmative counterparts in addition to the marker(s) of negation are called asymmetric by Miestamo (2005). The marking of NP participants is explicitly excluded from the scope of his study, and no systematic typological knowledge is thus available on the asymmetry between affirmatives and negatives as regards the marking of nominal participants. The aim of this study is to fill this gap. The effects of negation on the marking of nominal participants are surveyed in an areally and genealogically balanced sample of 240 languages.

Relevant effects of negation are found in case marking, in the use of articles and other determiners, in the use of class markers, and in focus marking. A comprehensive picture of the asymmetries found in the marking of nominal participants is presented in the paper. The types of asymmetry are cross-linguistically rather uncommon, being limited to specific areas and families. #

It is well-known that negation can affect case marking. In Finnish affirmatives, the object may be, on the one hand, in the genitive or nominative depending on the morphosyntactic environment, or, on the other hand, in the partitive, but in negatives only the partitive is possible. Similar case asymmetry is found in a number of European languages (Finnic, Baltic, Slavic, Basque): NPs in the scope of negation are marked with a case that has a partitive-marking function (partitive or genitive), either obligatorily or as a matter of preference. Such case alternations are however not found outside Europe in my sample.

Negation is found to affect the use of articles and other determiners, e.g., in French and in some Oceanic languages, such as Araki. In Araki (Alex François, pc.), realis affirmatives have bare NP objects and the verb bears a referential object marker and person-number crossreference. The object may be further specified as indefinite by the specific-indefinite marker *mohese*. In the negative, there is no cross-reference on the verb and the object is marked by the nonspecific (partitive) indefinite marker *re*. Referential marking and cross-reference are possible on the verb in negatives, too, but then the reading is definite, and *re* does not occur. The specific indefinite marker *mo-hese* is impossible in negatives and the non-specific indefinite *re* is impossible in realis affirmatives.

Class markers are affected by negation, e.g., in some Bantu languages. In Xhosa (Taraldsen 2010), noun class prefixes have a full form in affirmatives. In negatives, a shorter form of the class prefix can be used and the NP then gets an indefinite non-specific reading. The full form can be found in negatives if the object prefix also appears on the verb, i.e. when the object is definite (or, more rarely, specific indefinite).

Effects of negation on focus marking are found, e.g., in a number of African languages, in which negatives are treated as inherently focused. In Aghem (Larry Hyman, pc.), this results in marking the NP in the scope of negation as oblique.

The types of asymmetries between affirmation and negation found, e.g., in Araki and Xhosa are connected to the referentiality of the NPs, and can be explained by the discourse properties of negation (cf. Givón 1978). I propose that the partitive of negation is also connected to referentiality and motivated in the same way. The functional motivations for the asymmetry affecting focus marking will also be addressed in the paper.

References

- Givón, Talmy. 1978. Negation in language: Pragmatics, function, ontology. In *Syntax and Semantics. Vol. 9. Pragmatics*, Peter Cole (ed.), 69–112. New York: Academic Press.
- Miestamo, Matti. 2005. *Standard negation: The negation of declarative verbal main clauses in a typological perspective*. Berlin: Mouton de Gruyter.
- Taraldsen, Knut Tarald. 2010. The nanosyntax of Nguni noun class prefixes and concords. *Lingua* 120: 1522–1548.

Theme session: Typological Universals and Diachrony

Deconstructing Teleology: The place of synchronic usage patterns among processes of diacronic development

Views vary on the role of implicational hierarchies in shaping language. One is that hierarchies are primary, described by Corbett (2011): 'Since a hierarchy constrains what is a possible language, it is also a constraint on language change, because languages move from one possible state to another'. Another is the reverse, characterized in the theme session description: hierarchies simply 'capture the outputs of independent diachronic processes.' Here it is argued that explaining the typological patterns we find depends crucially on understanding the various kinds of interactions that can hold between hierarchies and processes of language change.

The point of departure is the most frequently-cited set of hierarchies, the animacy/topicality/referential hierarchies. They take forms like the following:

speaker > hearer > 3rd person pronouns > kinsmen > other humans > higher animals > lower animals > inanimates

(Additional features such as definiteness and count may also come into play.) Hierarchies from this set have been implicated in a number of areas of structure. One is differential number marking: if a language distinguishes number on forms at any point in the hierarchy, it will also differentiate number on all categories to the left (first noted in Smith-Stark 1974). Another is reference within the verb: if the referent of some form on the hierarchy is marked on the verb by pronominal or agreement affixes, so too will all forms to its left on the hierarchy (noted early in Moravcsik 1974). A third is case marking: if members of a category carry accusative case marking, so will all categories to its left; if members of a category carry ergative case marking, so will all categories to its right (Silverstein 1976). A fourth, of a slightly different type, involves hierarchical systems: if only one argument is identified by marking in the verb, the leftmost on the above scale will take precedence over all those to its right (with the caveat that in many systems, the order is 2nd persons > 1st) (Mithun 2012 and others).

Here it will be shown that these patterns display a variety of relationships between the hierarchies and diachrony, with examples from a range of families including Austronesian, Wakashan, Salishan, Chimariko, Yana, Pomoan, Siouan, Uto-Aztecan, and Iroquoian. In some cases, a need felt by speakers can drive the development of number distinctions over time. Evidence comes from language contact: bilinguals accustomed to number distinctions in one of their languages may create them in the other, through reanalysis of existing native forms. The forms themselves are drawn from a variety of sources, sources which determine the entry point for the distinction. In other cases, animacy/topicality/referential hierarchies play a more indirect role. They reflect recurring patterns of expression, which in turn serve as a foundation for the routinization of grammaticalization. Such developments can involve several different kinds of processes revolving around topicality, among them a propensity for passivization and the use of pronouns for given (topical) participants as opposed to full lexical expressions.

References

- Corbett, Greville 2011. *Implicational hierarchies*. *Handbook of Language Typology*. Jae Jung Song, ed. Oxford.
- Mithun, Marianne 2012. *Core argument patterns and deep genetic relations*. *Typology of Argument Structure and Grammatical Relations*. P. Suihkonen, B. Comrie, and V. Solovyev eds. *Studies in Language Companion Series*. Amsterdam: John Benjamins. 257-294.
- Moravcsik, Edith 1974. *Object-verb agreement*. *Working Papers in Language Universals* 15:25-140.
- Silverstein, Michael 1976. *Hierarchy of features and ergativity*. *Grammatical categories in Australian languages*. R.M.W. Dixon, ed. 112-71. Canberra: Australian Institute of Aboriginal Studies.
- Smith-Stark, T. Cedric 1974. *The plurality split*. *Papers from the 10th Regional Meeting of the Chicago Linguistic Society*. M. La Baly, R. Fox, and A. Bruck eds. 657-71.

Compound structures in alternate sign languages: new insights from Khoisan hunting signs

While the field of sign language typology remains sparsely researched, recent studies have established a dichotomy of “primary” and “alternate” sign languages (Kendon 2004; Zeshan 2008; Pfau 2012). Primary sign languages are full-fledged systems acquired by deaf people as their L1, while alternate sign languages are “kinesic codes” (Kendon 2004), developed by the hearing members of a speech community for use in special circumstances that preclude vocal communication. Among the latter are monastic sign languages, the Aboriginal sign languages of Australia, and Plains Indian Sign Language used among the native American groups of the Great Plains region. All languages of this type show certain similarities in structure. Recently, hunting signs used among several Kalahari Khoe-speaking groups of Southern Africa have been found to belong to the type of alternate sign systems (Fehn & Mohr 2012).

This paper analyses compound structures in a subset of five alternate sign systems. While the introductory analysis discusses differences and similarities of compound structures in those alternate sign languages linguistically documented, the paper also includes recent findings from Khoisan hunting signs (Mohr & Fehn 2012) used among the Tsǀxixa and ||Ani. The characteristics analysed are syntactical/phonological structure, morphological composition and semantic set-up. Compounds in primary sign languages phonologically and syntactically appear as one sign (Meir 2012). Although alternate sign language compounds appear to form one unit syntactically, i.e. their inner order cannot be reversed (Pfau 2012), phonological assimilation and reduction processes cannot be observed (cf. (1) from Pfau 2012).

- (1) HARD^WATER [CisSL]
„ice”

Therefore, the compound could not only be interpreted as „ice” but also as „hard water”. This seems to be a major distinguishing criterion from primary sign language compounds.

Concerning the morphological structure of the signs, compounds in Aboriginal sign languages (Kendon 1988), e.g., are usually bimorphemic, Plains Indian Sign Language exhibits trimorphemic compounds among others (Davis 2010), while monastic sign languages exhibit compounds consisting of up to nine parts (cf. (2) from (Barakat 1975).

- (2) SECULAR+TAKE+THREE+O+WHITE+MONEY+KILL+CROSS+GOD [CisSL]
„Judas”

Khoisan hunting sign compounds are either bi- or trimorphemic (cf. (3) from ||Ani hunting signs).

- (3) BEAK-CROOKED+DOT.PL+WINGS [||uen]
„guinea fowl”

Finally, Kendon (1988) mentions universal semantic patterns for several alternate sign languages, such as an internal ordering from the constituent with the most general meaning as the initial component to the constituent with the most specific meaning as the final one. This structure also applies to the Khoisan hunting signs, with slightly different orderings for mammal and bird referents (cf. (4) and (5)).

- (4) Mammals: [CL] horns/ears + fur pattern
- (5) Birds: classifier/beak [+ pattern of feathers] + wings

In conclusion, compounds across alternate sign languages show similar structures with respect to the phonological, syntactic, morphological and semantic level. These results are central for the young field of sign language typology. Firstly, they confirm the classification of the Khoisan hunting signs as alternate sign system. Secondly, they strengthen the establishment of an alternate sign language class as opposed to a primary sign language type based on structural grounds.

References

- Barakat, R. 1975. *The Cistercian sign language: a study in non-verbal communication*. (Cistercian Study Series 7) Kalamazoo, Mich.: Cistercian Publications.
- Davis, J. 2010. *Hand Talk: Sign Language Among American Indian Nations*. Cambridge: Cambridge University Press.
- Fehn, A.-M. & Mohr, S. 2012. Phonologie ikonischer Tiergesten bei zwei verwandten Kalahari Khoesprachigen Gruppen (Tsǀxixa und ||Ani-Khwe). Paper presented at the 12. Afrikanistentag, Köln.
- Kendon, A. 1988. *Sign Languages of Aboriginal Australia: Cultural, Semiotic and Communicative Perspectives*. Cambridge: Cambridge University Press.
- Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Meir, I. 2012. Word classes and word formation. In Pfau, R., Steinbach, M. and Woll, B. (eds.) *Sign Language. An International Handbook*. Berlin/New York: De Gruyter Mouton, 78-112.
- Mohr, S. & Fehn, A.-M. 2012. CL-bigbird + BEAK-crooked + WINGS – Compounding strategies in hunting signs of the ||Ani-Khwe in Botswana. Paper presented at the Kölner Afrikawissenschaftliche Nachwuchstagung IV, Köln.
- Pfau, R. 2012. Manual communication systems: evolution and variation. In Pfau, R., Steinbach, M. and Woll, B. (eds.) *Sign Language. An International Handbook*. Berlin/New York: De Gruyter Mouton, 513-551.
- Zeshan, U. 2008. Roots, Leaves and Branches – The Typology of Sign Languages. In Quadros, R. M. de (ed.) *Sign Languages: Spinning and Unraveling the Past, Present and Future. 45 Papers and 3 Posters from the 9th Theoretical Issues in Sign Language Research Conference*. Petrópolis: Editora Azul, 671-695.

Theme session: Linked Data in Linguistic Typology

Typology with graphs and matrices

In this talk we demonstrate how to leverage Semantic Web technologies to transform data in any number of typological databases, e.g. WALS (Haspelmath et al, 2008), Autotyp (Bickel & Nichols, 2002), PHOIBLE (Moran, 2012), ODIN (Lewis, 2006) or individual databases -- along with metadata from Ethnologue (Lewis, 2009), LLMAP (ILIT), Multitree (ILIT, 2009) and Glottolog (Nordhoff et al, 2012) – into Linked Data. This is the vision of the Linguistic Linked Open Data Cloud (LLOD; Chiarcos et al, 2012).

Once data from these databases are converted into a homogeneous format, i.e. RDF graph data structures, the contents of these disparate datasets can be merged into one large graph, which allows for their data to be queried in a federate search fashion, in line with how we currently search the content of the Web through popular search engines.

We will illustrate how a user can query and retrieve information about a particular language, from multiple databases, via a language's ISO 639-3 code. For example, a user might be interested in accessing all typological variables described by various databases for a particular language, e.g., word order data from WALS, genealogical information and phonological word domains from Autotyp, and phoneme inventory data from PHOIBLE, etc.

We show how the results of such queries can be combined and output into a matrix format that mirrors recent work in multivariate typology (cf. Witzlack-Makarevich, 2011; Bickel, 2011). By outputting the results of users' queries across different databases into table-based matrix formats, the results can be directly loaded into statistical packages for statistical analyses, and published algorithms can be directly applied to them and tested, e.g. statistical sampling procedures (cf. Cysouw, 2005) and statistical approaches to determine universal preferences, e.g. Family Bias Theory (Bickel, 2011). Furthermore, when typological data are output into tables, state-of-the-art approaches using linear algebra to transform matrices into new datasets can be applied (Mayer & Cysouw, 2011).

Estimating diachronic preferences of phonological features cross-linguistically

In this paper we apply Family Bias Theory (FBT, Bickel 2011, in press) to the PHOIBLE phonological typology database (Moran 2012) to investigate the diachronic developments of phonological features in the world's languages. The idea behind FBT is that the synchronic distribution of a typological variable within language families allows estimating systematic preferences or dispreferences (i.e. biases in certain directions) in the history of the variable under various conditions, such as linguistic areas. For this, it is necessary that families are sampled densely (as is the case with PHOIBLE), but FBT also includes statistical extrapolation to small families and isolates.

Using FBT's statistical approach, we look for universal (dis)preferences in the diachronic development of certain phonological features in PHOIBLE, which contains a broad sample of over 1600 segment inventories and a set of 33 distinctive features expanded from Hayes (2009) to attain full typological coverage of cross-linguistic segment types. We use a dimensionality reduction algorithm to identify the language-specific system of distinctive features that is minimally required to encode each language's inventory of contrastive segments. This approach is in line with an emergentist theory of features (cf. Mielke 2004), which in contrast to features being innate and part of UG, posits features as learned and language-specific. We then use FBT to detect universal and area-specific (dis)preferences in the resulting feature systems.

Results: we identify universal biases towards the features [continuant, dorsal, front, high, labial, low, nasal, period glottal source, sonorant, syllabic, coronal] and against [consonantal, back, spread glottis, long, tap, labiodental, short, retracted tongue root, advanced tongue root, fortis, lowered larynx implosive, raised larynx ejective, round] being required to capture contrasts in individual languages. These preferences are statistically independent of geographic area and are largely in line with notions of universal perceptual saliency, while features like [consonantal] are dispreferred because they are mostly predictable from other features in the system.

One third of the features with universal biases also so show additional (but not interacting) differences across areas [period glottal source, sonorant, syllabic, coronal; lowered larynx implosive, raised larynx ejective, round]. Features with no universal biases all show area-specific biases [distributed, strident, trill, anterior, tense, lateral, constricted glottis, approximant].

We conclude that most of the 33 features are subject to strong universal constraints in their development, while about 50% appear to have spread in certain areas, some with and some without being subject to universal constraints in addition.

References

- Bickel, B. 2011. Statistical modeling of language universals. *Linguistic Typology* 15, 401-414.
- Bickel, B. In press. Distributional biases in language families In: Bickel, B., L. A. Grenoble, D. A. Peterson, & A. Timberlake (eds.) *Language typology and historical contingency: a festschrift to honor Johanna Nichols*. <http://www.spw.uzh.ch/bickel-files/papers/stability.fsjn.2011bickelrevised.pdf>
- Hayes, B. 2009. *Introductory Phonology*. Blackwell.
- Mielke, J. 2004. *The Emergence of Distinctive Features*. PhD thesis, The Ohio State University.
- Moran, S. 2012. *Phonetics Information Base and Lexicon*. PhD thesis, University of Washington.

Theme session: Quantitative Linguistic Typology

Word Order Variants and Ellipsis of Subjects and Objects in English, Evenki, Chukchee and Russian folklore narrative texts (quantitative approach)

It was known decades ago that word order (WO) in Russian and Chukchee declarative sentences is more or less free, whereas in English and Evenki it is more or less fixed. Less is known about the possibility of omitting the main nominal parts of the sentence (subjects and direct objects) in the above mentioned languages although English as a non-pronoun-drop-language is opposed to all other three languages under consideration.

Only quantitative analysis of vast text samples can help define (1) the degree of flexibility vs. rigidity of WOs, and (2) the ratio of elliptic constructions in folklore texts of the languages in question. The data for the quantitative analysis included text samples comprising 1700 English transitive and intransitive finite verb forms (FVF), about 1000 Evenki FVFs, about 1870 Russian FVFs and about 2500 Chukchee FVFs (the source texts are given in the references). Only narrative fragments of the source texts were taken into consideration which means that questions and imperatives were excluded from the analysis. Since only constructions with V_{fin} (VF) were calculated, participial and converbal forms and constructions were also disregarded.

There are eight theoretically possible WO variants of non-elliptic (NE) constructions – two with intransitive verbs (SV and VS) and six with transitive verbs; plus five theoretically possible elliptic (EL) constructions (omitted elements are put in brackets) – SV(O) / VS (O); (S)OV / (S)VO and (SO)V. All these types are discussed in detail in the presentation.

The quantitative analysis of these types gave the following results (the poster will include the full table of variant frequencies):

1. The ratio of NE vs. EL constructions in four samples: English -- 82 % (NE) – 18 % (EL); Evenki – 46.2 % (NE) – 53.8 % (EL); Chukchee – 33.2 % (NE) – 66.8 % (EL); Russian – 44.2 % (NE) – 55.8 % (EL).
2. The ratio of WO variants in NE constructions with V_{in}(transitive): English: SV (41.5 %) – VS (10.5 %); Evenki: SV (31.5 %) – VS (2.3 %); Chukchee: SV (18.2 %) – VS (7.3 %); Russian: SV (16.4 %) – VS (15 %).
3. The ratio of WO variants in NE constructions with V_{transitive}: English: SVO (29 %) – SOV (0 %: 1 % goes to OSV of the famous type ‘Talent he had, money he hadn’t’); Evenki: SVO (3 %) – SOV (9 %); Chukchee: SVO (2.7 %) – SOV (2.56 %); Russian: SVO (5.8 %) – SOV (3.6 %).
4. The ratio of WO variants in EL constructions with V_{transitive} (O is omitted): English: SV (3.23 %) – VS (0.12 %); Evenki: SV (6.1 %) – VS (0 %); Chukchee: SV (11 %) – VS (1.9 %); Russian: SV (4.24 %) – VS (1 %).

The most interesting is the frequency of V-only constructions (with both S and O omitted; the percentage is given in the diminishing order): Chukchee (12.2 %) > Russian (10.36 %) > Evenki (9.5 %) > English (1.73 %). It should be noted that Chukchee is the only language of our sample that has subject-object agreement on the finite verb form.

The data presented above clearly shows that quantitative method helps to get precise information pertaining to such language properties as WO characteristics and ellipsis possibilities.

References

Once upon a time. Progress Publishers. Moscow. 1976.

Evenkijsky folklore [Evenki Folklore]. Compiled by M. Voskobochnikov. Leningrad. 1960.

Skazki [Fairy-tales]. Compiled by I. Pan'kin. Tula. 1976.

Tynetegyn. Chavchyven Lymngylte [Chauchu's Fairy-tales]. Compiled by L. Belikov. Magadan. 1961

The grammaticalisation pathways of the GET verb [tie53] in the Shaowu dialect in southern China

Much investigation on the argument structure and realisation of GET, GIVE and TAKE verbs has been carried out over the past decades (c.f. Newman 1996, Gronemeyer 1999, Diedrichsen 2012, Lenz & Rawoens 2012, Nolan 2012, Tragemel & Habicht 2012, inter alia). These analyses give a special focus to these verbs in Indo-European languages. My paper offers a different set of typological data, hence perspectives, from a Sino-Tibetan language called Shaowu of the Sinitic branch, spoken in Fujian province in southern China, which I collected over the past four years in the field. In particular, I look at the polysemy of the GET verb [tie53] 得 in Shaowu, and attempt to explain its multi-functionality in terms of its diachronic development and change in argument structure. I also examine how its different syntactic configurations coerce gradual semantic change (c.f. Zhang Min 2011, Chappell 2012). Finally, I propose a polygrammaticalisation pathway that is unique to Shaowu, and compare it to the GET verbs in Mandarin, Cantonese, English and French.

The morpheme [tie53] starts out as a concrete lexical verb meaning ‘to get’, ‘to obtain’ (as instanced in example 1). This mono-transitive lexical [tie53] has then developed into a di-transitive lexical [tie53], meaning ‘to give’ (as in example 2). This curious antonymic sense occurs within the di-transitive construction, the mechanism of which will be accounted for, and which may explain the interesting fact as to why Shaowu does not have a basic verb of ‘giving’.

(1) As a mono-transitive lexical verb ‘to get’ [S+V_{get}+O]

□他/她	得	到	了	蜀	个	奖。
xu ³⁵	tie ⁵³	tau ²¹	ə ⁰	ei ²¹	kei ²¹	tsioŋ ⁵⁵
3SG	get	reach	PFV	one	CLF	prize

‘S/He got/obtained a prize.’

(2) As a di-transitive lexical verb ‘to give’ [S+V_{give}+IO+DO]

□我	得	□他/她	十	块	票	儿。
xaj ³⁵	tie ⁵³	xu ³⁵	ein ²¹	k ^h uai ²¹	p ^h iau ²¹³	ə ⁰
1SG	give	3SG	ten	CLF	bank note	

‘I gave him/her ten RMB (monetary unit in China).’

In addition, it can also be used as a causative verb ‘to make’ (example 3), or a permissive causative ‘to let’ (example 4), or a modal auxiliary meaning ‘to allow’.

(3) As a ‘make’ causative verb [S+CAUS_{make}+O+VP]

大	家	得	新	姐	夫	向	人	客	敬	酒。
t ^h ai ³⁵	ka ⁰	tie ⁵³	sən ²¹	tsia ⁵⁵	fu ⁰	xiŋ ²¹	nin ²²	k ^h a ⁵³	kin ²¹	tsou ⁵⁵
Everyone		CAUS	bridegroom		ALL		guest		toast	wine

‘Everyone made the bridegroom toast to the guests.’

(4) As a permissive ‘let’ causative verb [S+CAUS_{let permissive}+O+VP]

娘	佬	得	小	团	子	去	外	头	玩。
niŋ ²²	lau ⁰	tie ⁵³	ei ^h iau ⁵⁵	kin ²¹	tsə ⁰	k ^h ɔ ²¹	uai ³⁵	xəu ²¹	kau ⁵⁵
mother	CAUS	small	child	DIM	go	outside		play	

‘The mother let the little child go to play outside.’

The multi-functional [tie53] has progressed along various grammaticalisation pathways to be used as a benefactive marker (example 5), a dative marker (example 6) and a passive marker (example 7), as

well as a verb complement marker. Most, if not all, of these pathways are found in some of the world's languages, documented in Heine & Kuteva (2002).

(5) As a benefactive marker [S+BEN+O+VP]

(□你)	得	□我	去	买	菜。
xiej ³⁵	tie ⁵³	xaj ³⁵	k ^h ɔ ²¹	mie ⁵⁵	t ^h ə ²²
2SG	BEN	1SG	go	buy	food

'(You) go and buy food for me.'

(6) As a dative marker [S+V+DO+DAT+IO]

□我	拿	了	个	本	书	得	□他/她。
xaj ³⁵	na ²²	ə ⁰	kei ²¹	pən ⁵⁵	ɛy ²²	tie ⁵³	xu ³⁵
1SG	take	PFV	one	CLF	book	DAT	him/her

'I gave a book to him/her.'

(7) As a passive marker [Patient+PASS+Agent+VP]

□我	得	□他/她	骂	了。
xaj ³⁵	tie ⁵³	xu ³⁵	ma ²¹³	ə ⁰
1SG	PASS	3SG	scold	PFV

'I was scolded by him/her.'

It is the aim of this paper to explore these various roles of [tie53] by looking at its syntactic configurations and semantic functions, as well as how they interact with each other. A diachronic account will be given to explain the multi-faceted synchronic properties of [tie53], against a backdrop of typological features belonging to what Chappell (2012) classified as transitional area between the North and the South.

Theme session: Linked Data in Typology

Crowdsourcing WALS

WALS (Haspelmath et al. 2005) is a hallmark of typology and has revolutionized the way how linguists do typology (Donohue 2006).

Languages are squared with features. At the intersection, we find feature values. E.g. the language English has the feature 'word order' with the feature value 'SVO'. WALS lists 192 features and 2678 languages. However, the resulting data matrix is very sparse, and instead of the possible 514176 data points, there are only about 68000, or 13%.

Already at 13% filling, the matrix is interesting for typologists, and many researchers make use of WALS data, leading to many papers, some of them very influential, e.g. Atkinson (2011). A less sparse matrix is a clear desideratum for typology.

The WALS team do not have the capacity to fill in all the data points, but there are many projects out there, without institutional links to WALS, which collect data according to features defined by WALS and languages as defined by WALS. These projects either add additional feature values for languages which were lacking a value for a certain feature. Or they add a novel feature and add feature values for languages, using WALS codes for the languages. The problem is how these additional data can be made available to other users of WALS.

The WALS database is closed. There is no interface and no process to ingest new data into WALS. The reason for this is that MPI-EVA will not allow write access to its databases from the outside, even less so for random people. Furthermore, WALS has a very high scientific reputation, which could be tarnished by low quality contributions.

The solution for these problems is that data producers store their data points in their own web space according to Semantic Web principles. The data points are then registered and made available to the WALS project for harvesting, including provenance data. The WALS project then makes available the harvested data on the WALS site under a specific label, e.g. "WALS community", in addition to the WALS core already available. Users can then choose whether they want to query only the curated data by the WALS editors, or whether they are also interested in data from other data providers.

This procedure has the following advantages

- Every project manages its own data points. No security risks for central WALS core server
- Clear indication of provenance
- Data producers do not have to care about web server and database administration. They only have to provide their data points in the specified format, which is reasonably trivial
- Shared implementation allows for easy aggregations
- Clear definition of work flow allows automation of processes

This talk will explain the general setup of the crowdsourcing project and show a proof of concept where "WALS core" and "WALS community" are accessed in one query.

References

- Atkinson, Quentin. 2011. "Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa". *Science* 332, 346-349.
- Donohue, Mark. 2006. *Review: Typology: Haspelmath, Dryer, Gil & Comrie (2005)*. <http://linguistlist.org/pubs/reviews/get-review.cfm?SubID=71168>.
- Haspelmath, Martin, Dryer, Matthew, Gil, David & Comrie, Bernard (eds.). 2005. *World Atlas of Language Structures*. Oxford: Oxford University Press.

Attributive and depictive uses of caritive in Bashkir

The Bashkir language (Altaic, Turkic) has an affix *-hǟǟǟ* (with variants *-heǟ / -hoǟ / -höǟ*) that expresses caritive meaning ‘without’. It is highly productive and combines with any noun and personal pronouns. There are at least two types of uses of caritive: attributive and depictive¹. The attributive use implies that the word with suffix *-hǟǟǟ* is a syntactic dependent of a noun that it characterizes. In Bashkir it means that it takes place in preposition of a noun:

- (1) *Anda min bala-lar-hǟǟǟ duθ-tǟ kür-ǟe-m.*
 there I child-PL-CAR friend-ACC see-PST-1SG
 ‘There, I met a friend who doesn’t have children’

The depictive use implies that the caritive form functions as copredicative and can’t be considered as a syntactic dependent of a noun, although semantically, it characterizes the referent expressed by the noun.

- (2) *Anda min Bolat-tǟ bala-lar-hǟǟǟ kür-ǟe-m.*
 there I Bulat-ACC child-PL-CAR s ee-PST-1SG
 ‘There, I met Bulat without (his) children.’

Depending on the attributive or depictive use, the caritive has different grammatical features. For example, in attributive use the word with the caritive marker can’t (with some exceptions) have dependent adjectives, pronouns, numerals:

- (3) **Bǟl kitap-hǟǟǟ bala mäktäp-kä kil-de.*
 this book-CAR child school-DAT come-PST
 ‘A child without this book came to school.’

At the same time the depictive use allows all these dependents:

- (4) ^{ok} *Bala bǟl kitap-hǟǟǟ mäktäp-kä kil-de.*
 child this book-CAR school-DAT come-PST
 ‘A child came to school without this book.’

Personal pronouns can have the caritive marker (like *heǟ-heǟ* you-CAR ‘without you’) only in depictive contexts.

Besides, the caritive form never allows any expression of possession such as personal-possessive markers (** kitap-tǟm-hǟǟǟ* book-P.1SG-CAR ‘without my book’) or genitive form of a noun (** Bolat-tǟn kitap-hǟǟǟ* Bulat-GEN book-CAR ‘without a book of Bulat’). This phenomenon is supposed to be connected with the nature of caritive form: the marker *-hǟǟǟ* is not a case marker and has some derivational properties, that is why it has some morphological restrictions which case markers do not have.

In the talk the grammatical features of the Bashkir caritive form will be discussed in detail considering attributive and depictive uses. The semantic and pragmatic difference between two types of uses will also be taken into consideration. Different grammatical properties connected with the two uses of the caritive can probably be explained by the fact that the attributive use usually implies non-referential object (cf. (1) where children do not exist) while dependent adjectives, pronouns, numerals are often used with referential nouns, i.e. in depictive use.

Abbreviations

1 – 1ST person; ACC - accusative; CAR - caritive; DAT - dative; GEN - genitive; P - possessive; PL - plural; PST – past; SG - singular.

References

J. van der Auwera, A. Malchukov. A semantic map for depictive adjectivals // E. Schultze-Bernd & N. P. Himmelmann (eds.). *Secondary predication and adverbial modification: the typology of depictives*. Oxford: Oxford University Press. 2005. Pp. 393-421.

Some curious and subtle properties of grammatical objects in Aguaruna (Jivaroan)

Aguaruna is a Jivaroan language spoken mainly in Amazonas, Peru. Morphologically it is suffixing and agglutinating and shows both head and dependent marking. Unmarked constituent order is predicatefinal, and clause-chaining is pervasive. Grammatical relations centre on Subject and Object, and basically follow accusative alignment. Morphological (case-marking and verbal indexing) and syntactic properties of Subject are uncontroversial, but Object is less clear. Two phenomena in particular stand out as typologically interesting: (i) split marking of objects and (ii) status of multiple objects.

Split marking

There is a scenario-conditioned split in accusative case marking (see Witzlack-Makarevich 2011 §8.6 for discussion), whereby third person objects remain unmarked if the subject is first person plural or second person. Overall (2007) relates this to a hierarchy $1sg > 2sg > 1pl/2pl > 3$, but in fact this is not sufficient to explain scenarios with 1pl acting on 2sg/pl or vice versa. This paper will report on new fieldwork data and attempt to clarify the details of this typologically unusual system.

Multiple objects

Case marking of all objects (notional direct and indirect objects as well as those added by applicative derivation) is identical. Syntactic processes are even less selective, with relativisation and nominalisation simply contrasting Subject with “non-Subject”, which may include locations and other oblique participants. This suggests that Aguaruna is a symmetrical language, in the sense of Bresnan & Moshi (1990). There is only one morphological slot for marking SAP objects on the verb, and speakers paraphrase to avoid competition for this slot. These avoidance strategies, together with the marking of what Haspelmath (2007) labels ditransitive inverse situations, suggest a ranking of Beneficiary/Recipient over Theme, which in turn presupposes a syntactic distinction between these roles. However the very marginal role of this distinction in the grammar raises the question of what it means for a language to be asymmetrical (cf. Zariquiey 2011 on Kashibo-Kakataibo).

References

- Bresnan, Joan & Lioba Moshi (1990) ‘Object asymmetries in comparative Bantu syntax.’ *Linguistic Inquiry* 21: 147–85
- Haspelmath, Martin (2007) ‘Ditransitive alignment splits and inverse alignment.’ *Functions of Language* 14(1):79-102 (special issue on ditransitives, guest edited by Anna Siewierska)
- Overall, S. E. (2007) ‘A grammar of Aguaruna’ PhD dissertation, RCLT La Trobe University, Melbourne, Australia
- Witzlack-Makarevich, A. (2011) ‘Typological variation in grammatical relations’ PhD dissertation, Universität Leipzig, Germany
- Zariquiey Biondi, Roberto (2011) ‘A grammar of Kashibo-Kakataibo’ PhD dissertation, RCLT La Trobe University, Melbourne, Australia

Nominalized predicates and the Tibeto-Burmanization of Nepali

Tibeto-Burman languages in the Central Himalaya exhibit a typological profile for nominalized verbs that Bickel (1999) has called “the Standard Sino-Tibetan Nominalization (SSTN) pattern”, where nominalized forms are used for verbal complements, attributive adnominals, relativization, and frequently as main verbs (Watters, 2008; Noonan, 2011 inter alia). The pattern has been much discussed in TB literature, however the existence of very analogous forms has attracted less attention in another major language of the area: Nepali.

The intense contact between Nepali and Tibeto-Burman languages is documented over many centuries (Driem, 2001 inter alia), and the striking parallels of Nepali’s use of nominalized verbs with the “SSTN pattern” indicate that it could well be considered an areal as well as a Tibeto-Burman genetic feature. In conforming to the prevailing linguistic profile of the Central Himalaya, Nepali has essentially abandoned the distinction between “finite” verbal morphology for main clauses and “non-finite” morphology for dependent clauses which generally holds in other New Indo-Aryan languages (Masica, 1993) and throughout Indo-European.

This paper will focus on one of the most aberrant phenomena in Nepali from an Indo-Aryan/Indo-European perspective: use of nominalized predicates as independent verbs to achieve certain pragmatic effects. For example, to say “I came yesterday” a Nepali-speaker can choose between *hijo ā-ẽ* (yesterday come-PFV.1P) or *hijo ā-eko* (yesterday come-PST.PTCPL), the first being a “finite” verb with person agreement and the second a nominalized form. Whereas the first would give the proposition full narrative force as a foregrounded action, the second would indicate an action which is either backgrounded or somehow topical in the discourse context.

Using discourse-based data from the Nepali spoken corpus, I will show that Nepali’s usage of nominalized verbs as a discourse strategy more closely resembles typical Tibeto-Burman rather than typical Indo-Aryan syntactic/pragmatic patterns. This case stands as yet more evidence that areal factors such as substrata and contact can play at least as large a role as genetic affiliation in determining the typological profile of a language, and on occasion can draw it far from the “standard” typology of its family (see Noonan, 2010; Donohue, 2012).

References

- Bickel, Balthasar. 1999. *Nominalization and focus constructions in some Kiranti languages*. In Yogendra P. Yadava & Warren Glover (eds.). *Topics in Nepalese Linguistics*, 271-96. Kathmandu: Royal Nepal Academy.
- Donohue, Mark. 2012. *Typology and areality*. *Language Dynamics and Change* 2: 98–116.
- Driem, George van. 2001. *Languages of the Himalayas* (2 vols). Leiden: Brill.
- Masica, Colin. 1993. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Noonan, Michael. 2010. *Genetic Classification and Language Contact*. In: Hickey, Raymond (ed.) *The Handbook of Language Contact*, 48-65. Singapore: Wiley-Blackwell.
- Noonan, Michael. 2011. *Aspects of the Historical Development of Nominalizers in the Tamangic Languages*. In: Foong Ha Yap, Karen Grunow-Harsta and Janick Wrona (eds). *Nominalization in Asian Languages: Diachronic and typological perspectives*, 195-214. Amsterdam and Philadelphia: John Benjamins.
- Watters, David. 2008. *Nominalization in the East and Central Himalayish Languages of Nepal*. *Linguistics of the Tibeto-Burman Area* 31(2): 1-43.

Theme session: Typological hierarchies in synchrony and diachrony

The interaction of hierarchies of number, animacy and morphosyntax in Meso-Melanesian

Across the Meso-Melanesian (MM) subgroup of Oceanic (Papua New Guinea), a hierarchy of expression of number interacts with hierarchies of animacy and morphosyntactic exponence in ways that are unusual and significantly under-reported in the literature. This paper presents the details of some of these interactions.

Some MM languages give neat evidence for differential number marking on the basis of hierarchies of number and animacy. In Vinitiri (Van Der Mark 2007), dual, trial and plural are expressible by pronoun or inflection. With humans, plural marking is obligatory while dual and trial are optional, but not equally so: dual occurs significantly more frequently than trial. Rather than simply supporting number hierarchies such as Corbett (2000:38) as an implication hierarchy of the exponence of number categories in a language, this therefore demonstrates a hierarchy of preference of use in a language that allows expression of multiple categories, providing even stronger evidence for the psychological reality of the number hierarchy. Vinitiri also displays differential likelihood of expression of optional categories on the basis of a morphosyntactic hierarchy where expression of dual or trial is more likely in pronouns than in agreement, as in (1). Moreover, animacy also plays a role in this system. While the facts above are true for humans, non-singular inflection is impossible with inanimates, singular occurring regardless of number (agreeing with Corbett 2000:55-57). However, lexical expressions of number such as a dedicated plural quantifier can occur, giving apparent number contradictions, as in (2), again demonstrating differential number marking on the basis of both animacy and morphosyntax.

- (1) *Mi mutu βuse bur̄si u-ra=ra pisa*
 1EXC.PL.SBJ chop throw.away fall to-DIR=ART ground
 ‘We [three] chopped [it] away onto the ground,

na-muru mitalu mutu-iau a urū-na-p̄kan̄.
 LOC-follow 1EXC.TR.SBJ chop-1SGOBJ ART two-LIG-piece
 then we three chopped me a piece.’

- (2) *Supu di ḡ k̄li ra=um̄n̄ tunu.*
 PURP 3PL.SBJ PST dig ART=PL tunnel
 ‘They were supposed to dig tunnels.

βare mi ḡ kisi, mi ḡ launu ta-n̄.
 PURP 1EXC.PL.SBJ PST stay 1EXC.PL.SBJ PST live LOC=3SG.PSSR
 So we stayed, we lived in it [the tunnels].’

Animacy also interacts with number in unusual ways in the phenomenon of inverse number marking. Some MM languages divide nouns into animacy-based classes in which the singular article of one class marks plural in the other and vice versa (Baerman 2007:40-41; Corbett 2000:163-165; Palmer 2012:449-451). The role of animacy in inverse number varies between languages, but in several animacy interacts with number variably on a number of bases, a fact not reported in the literature. In Nehan (Logan et al 2013), class boundaries fall at different points on the animacy hierarchy on the basis of grammatical role. While human and body-part terms are always in an “A-Class”, and

inanimates in an “O-Class”, the boundary varies with non-human animates. For core arguments, some animates are A-class and some O-Class. With obliques, however, all animates are A-Class.

In Teop (Mosel & Spriggs 2000), class membership again divides at differing points in the animacy hierarchy, but on a completely different basis. In this language three classes exist, an “E-Class”, an A-Class and an O-Class. Corbett (2000:164-165) treats the E-Class as distinct only in singular, with plural expressed using the same article as the A-Class. However, a distinct E-Class plural article does in fact occur, but with a reduced range of items. In singular, personal names, kin terms, important humans (‘chief’, ‘priest’ etc) and domestic animals are E-Class, while ordinary humans and items further down the hierarchy are A-Class. In plural, however, only personal names and kin terms are E-Class. Important humans and domestic animals are A-Class.

(3)

	<u>E-Class</u> names, kin	important humans, domestic animals	<u>A-Class</u> humans, animates	<u>O-Class</u> inanimates
SG	<i>e</i>		<i>a</i>	<i>o</i>
PL	<i>ere</i>	<i>o</i>		<i>a</i>

The Teop system is also significant in the differential ranking of humans in the hierarchy, with domestic animals placed below important humans, but above ordinary humans, an arrangement not represented in Corbett’s (2000:56-66) typology or similar hierarchies, and at odds with Silverstein’s (1976) basic split between humans and animates (see Allen 1987:57; Comrie 1989:185,194-197; Nichols 1992:160–161; Smith-Stark 1974).

Understanding optional plural marking: a cross-linguistic study on speech production

Plural marking is optional in a large number of languages (Corbett 2000); the research question of this study is which factors determine this optionality. As a starting point, we assume that there are at least two sources of typological difference that interact with number marking: (a) languages differ with respect to the possible conceptualizations of the nominal lexicon, in particular with respect to the possibility to use (particular subsets of) nouns in order to refer to atomic entities (Lucy 1992, Chierchia 1998); (b) languages differ with respect to the locus of encoding plurality of the nominal referents, either within the noun phrase and/or through cross-reference markers on the verb (Corbett 2000).

LANGUAGES: We compare four languages with optional plural marking that illustrate these dimensions of diversity: Urum (Turk, Georgia), Fongbe (Niger-Congo, Benin), Cabécar (Chibchan, Costa Rica), Yucatec Maya (Mayan, Mexico). Obligatory numeral classifiers suggest that the noun denotation cannot be type-shifted to a denotation referring to atomic entities: this is the case for Yucatec Maya and Cabécar (and not for Urum and Fongbe). Plural cross-reference markers on the verb are available in Urum for subjects and in Yucatec Maya for subjects and objects (and not in Urum and Fongbe).

METHOD: After an outline of the crucial facts concerning the properties of nouns and of number marking in these languages (elicited and corpus data), we present the results of a cross-linguistic experimental study (18 native speakers per language; speech production based on a translation task; 24 tokens per speaker; a reliability test with corpus data has been conducted for Urum; Yucatec data is currently collected). This study examines the effects and interaction of two crucial factors: animacy (human vs. non-human) and syntactic function (subject, object, adjunct).

RESULTS: All languages display an animacy effect, such that plural is significantly ($p < .05$; GLMM) more frequent with human nouns than with non-human nouns. The comparison between syntactic functions shows (a) that there is no evidence for a compensation effect of plural marking on the verb and (b) that there is a general pattern subject > object > adjunct (plural frequencies) – however not reaching the significance level in most comparisons. Crucially, we observe a significant difference between languages, such that plural marking in Cabécar appears generally less frequently than in the other languages.

CONCLUSIONS: Cross-linguistically consistent tendencies can be presumably traced back to cognitive asymmetries, which apply to the asymmetries between several classes of entity (human vs. nonhuman) or between different levels of salience of syntactic functions. The fact that the examined languages differ in the overall frequency of plural marking is challenging: the grammatical facts lead to the hypothesis that if the noun (non-specified for number) does not denote atomic entities, plural reference is more likely to be established without plural marking.

References

- Chierchia, G. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339-405.
Corbett, G.G. 2000. *Number*. Cambridge: CUP.
Lucy, J. 1992. *Grammatical Categories and cognition*. Cambridge: CUP.

Metonymy in upper and lower limb nomenclature

A considerable amount of effort has been placed on describing PART TO WHOLE metonymy in body-part terminology, as demonstrated by Old Irish *lām* 'hand' and Modern Irish *lamh* 'hand', 'arm'. While it is generally agreed that languages often use PART TO WHOLE to name terms associated with the limbs, such as 'arm' or 'leg', other metonymies remain under described in the body-part domain. The naming of the arms, hand, feet and legs as units has dominated discussions of metonymic changes in limb nomenclature, resulting in little cross-linguistic observations regarding how languages name other body-parts associated with the limbs, such as the elbows or ankles. Using etymological data, Wilkins (1996) observes that there is also a natural tendency for languages to develop body-part terms from verbal actions associated with them, such as the development from a term meaning 'walk' to mean 'foot' or 'grasp' to mean 'hand'. In the case of semantic change, it is reasonable to hypothesize that languages utilize metonymic processes to name other body-parts associated with the limbs. The question is what types of metonymies languages use to name the elbows, wrists, ankles and knees.

In a genetically- and areally-balanced sample of 153 non-Indo-European languages, it is found that the notion of bending or turning is prevalent in the terms for elbow, wrist, and ankle. Examples such as Emai *uguobo* [nominalizing prefix.bend.hand] 'elbow' show how terms develop from verbal actions associated with them, whereas examples such as Q'eqchi' *kux uq'm* [neck hand] demonstrate an extension of other body-parts that also bend or turn. Using roughly 85 terms meaning 'elbow', 'wrist' and 'ankle' with morphological glosses from my cross-linguistic data, I show that there is a cross-linguistic tendency to name parts of the limbs by their physical functions.

These results not only affirm the claim made by Wilkins (1996), they also support using etymological data as a useful tool to help identify cross-linguistic metonymies. As has been shown with cross-linguistic studies, semantic change is regular, thus it is not surprising that languages use the same types of metonymies to talk about body-parts as for other objects in the world, such as tools, or animals.

References

- Hilpert, Martin. 2007. *Chained metonymies in lexicon and grammar: a cross-linguistic perspective on body part terms*. In G. Radden et al. (eds.), *Aspects of Meaning Construction*. Amsterdam: Benjamins, 77-98.
- Wilkins, David P. 1996. *Semantic change and the search for cognates*. In M. Durie & M. Ross (eds), *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. New York: Oxford University Press. 264-304.

Theme session: Lexical typology of qualitative concepts

Words of hardness and softness: towards lexical typology

This paper presents an on-going study of the semantic domain of softness/hardness. At present we have data on eight languages: Russian, French, Chinese, Korean, and four Uralic languages (Finnish, Hungarian, Komi and Nenets).

The qualities in this semantic domain are patientive, i.e., an object can only be described as hard or soft after interaction with the subject, so for our study we used a typological questionnaire containing a list of situations in which hardness/softness of an object is experienced (touching, chewing, contact with clothes or furniture, etc.)

In our study lexical systems are classified as rich, poor, or average according to the number of words discovered in the semantic domain in question (after [Maisak, Rakhilina eds. 2007]). **Poor systems** contain a single pair of antonyms (Komi *choryd* 'hard' vs *n'ebyd* 'soft'). **Average systems**, containing a total of three words, with synonyms for either 'hard' (Russian *tv'ordyj* and *zhestkij*, 'hard', vs. *m'agkij* 'soft') or 'soft' (French *dur* 'hard' vs. *mou* and *moelleux*, 'soft'), reveal semantic distinctions which fall into two categories:

1) **perceptive** – immanent vs. experiential

The perceptive parameter distinguishes between qualities which are either immanent or dependent on the perception of a subject. For example, in Russian *tverdyj* describes the ability to resist deformation, and may imply contact of the object with instruments or quasi-instruments (hands), while *žestkij* describes the sensation of hardness which is impressed on a subject, e.g. through chewing, using furniture, wearing clothes etc, cf. the notion of experiential adjectives in [Kustova 2004].

2) **attitudinal** – desirable vs. undesirable

The attitudinal parameter distinguishes between qualities which a subject considers desirable or undesirable. For example, French system contains two terms for 'soft': *mou*, meaning softness as pleasant, comfortable or otherwise positive, such as soft (fresh) bread, soft (comfortable) cushions; while *moelleux* tends to have negative connotations, such as soft (overcooked) potatoes etc.

Based on the distinctions found in "average" systems, we can hypothesize the existence of systems with four terms, either perceptive (e.g., 'objective hardness' vs. 'experienced hardness' or 'objective softness' vs. 'experienced softness') and/or attitudinal (e.g., 'desirable hardness' vs. 'undesirable hardness' or 'desirable softness' vs. 'undesirable softness').

Our study also revealed **rich systems** containing more subtle distinctions. Korean, the richest system we have encountered in our study, has three words for 'hard' and seven for 'soft'. It uses both the perceptive and attitudinal oppositions and also introduces additional distinctions: hard shell (shellfish) vs. solid body (stone); visual vs. tactile perception of softness; the result of interaction with an object (e.g., rebounding (pillow) vs. penetrating (jelly)), etc. These distinctions also show cross-linguistic consistency.

The metaphoric meanings of 'soft' and 'hard' are derived in a typologically consistent way, and maintain the core semantic distinctions organizing the perceptive and attitudinal systems, e.g, French *mou* 'soft (undesired)' → *un élève mou* 'a dull, slack student'.

A further interesting feature of the adjectives under consideration is their ability in some languages to function as intensifiers (e.g. English *hard drinker*), which can be regarded as the first step towards

grammaticalization, for more detail see [Rakhilina ed. 2010]. Note also that grammaticalization is a fairly rare discussed phenomenon for adjectives, cf. [Heine, Kuteva 2002].

References

- Heine B., Kuteva T. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press
- Kustova G. 2004. *The Types of Derivative Meanings and Mechanisms of Language Expansion*. Moscow, Jazyki Slavjanskih kul'tur
- Majsak T., Rakhilina E. (eds) 2007. *Verbs of AQUA-motion: lexical typology*. Moscow, Indrik
- Rakhilina E. (ed.) 2010. *Construction Linguistics*. Moscow, Azbukobnik, 2010

Theme session: Linked Data in Linguistic Typology

Noun Modifying Clause Constructions in Dravidian

The Dravidian languages present Noun Modifying Clause Constructions which show a number of typological similarities with the Japanese ones, notably its versatility and the absence of explicit indication of syntactic and semantic relation between the head noun and the dependent clause.

The core device use in all the languages except Brahui (North-Dravidian) is a participle, varying in tense-aspect and polarity and suffixed with an adjectival marker (mainly *-a, *-i). The clause ending in this participle precedes the head noun. Most of the arguments (core, oblique, peripheral...) of the dependent clause can be taken as head:

Telugu: annam tin-e ceyyi 'the hand with which one eats'
 food eat-hab-adjz hand

Badaga: konju ginju ella murida sadda 'the sound of all the branches breaking'
 branch ECHO all break-past-adjz sound

To the noun head can be substituted a pronominal derivative, suffixed to the participial form of the verb. Usually this suffix varies in gender and number, but is restricted to the 3rd person (except in a few languages which allow 'I/you... who...'). A common use of the 3rd person neuter singular of this construction is to nominalize the clause: 'the fact that...'. Both types, with pro-head1 and with head noun2, occur in the following sentence:

Kannada: śiva mandirada munde basavaṅṅa maṅṅapa iruvudu1
 Siva temple-gen before Nandi-gen hall be-nonpast-nomz
 ellariḡu gottu iruva saṅṅati2
 all-dat known be-nonpast-adjz fact
 '[the fact] that a Nandi hall stands before a Siva temple [is] a fact known to all'

The paper will present the regular features of the construction as well as some restrictions and less common features. Among them is the extension of the use of the modifying clause with non-nominal heads, such as the adverbials, munde 'in front of, before', mēle 'on, after':

Kannada: maḷe banda mēle ḡle bandu ide 'after the rain came, the river has risen'
 rain come-past-adjz after river come-cnj is

Diachronic morphological data show some affinities between the verbal modifier (participle) and the nominal complement (genitive) both expressing a syntactic dependency on a head and the lack of semantic specification of the relation.

References

- Krishnamurti, Bh. 2003. *The Dravidian languages*. Cambridge.
 Steever, S. 1998, *The Dravidian languages*. Routledge.

Theme session: Lexical typology of qualitative concepts

Lexical sources for SPEED adjectives and adverbs

Among qualitative concepts, that of SPEED is likely to be one of the most complex, and, at the same time, tending to a universal lexicalization. Interestingly, within the process of lexicalization, it displays both very special patterns of metaphorization (cf. ‘quiet, soft, gentle’ → ‘slow’, as in Russian *tixij*) and develops non-trivial instances of grammaticalization (cf. ‘quick’ → ‘if’, as in Polish *skoro* or Russian *kol’ skoro* ‘as long as’). All these properties are highly relevant for linguistic theory (and lexical typology in particular) and make it particularly pertinent to pay attention to different lexicalization patterns moulding the concept of speed in the world’s languages.

The paper considers main conceptualization patterns for the two domains: those of high and low speed, respectively.

The domain of *high speed* is usually rich and well elaborated. For example, Russian disposes of more than 15 adjectives denoting high speed (as *bystryj, skoryj, šustryj, sporyj, prytkij, provornyj, stremitel’nyj* etc.); however, there is one dominant unit (*bystryj*) which covers all semantic varieties – other units are either more specialized or getting obsolete. Within this domain, a typical polysemy is between ‘velocity’ meaning (roughly, describing a higher intensity of a nonhomogenous process, as in *spread quickly* [at a high speed]) and ‘immediate’ meaning (roughly, describing a reduced interval between two events, as in *answer quickly* ≈ ‘answer immediately [after being asked]’). Less often, a third meaning can be added, namely, ‘early / premature’ (as in Japanese *haya*).

Typically, two main sources for lexical expression of high speed are detected. The first one is provided by prototypical “high-speed” situations of fast physical motion – such as falling, running, flushing, hitting, throwing, etc. Cf. Latin (and Romance) *rapidus* (< *rapere* ‘grasp, grab’), Lithuanian *ristas* (etymologically related to ‘run’, cf. C.-Sl. *ristati* ‘run, leap, ride’), Russian *šibkij* and Polish *szybki* (both etymologically related to ‘throw; hit’), Polish *prędko* (etymologically related to ‘flush, flow’), Czech *rychlý* (attested in all West Slavic, etymologically related to ‘move’ and ‘break down’), etc. The second one is related to habitual situations – properties of prototypically “quick” agents. Cf. Rus. *živoj* ‘alive’, *vesělýj* ‘merry, cheerful’ (esp. as quasi-imperatives of the type *živee!*, *veselee!* ‘<be> quick!’, ‘lively!’) and Chinese *kuài* ‘quick’ < ‘joyful, pleasant’. Both positive and negative connotations of high speed terms are possible: high speed may be perceived as an advantage (witness Russian *živoj*), as well as disadvantage (witness Slavic *nagl-*, combining, in different Slavic languages, meanings ‘quick / sudden’ and ‘insolent’). Negative connotations are well attested, e.g., also in French, cf. colloquial expressions such as *aller plus vite que la musique*, *il y va un peu vite*, etc.

The domain of low speed is usually less elaborated, but displays similar characteristics; here, typically “slow” situations and typical human qualities are at work as well.

For high-speed terms, a frequent grammaticalization is attested. We argue that the two main grammaticalization paths are intensifiers and temporal or conditional connectors developing from ‘velocity’ meaning and ‘immediate’ meaning respectively.

Theme session: General session or Generalized Noun Modifying Clause Constructions

Relative clauses in Mayan: Subject preference under ergativity

This paper presents and analyzes the processing of relative clauses (RCs) in Mayan languages Ch'ol and Q'anjob'al.

Background. A large body of work on the processing of relative clauses maintains that subject relatives are generally easier to process than object relatives [2; 3]. However, until recently, cross-linguistic evidence for the subject processing advantage (SPA) in RCs was limited to nominative-accusative languages. In such languages, it is hard to determine whether the processing advantage is due to subjecthood or surface cues like word order, case marking, or agreement. Ergative languages provide an excellent opportunity to tease apart grammatical function and surface cues, since they associate more than one case with the subject position. Unlike nominative-accusative languages, where the SPA and case cueing work in the same direction, in ergative-absolutive languages, the case cue may pull in the opposite direction of the SPA. However, results for Basque [1] and Avar [5] do not seem to show the SPA at all, which necessitates further work on the processing of RCs in ergative languages.

New results. In order to determine whether ergative languages exhibit the SPA, we studied the processing of RCs in Ch'ol and Q'anjob'al, two head-initial ergative languages that mark ergativity via agreement [4]. In both languages, the ergative is syntactic subject, as evident from several diagnostics. Both languages have RCs unambiguous between subject and object interpretation as well as ambiguous RCs. The results indicate a strong processing advantage for subjects. Participants were significantly slower in interpreting object RCs than subject RCs, regardless of whether the subject was absolutive (intransitive clause) or ergative. When presented with unambiguous stimuli, participants were less accurate on object RCs, and with ambiguous RCs, they strongly preferred the subject interpretation. Thus, the SPA appears to be present in Mayan.

Discussion. Our results provide novel evidence for the psychological reality of the notion 'subject', as well as support for the Keenan-Comrie hierarchy in ergative languages. We compare these new results with the findings on RC processing in Basque and Avar, where the SPA seems to be absent, and suggest the SPA applies in a more limited way than usually assumed; in many instances where the SPA has previously been claimed, it may be confounded by surface cues. We also argue that case morphology and agreement morphology are subject to different processing strategies.

References

- [1] Carreiras, M., J.A. Duñabeitia, M. Vergara, I. de la Cruz-Pavia & I. Laka. 2010. *Subject relative clauses are not universally easier to process: evidence from Basque*. *Cognition* 115, 79–92.
- [2] Keenan, E.L., & B. Comrie. 1977. *Noun phrase accessibility and Universal Grammar*. *Linguistic Inquiry* 8, 63-99.
- [3] Kwon, N.-Y., Y.-Y. Lee, P. Gordon, R. Kluender & M. Polinsky. 2010. *Cognitive and linguistic factors affecting subject/object asymmetry: an eye-tracking study of pre-nominal relative clauses in Korean*. *Language* 86, 546–582.
- [4] Larsen, T.W., & W.M. Norman. 1979. *Correlates of ergativity in Mayan grammar*. In *Ergativity: Towards a theory of grammatical relations*, ed. F. Plank, 347–370. London/New York: Academic Press.
- [5] Polinsky, M., C. Gomez Gallo, P. Graff & E. Kravtchenko. 2012. *Subject preference and ergativity*. *Lingua* 121, 120-140.

Theme session: Quantitative Linguistic Typology: State-of-the-Art and Beyond

Tracking Typology on the Micro-Level: Inheritance and Diffusion of Linguistic Features

Most linguistic phenomena show distinct geographical patterns, which has long been recognized in the literature. One of the main tasks of now-a-days typology is to explore how languages and their structures are distributed and to explain which factors determine those distributions (Bickel, 2007). A large number of competing factors responsible for the observed distributions and dependencies among the factors themselves make this task highly complex. In recent years, large amounts of digitally available data, both linguistic and non-linguistic, have enabled the researches to employ quantitative methods and explore alternative explanations. One of the less explored possibilities to tackle this problem is the employment of simulations, i.e. computational modeling, which enables the researches to test various hypotheses about language variation across time and space. Although well established method in some disciplines, like physics, simulations are still very cautiously being employed in linguistics.

In this talk I will present results of the experiment in which the spread of linguistic features is simulated for different areas of the world. The model comprises several general parameters related to geography, like mountain ranges and rivers, while simulating the spread of linguistic features. The main objective of this experiment is to test the influence of geography on the spread of linguistic features within and between languages. Real-world geographic information is explored using the Geographical Information Systems (GIS) in combination with the various Python modules developed for spatial data analyses. The approach taken here is not to simply visualize the spread of linguistic features on the map, but to condition the spread of linguistic features on various geographical factors, test alternative hypotheses, compare it to the real-world linguistic data and, at the very end, visualize the simulated processes.

The proposed simulations make use of the latest developments in geospatial information studies and large amounts of freely available digital data in order to simulate language change and shed more light on the underlying mechanisms responsible for the observed distributions of languages and their structures.

References

Bickel, B. 2007. Typology in the 21st century: major current developments. *Linguistic Typology* 11, 239–251.

Bi-absolutive construction in Nakh-Dagestanian: Unity and diversity

Introduction. Nakh-Dagestanian (ND) languages are characterized by ergative alignment, i.e., subjects of intransitive clauses pattern with patient arguments of transitive clauses with respect to case marking and agreement, with agent arguments getting ergative case.

Bi-absolutives in Nakh-Dagestanian. Almost all ND languages have a bi-absolutive (BA) construction, in which both arguments of the transitive clause are marked with the absolutive case (ABS), as in (1); in such a construction, the agent corresponds to the ‘higher’ ABS, and the patient to the ‘lower’ ABS. The ergative counterpart can be found in the same context (2), though the difference in case marking is associated with subtle differences in meaning. It has been observed that BA sentences are only found with verbs in non-perfective aspect. In this paper we present diagnostics of syntactic structure of the BA construction in ND, which reveal that despite surface similarities (as discussed in Forker 2012), there are structural subtypes among ND BA constructions requiring different analyses.

Possible Analyses. Syntactic properties of the BA construction have received at least two analyses: (a) pseudo-noun-incorporation (Forker 2010, 2012), and (b) bi-clausal structure, with the lower clause corresponding to a non-finite clause (Kazenin & Testelec 1999, Kazenin 2001). Despite accounting for some facts in some languages, neither analysis can be applied to ND languages across the board (cf. Forker 2012). We propose that the two analyses may be still viable, but only for some ND languages. We also propose a third analytical possibility for BA constructions: (c) restructuring or clause union (Wurmbbrand 2004). Under the restructuring analysis, the BA construction is monoclausal with several functional heads above VP.

We discuss the diagnostics that distinguish between the three analytical possibilities. Under (a) ‘pseudo’-noun incorporation (Massam 2001), the lower ABS cannot be expressed by a pronominal or demonstrative; the order of the patient argument and the verb must be fixed (patient-V); gapping constructions must be impossible (if patient and verb form a constituent, it should be impossible to omit just the verb in one of the clauses, given the structural parallelism requirement on gapping); the word order AgentABS PatientABS IO V must be impossible. To differentiate between (b) and (c), further structural tests are needed. The bi-clausal analysis presupposes that the patient argument and the lexical verb belong to a separate clause, to the exclusion of the agent argument. We can therefore expect that the patient cannot be bound by the agent (unless a language has logophors); separate clausal negation should be possible on the lexical verb; the causative derivation should be impossible, and topic/focus particles should be possible in the patient-verb clause. None of these characteristics are expected under restructuring. To illustrate the differences between three analyses, we present and compare data from three ND languages: Bagwali (analysis (a)), Chechen and Lak (b), and Tsez (c).

- (1) *Rasul qata bullali-sa-r.*
Rasul.1.ABS house.3.ABS 3.build:DUR-PART-3
- (2) *Rasul-lul qata bullali-sa-r.*
Rasul-ERG house.3.ABS 3.build:DUR-PART-3

‘Rasul is building a house.’ (Lak; Kazenin 1999: 101)

Theme session: Lexical typology of qualitative concepts

A NEW APPROACH IN 'OLD' STUDIES

This paper is an extensive typological investigation of 'old' on the basis of about a hundred languages. Unlike the previous projects (study of aquamotion predicates in Maisak, Rakhilina 2007, or predicates of pain in Reznikova, Rakhilina, Bonch-Osmolovskaya 2012), this research is dictionary-based, and the question is whether it is possible to carry out a reliable typological analysis using only dictionary entries (supported by expert comments where needed), if a researcher has a strong hypothesis of what a semantic field in question looks like.

The semantics of 'old' has been studied for Russian (Rakhilina 1999), English (Taylor 1992, Partee 1995) and Mandarin Chinese (Kuzmina 2011). It has been proven that there are four basic interpretations (frames) of 'old' depending on internal aspectual characteristics that different constructions "*old* + N" possess:

1. 'such that came into being long before the moment of speech [and has changed]' (e.g. "old oak")
2. 'created long before the moment of speech' (e.g. "old clothes")
3. 'such that came into being or was created long before the moment of speech and is no longer in use' (e.g. "old flat")
4. 'such that came into being or was created long before the moment of speech and is no longer in use: the epoch related to its creation has already passed away' (e.g. "old coins")

Further analysis has shown that these frames can be differently combined within a certain "basic" lexical item in different languages. These combinations are diverse and non-trivial, but certain regularities can also be arrived at.

Three lexical systems appear to be most widespread: a "dominant" system, a "binary" system and a "rich" system. In dominant systems, all the four semantic frames are included in the meaning of a single lexeme; it is typical for Slavic and Germanic languages, as well as for Mandarin Chinese. There can also be additional semantic properties which influence the lexical choice (consider Chinese, where two words 旧 and 久 have a positive and a negative connotation respectively). Binary systems can be found, among others, in Mari and Tatar and have two lexemes which are complementarily distributed to express a certain semantic opposition, e.g. animate-inanimate. In rich systems the four frames are distributed among four separate lexical items (examples are Quechua and Evenki). Note that different systems can be found even within one and the same language group: thus, among Turkic languages, Tatar and Turkish tend towards a binary system, while Yakut has a rich system (typical for other Siberian languages as well).

References

- Maisak, Timur & Ekaterina Rakhilina (eds.). 2007. *Glagoly dviženija v vode: lexičeskaja tipologija* [Verbs of Aquamotion: a lexical typology]. Moscow: Indrik.
- Kuzmina, Ksenija. 2012. *Denoting time in Chinese* (MA thesis). Moscow, Russian State University for the Humanities.

- Rakhilina, Ekaterina. 1999. Aspectual classification of nouns: a case study of Russian. In: W. Abraham, L. Kulikov (eds.). *Tense-aspect, transitivity and causativity*. Amsterdam: Benjamins, 341-350.
- Reznikova, Tatiana.; Rakhilina, Ekaterina.; Bonch-Osmolovskaya, Anastasia. 2012. Towards a typology of pain predicates. *Linguistics*, 50.3, 421–465.
- Partee, Barbara H. 1995. Lexical semantics and compositionality. In: L.Gleitman and M.Lieberman (eds.), *Invitation to cognitive science, part I: Language*. Cambridge (MA): MIT.
- Taylor, John R. 1992. Old problems: adjectives in Cognitive Grammar. *Cognitive Linguistics* 3.1, 1-35.

In search of extreme Tone inventories across Chinese dialects

Introduction

Chinese is considered to consist of numerous dialects (also called as *Sinitic Languages*, Chappell 2001). Meanwhile, Chinese is a contour tone language typically differentiated from register tone languages in, say, West Africa (Pike 1948). How many tones a Chinese dialect may at most/at least employ? What can we learn from extreme tone inventories across Chinese dialects? Through arduous work, the purpose of this study is to identify the extreme tone inventories in varieties of Chinese, aiming at contributions to tone universal studies.

1. Largest tone inventories across Chinese dialects

Largest inventories are controversial in researchers. According to Zhang (1987) and Bi (1994), Xinxu Cantonese is an 11-tone dialect and Hengxian Pinghuais a 10-tone dialect respectively. But the two dialects were both treated as 12-tone systems later (Tan, 2004; Zee, 2012).

The largest inventory maybe consists of 13 tones so far. Jinxian Gan was reported as a 7-tone dialect by Yan (1986, 1988). Bobai Cantonese was rectified from 9-tone to 10-tone dialect around 80 years ago (Wang 1931). But they are both labeled under 13-tone system by Cao ed. (2008). Rongxian Cantonese turns out to involve 13 tones definitely in this study while only 9 tones were accepted in past. It is clear that largest tone inventories are usually resulted from aspiration tones.

2. Smallest tone inventories across Chinese dialects

2-tone system, it seems self-evident in logic, is granted as the smallest tone inventory. Actually there are some dialects with 2 tones, e.g. Honggu dialect (Luo, 1999) and Wuwei dialect (Caoed.2008). An interesting case is Minqin Chinese which has been reported as 4-, 3-and 2-tone dialect in literature. Our study proposes that Minqin dialect has 2-tone system in the majority while it has 3-tone system as well. It is a rare language with multi tone systems.

Some researchers, however, implausibly claim that a handful of dialects make use of only one tone! Zhang (2003) reported Tianzhu Chinese and Minhe Chinese as 1-tone dialects. Though it sounds absurd, our study turns out Tianzhu Chinese has only one tone (a falling tone with rising onset). Acoustical experiment and statistical analysis were conducted to test the results. We argued that 1-tone languages are essentially different from non-tone languages.

3. Some special tones

Level tones and falling tones are analyzed acoustically in this section, including 4 level tones in Cantonese and Leihua (Lin 1995), 4 falling tones in Fuqing Min (Feng 1993. Zhu 2012 disagreed), 5 falling tones in Gurao dialect, and even 6 falling tones in Chaoyang dialect (Jin & Shi 2010).

4. General Discussion

Some methodologies and principles are discussed from extreme tone inventory studies: formal vs. substantial; inter-tone distance vs. intra-tone focalization; pitch range vs. tone inventory size, etc.

Results indicate: 1) Lots of extreme inventories are controversial. Inventory identification is needed obligatorily under CONSISTENT principles. 2) Both of largest and smallest inventories share characteristic of high variations though differentiating in inter-tone and intra-tone variations. 3) Markedness theory is not always convincing. To some extent contour diversity is negatively related to tone categories. 4) No explicit effects of tone inventory size (i.e. correlation related to segmental inventory size, syllable complexity, population size and climate, etc.) are observed.

5. Conclusion

Study on extreme inventories of Chinese (as an enormous contour tone language), is an important supplement to language universals, and are of much significance to language typology.

Theme session: Typological hierarchies in synchrony and diachrony

Deconstructing the person hierarchy as an explanation of the synchrony and diachrony of Tupi-Guarani indexing systems

Among the typological hierarchies, the hierarchies called 'empathy / animacy / saliency / referential or inherent topicality / nominal / indexability hierarchy' (Kuno and Kaburaki 1977, Comrie 1981, DeLancey 1981, Klaiman 1991, Dixon 1994, Givón 1994, Bickel and Nichols 2007) all posit a person hierarchy, generally $1 > 2 > 3$. A remarkable application of the person hierarchy in descriptive linguistics lies in the explanation of hierarchical and inverse indexing systems (Nichols 1992, Zúñiga 2006). Tupi-Guarani languages are cited as perfect examples of a hierarchical indexing system, where the relative ranking of A and P on the hierarchy determines the selection of the person markers (see for ex. Payne 1997), as are Algonquian languages for the inverse systems including a direction marker. The person hierarchy is systematically used as an explanation of the indexing system in Tupi-Guarani studies, be they comparative (Montserrat and Soares 1983, Jensen 1998), general (Payne 1994) or on individual languages. The person hierarchy is also suggested as an explanation for the supposed development of the hierarchical system out of an ergative system (Jensen 1998). This paper questions the relevance of the person hierarchy as a synchronic and diachronic explanation for the Tupi-Guarani systems.

A recent survey of 28 Tupi-Guarani indexing systems surprisingly shows that only two of these systems (those of Ava-Canoeiro and Kayabi) can be said to follow perfectly the 'model' of a hierarchical indexing system based on a $1 > 2 > 3$ hierarchy as outlined above. Some languages follow different hierarchies or different indexing systems, but the great majority of them show a clear $SAP > 3$ hierarchy along with an opaque marking of local configurations (when a speech act participant is acting on another speech act participant). This opacity is such that the systems can hardly be said to derive from a person hierarchy. This nontransparency of local configurations has been well studied for Australian and Amerindian languages by Heath (1998). It is explained as avoidance of pragmatically sensitive combinations (resembling the common pragmatic restrictions on the use of transparent 2SG pronominals), interpretable as face threatening acts, as Brown and Levinson (1987) put it (Siewierska 2004)). Heath suggests that linguists have 'denying' reactions in front of this opacity, for instance imposing hierarchies with artificial segmentation and labeling of surface morphemes. Our survey confirms this 'meta-analysis'.

The relevance of the person hierarchy is questionable when it explains only part of the facts that it is supposed to cover, and when it is reduced to $SAP > 3$. In the case of Tupi-Guarani languages, the $1 > 2 > 3$ hierarchy is an inefficient tool and it does not explain the synchronic facts in a simple manner. A close examination of Algonquian data led to the same conclusion (Macaulay 2009), and an alternative explanation of the system was offered by Zúñiga (2008).

The variation among the 24 languages of the survey that manifest some hierarchical organization (at least $SAP > 3$) is great: there are five types of $2 _ 1$ encoding, and 8 types of $1 _ 2$ encoding. Considering this variation, it seems very speculative to reconstruct an indexing system based on a clear $1 > 2 > 3$ hierarchy. The genesis of the present-day systems has not yet been satisfactorily explained (Gildea 2002). Investigating the presupposed role of the person hierarchy in the development of Tupi-Guarani systems is the next step, after having deconstructed its role as both a descriptive tool and an explanatory device for synchronic data. I suggest it may only be an epiphenomenon of the distinction between nouns and pronouns, on the basis of Anambé data.

References

- Bickel, Balthasar et Nichols, Johanna. 2007. "Inflectional morphology", in *Language typology and syntactic description*, Vol. III: Grammatical Categories and the Lexicon, Timothy Shopen (ed), Cambridge: CUP, 169-240.
- Brown, Penelope et Levinson, Stephen. 1987. *Politeness*, Cambridge: Cambridge University Press.
- Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*, Oxford: Basil Blackwell.
- DeLancey, Scott. 1981. "An interpretation of split ergativity and related patterns", in *Language*: 57, 626-657.
- Dixon, Robert M. W. 1994. *Ergativity*, Cambridge studies in linguistics ; 70, Cambridge ; New York: Cambridge University Press, xxii, 271 p. p.
- Gildea, Spike. 2002. "Pre-Proto-Tupí-Guaraní Main Clause Person-Marking", in *Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL*, Vol. Tomo I, Ana Suely Cabral and Aryon Rodrigues (eds), Belem: Editoria Universitária U.F.P.A.
- Givón, Talmy. 1994. "The pragmatics of de-transitive voice: functional and typological aspects of inversion", in *Voice and inversion*, Talmy Givón (ed), Amsterdam: John Benjamins, 3-44.
- Heath, Jeffrey. 1998. "Pragmatic skewing in 1-2 pronominal combinations in Native American Languages", in *IJAL*: 64.2, 83-104.
- Jensen, Cheryl. 1998. "Comparative Tupí-Guaraní Morpho-syntax", in *Handbook of Amazonian languages*, Vol. IV, Desmond Derbyshire and Geoffrey Pullum (eds), Berlin: Mouton de Gruyter, 490-603.
- Klaiman, M.H. 1991. *Grammatical Voice*, Cambridge Studies in Linguistics 59, Cambridge: Cambridge University Press.
- Kuno, Susumo et Kaburaki, Etsuko. 1977. "Empathy and syntax", in *Linguistic Inquiry*: 8.4, 627-672.
- Macaulay, Monica. 2009. "On prominence hierarchies: Evidence from Algonquian", in *Linguistic Typology*: 13.3, 357-389.
- Monserat, Ruth et Soares, Marília Faco. 1983. "Hierarquia referencial em línguas Tupi", in *Ensaio de linguística*.9, 164-187.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*, Chicago/London: University of Chicago Press.
- Payne, Doris. 1994. "The Tupi-Guarani inverse", in *Voice: Form and Function*, Barbara Fox and Paul Hopper (eds), Amsterdam/Philadelphia: John Benjamins, 313-340.
- Payne, Thomas. 1997. *Describing morpho-syntax. A guide for field linguists*, Cambridge: Cambridge University Press.
- Siewierska, Anna. 2004. *Person*, Cambridge Textbooks in Linguistics, Cambridge: Cambridge University Press, 327 p.
- Zúñiga, Fernando. 2008. "How many hierarchies, really? Evidence from several Algonquian languages", in *Scales*, Marc Richards and Andrej Malchukov (eds), Leipzig: University of Leipzig, 99-129.
- Zúñiga, Fernando. 2006. *Deixis and Alignment. Inverse systems in indigenous languages of the Americas*, Typological Studies in Language 70, Amsterdam/Philadelphia: John Benjamins, 300 p.

A comparison of word order and the intonational phrasing of noun phrases in two languages

This paper will examine the relationship between word order and the intonational phrasing of noun phrases (NPs) in two Australian languages with vast differences in grammatical structure; one a polysynthetic head-marking language and the other a dependent-marking language with extreme morphological marking of dependents. It might be assumed that these differences in grammatical structure may be reflected in differences in intonational phrasing. This is not the case, however, with intonational phrasing behaving similarly in the two languages, providing evidence for the independence of intonation and grammatical structure in the linguistic system.

Findings of a corpus of spontaneous speech reveal that NP ellipsis is common in both languages, as found in Australian languages generally (Bowe, 1990; Bower, 2008; Evans, 2003), suggesting that the very presence of an NP indicates that its status is marked in the discourse (e.g. Bower, 2008; Mushin, 2005). Furthermore, the word order of clauses is very similar in the two languages. In both languages, the position of the NP in the clause serves distinct discourse functions. Pre-verbal NPs serve to introduce a new topic or new information to the discourse, provide contrast, or add drama or emphasis to the narrative, while post-verbal NPs serve to elaborate, highlight or clarify referents. These findings are in line with studies of other Australian languages (Bowe, 1990; Evans, 2003; Laughren, et al., 2005; Mithun 1987; Simpson & Mushin, 2008).

Some differences between the two languages do arise, however, in terms of the intonational phrasing of the argument and adjunct NPs in certain environments. This paper will examine the similarities and differences of intonational phrasing of NPs in relation to word order and discourse structure.

References

- Bowe, H. J. (1990). *Categories, constituents and constituent order in Pitjantjatjara: an Aboriginal language of Australia*. London; New York: Routledge.
- Bower, C. (2008). Bardi arguments: referentiality, agreement, and omission in Bardi discourse. In I. Mushin & B. Baker (Eds.), *Discourse and grammar in Australian languages*. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Evans, N. (2003). *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune* (Vol. 1 & 2). Canberra: Pacific Linguistics, Australian National University.
- Laughren, M., Pensalfini, R., & Mylne, T. (2005). Accounting for verb-initial order in an Australian language. In A. Carnie, H. Harley & S. A. Dooley (Eds.), *Verb First: on the syntax of verb-initial languages*. Amsterdam: John Benjamins Publishing Company.
- Mithun, M. (1987). Is basic word order universal? In R. S. Tomlin (Ed.), *Coherence and grounding in discourse* (Reprinted in D. Payne (ed.) 1992 *Pragmatics of word order flexibility*, 15-61. Amsterdam: John Benjamins Publishing Company ed., pp. 281-328). Amsterdam: John Benjamins Publishing Company.
- Mushin, I. (2005). Word order pragmatics and narrative functions in Garrwa. *Australian Journal of Linguistics*, 25(2), 253-273.
- Simpson, J., & Mushin, I. (2008). Clause-initial position in four Australian Languages. In I. Mushin & B. Baker (Eds.), *Discourse and grammar in Australian languages*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

The phonologically exceptional continent: a large cross-linguistic survey reveals why Australia is, and is not, typologically unusual

Australian languages are known for their very low level of phonological diversity. Yet how and why just one continent should be so homogeneous is not understood. We report on results emerging from the first large scale study of Australian morphophonemics, and show that the “Australian problem” does not extend to all corners of the phonology.

Background Existing phonological surveys of Australian languages have focused on phoneme inventories, static phonotactics and stress patterns. However, to better understand the Australian problem we require more information, preferably both synchronic and diachronic, and thus a promising domain of investigation is morphophonemic alternations: synchronic phenomena which preserve a strong signal of prior changes.

Data The AusPhon---Alternations database is the first large scale survey of segmental morphophonemic alternations in Australian languages. Alternations are coded in a commensurate manner, irrespective of their description in source materials as ‘allomorphy’ or ‘(morpho)phonological rules’. In order to survey information from a wide band of time depths, we will not distinguish here between productive and nonproductive alternations, but focus instead on the alternations’ content. At time of writing, 70 linguistic varieties and ca. 1600 alternations have been coded for.

Emerging findings NO ‘AUSTRALIAN TYPE’ In Australia, segment inventories, phonotactic constraints and stress patterns show only minor variation across the vast majority of languages and language families. In contrast, there is no comparable, widespread sharing of segmental morphophonological alternations. The following patterns do recur across languages, but the rate of incidence is low. 1. STOP LENITION A pattern of sonority---conditioned stop lenition, identified in earlier research, is not uncommon: stops alternate with glides or zero, with stops appearing after occlusives, and glides appearing after continuants. 2. PLACE ASSIMILATION Assimilation in place of articulation is rare, however not as rare as one might expect once phonotactic factors are taken into account. Namely, since phonotactic constraints typically permit only few sonority (or manner) sequence types, and since geminates are generally not permitted, what would have been place assimilation typically results in complete deletion, as for example in /ɲn/ → /nn/ → /n/. 3. C VS V DELETION IN CLUSTERS An asymmetry arises when one examines the deletion of consonants from underlying consonant clusters versus vowels from vowel clusters: consonant clusters show no strong preference for deleting stem versus affixal material, whereas vowel deletion tends to remove vowels stem---finally, and to preserve them affix---initially. This is typologically interesting, in that it contradicts the assumption in some phonological research, that ‘faithfulness’ to stems universally outranks faithfulness to affixes.

Conclusions/perspective The typological homogeneity of Australian language phonologies does not extend to morphophonology. Nevertheless, our observations suggest new insights into those aspects of phonology which are highly uniform: the lenition of stops to glides is inventory---preserving; and assimilation is rare except when it feeds deletion, which preserves phonotactic patterns. Though these effects are small and infrequent, in the long run they may contribute to the temporal stability of the most widespread phonological patterns.

How to design a dataset which doesn't undermine automated analysis

The twenty-first century brings the science of big data. For typologists, two new theoretical tasks have arisen. We must ascertain which new methods are appropriate for the modeling and analysis of language. And we must ascertain which linguistic data, and which methods of representing it, are mathematically appropriate for those methods. We argue that much of data with which typologists are both familiar and comfortable is liable to violate essential preconditions that computational methods regularly place on their input. We review one computational study that employs a dataset which at first glance looks well constructed and examine where, how, and why the dataset undermines the assumptions of the model; we then examine solutions.

The rapid rise of computational statistics presents challenges to a small discipline like linguistics, where even the most outstanding typologists rarely run a lab employing its own information scientist. Moreover, premature studies which do violence to linguistic data but are published in highly visible outlets have muddied the water unnecessarily. The science of the statistics is sound, but the application of it to linguistic data has become contentious. On the positive side, such studies have generated a healthy debate about which statistical models might be appropriate for the analysis of language.

Less attention has focused on the fact that automated methods, even when appropriate, are brittle. They produce meaningful results only if the input data meets stringent mathematical preconditions. Again, distractions abound as attempts are made to analogize from molecular biology to language, when what is at stake for the use of statistical methods is more general and fundamental. Namely, computational statistical methods place strong constraints on the dependencies which may exist between data points; very often, **independence** is required. In linguistics however, a drive for elegance has led us to cultivate analyses in which individual parts are highly interdependent. Without careful scrutiny, these dependencies will carry over into typological datasets, rendering their analysis by most computational methods invalid (or at best, degraded) from the outset.

We illustrate how these issues play out in a study of 121 languages × 160 typological survey-questions (Reesink et al. 2009). We then propose the following methodological guidelines:

1. **Use micro answers rather than macro.** Answers to many 'macro' questions in linguistics, e.g. "are there prenasalised stops?" are arrived at by weighing up answers to multiple, antecedent 'micro' questions, e.g. "does [NC] appear word initially?"; "does /NC/ contrast with /N+C/?". Confusingly, two languages may answer macro questions identically while having none of their micro answers in common. Micro answers are more valuable data.
2. **Identify dependencies.** Macro answers may share underlying micro answers. E.g. "are there prestopped nasals?" and "are there closed syllables?" are both sensitive to micro questions about intervocalic clusters. This gives rise to dependencies.
3. **Minimize and track dependencies.** Dependencies should be minimized, and where they remain, must be kept track of. Doing so enables one to select multiple *subsets* of the data which are independent. Application of these principles will improve the suitability of linguistic datasets for the advanced statistical methods now dramatically impacting the quantitative sciences.

Theme session: Lexical typology of qualitative concepts

Qualities of size: towards a typology¹

Every physical object has three dimensions. However, languages tend to have more than three qualitative lexemes describing size. Traditionally this phenomenon is accounted for by combining parameters: it is considered that several elementary semantic components (the direction of measuring, horizontal vs. vertical orientation of an object, etc.) in different combinations make up the meaning of a lexeme (see Bierwisch, Lang 1989, Lang 1995, Wierzbicka 2007).

Such approach is effective, but it doesn't explain all the cross-linguistic differences in the behavior of lexemes denoting size and doesn't make it possible to predict the restrictions imposed on their combinability. Therefore we have created a novel questionnaire based on the principle of "topological classification" (see Talmy 2000 and Rakhilina 2000 on the domain of dimensions). A topological type embraces both form and spatial orientation of a physical object, and takes into account their functional characteristics and the position of the observer (cf. vertical and horizontal surfaces, containers, balls, prolonged objects fixed at their bottom, etc.).

Our research is already completed for six languages of different groups and families: Russian, Serbian, English, French, Mandarin Chinese and Khanty² (the Tegi dialect); the research on Japanese, German, Spanish, Latin, Welsh, Estonian and Komi-Zyrian is in progress. Evidence shows that topological classes successfully describe the size domain: they are regularly reproduced as the basis for the core oppositions in this semantic zone, but every language combines them on its own.

Thus, for instance, in Mandarin Chinese the thickness of long objects (like columns, ropes, pipes) is described with lexemes *cū* and *xì* (*cū tiěsī* 'a thick wire', *xì gùnzi* 'a thin stick'), whereas the thickness of flat objects (like books, boards, a snow layer) is referred to by another pair of antonymic lexemes: *hòu* – *báo* (*hòu mùbǎn* 'a thick board' – *báo zhǐ* 'thin paper'). Russian however doesn't draw a borderline between these two topological types (in so far as adjectives of size are concerned) and puts them into the same class with the only pair of adjectives for their description: *tolstyj* and *tonkij* (*tolstaja provoloka* 'a thick wire', *tolstaja doska* 'a thick board') – *tonkaja palka* ('a thin stick'), *tonkaja bumaga* ('thin paper').

On the other hand, in Serbian long objects fall into two classes: hollow objects with the interior functional surface (for example, a passage) and objects without any functional interior surface (for example, a girl's braid). The objects of the first type are described with the lexeme *uzan* (*uzani hodnik* – 'a narrow passage'), while the objects of the second type are referred to as *tanak* (*tanka pletenica* – 'a thin braid'). On the contrary, in French only one lexeme – *mince* – is used to describe both types of objects (*un couloir mince* - 'a narrow passage', *une ficelle mince* - 'a thin rope').

Our material shows that topological types are combined differently not only in different languages, but also in different lexemes within one language. For example, the zones of large and small sizes tend to be asymmetrical: there is usually a simpler classification of topological types in the zone of small size. Thus, Tegi Khanty draws no distinction between vertically oriented rigid objects (trees, walls) and containers (for example, a river or a bowl) usual for European languages (cf. English *low* vs. *shallow*, German *niedrig* vs. *seicht*, Russian *nizkij* vs. *melkij*): in Tegi Khanty both a low tree and a shallow river are described with the adjective *tef*. The domain of small sizes is also less elaborated in French and Latin.

¹ work supported by grant РФФИ №11-06-00385-a

² work supported by grant РФФИ № 13-06-00884

References

- Bierwisch, Lang (eds.), 1989. Dimensional adjectives: grammatical structure and conceptual interpretation. Berlin: Springer.
- Lang, E. 1995. Basic dimension terms: a first look at universal features and typological variation // E. Lang (ed.), *FAS papers in linguistics*. B., 66-100.
- Rakhilina, E.V. 2000. Kognitivnyj analiz predmetnyx imen: semantika i sochetaemost' [Cognitive analysis of names of objects: semantics and combinability]. M.: Russkiye slovari, pp. 118 – 152.
- Talmy L. 2000. How language structures space // Talmy L. *Toward a Cognitive Semantics*. Vol. I. Cambridge.
- Wierzbicka, A 2007, 'Shape and colour in language and thought', in Andrea C. Schalley and Drew Khlentzos (ed.), *Mental States: Volume 2: Language and cognitive structure*, John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 37-60.

Syntactic and semantic agreement in the Gújjolaay Eegimaa noun class system

Typological research on agreement systems has revealed that syntactic agreement is more canonical than semantic agreement (Corbett 1991, 2006).

The distinction between syntactic and semantic agreement types is usually illustrated in English by the contrast in; *the committee has/have decided*, where *has* agrees with the formal property of the controller, whereas *have* agrees with its semantic properties. Eegimaa, a Jóola language of the Atlantic family of the Niger-Congo phylum spoken in Southern Senegal also exhibits such a phenomenon as illustrated by the hybrid noun *bá-jur* ‘young woman’, which triggers syntactic agreement on the NP, but human class semantic agreement on the predicate.

- (1) bá-jur babu a-kkoŋ-ut
 CL5-young.woman CL5:DEF CL1-cry-NEG
 ‘The young woman did not cry’

Corbett (1991, 2006) demonstrates that the opposition between syntactic (canonical) and semantic (non-canonical) agreement follows rules of an agreement hierarchy (attributive > predicate > relative pronoun > personal pronoun). The hierarchy predicts that levels on the left are more likely to show syntactic agreement whereas those on the rightmost side will show semantic agreement. The goal of this paper is to discuss agreement mismatches which occur in Eegimaa lexical hybrids and location nouns, and their relevance to the agreement hierarchy predictions.

In addition to the most familiar hybrid nouns, Eegimaa also has typologically fascinating type of semantic agreements, which occur with location nouns, and which are scarcely discussed in the literature. These agreements, which can use agreement markers of up to four different noun classes, may be termed *constructional mismatches* following Corbett (2006).

Similar to Eegimaa hybrids, constructional mismatches also trigger different agreement markers at the same time. However, two types of semantic agreement may be distinguished here in addition to the syntactic agreement illustrated in (2) below. One with the human plural agreement (Class 2) attested on predicates as illustrated in (3), and another one expressing general location with class 5, precise location with class 13 and location inside with class 14 as in (4) below, and which is only found on relative pronouns and personal pronouns.

- (2) é-suh yayu y-o na-juh me e-tos-ut
 CL3-village CL3:DEF CL3-PRO 3SG-see SUBORD CL3-move-NEG
 ‘The village that he saw has not moved’

- (3) é-suh yayu bug-o na-juh me gu-tos-ut
 CL3-village CL3:DEF CL2-PRO 3SG-see SUBORD CL2-move-NEG
 ‘The people that he saw have not moved’

- (4) é-suh yayu b-o/t-o/d-ó nú-pull-o me e-tos-ut
 CL3-village CL3:DEF CL5-PRO/CL13-PRO/CL14-PRO 2SG-come.out-DIR SUBORD CL3-move-NEG
 ‘The village where he came out from has not moved’

This paper augments the typology of agreement by examining rare data from an African language and by providing a detailed analysis of its three-way agreement mismatches with locatives. I discuss the different levels of the agreement hierarchy where semantic agreement types occur, and show that both hybrids and constructional mismatches conform to the predictions made by the agreement hierarchy. I also provide a

detailed examination of the different constructions where agreement mismatches such as those in (3) and (4) are attested, and discuss the semantic contrasts indicated by these mismatches.

Typology and planning scope in sentence production: eye tracking evidence from Tzeltal and Tagalog

The explanatory potential of language processing for linguistic typology has been advocated by functionalist typologists and psycholinguists alike [1,2,3,4]. Building precise, testable theories of the relationship between real-time language use and typological patterns is, however, hampered by a lack of empirical research on language processing in the vast majority of the world's languages [5]. As a small step towards expanding the typological scope of processing research, we investigate the processes underlying simple sentence production in two verb-initial languages, Tzeltal (Mayan) and Tagalog (Austronesian), using a picture-description/eye tracking task. Our goal was to determine to what extent the time-course of sentence production is affected by differences in basic word order and verbal morphology.

In Tagalog, the verb agrees in semantic role with the “privileged syntactic argument” (PSA), which may be either the agent or patient. In Tzeltal, finite verbs obligatorily cross-reference core verbal arguments by means of person markers which vary depending on the verb's transitivity.

We used a picture description/eye tracking paradigm [6] to investigate sentence planning in the two languages. In this paradigm, participants describe drawings of simple events while their speech and gaze are recorded. It has been demonstrated that eye gaze and speech are tightly correlated in such tasks, and therefore can offer a window into the time-course of sentence planning (for English, cf. [7]). Previous studies suggest that in English, sentence planning may be lexically incremental [8], i.e., speakers may start to speak having encoded only one argument of the to-be-uttered sentence, delaying the encoding of further arguments and the relations among them until after speech onset. This kind of incremental planning is supported by English syntax: verbs are in sentence-medial position and at least for full NPs there are no dependencies overtly marked on the initial subject argument. In verb-initial structures, by contrast, speakers must plan the verb before speech onset. By hypothesis, the morphological information encoded on the verb in verb-initial languages should influence the scope of sentence planning units, because any dependencies morphologically marked on the verb must be planned prior to sentence onset.

34 native Tzeltal speakers and 54 native Tagalog speakers participated in a picture description task in which they had to describe target pictures of simple transitive events that were embedded in a list of intransitive filler pictures; the participants' speech and gaze were recorded (120Hz sampling rate).

In Tagalog a character was more likely to be fixated in an early time window of up to 600ms after stimulus onset if it was selected to be the PSA than if it was the non-PSA. We interpret this as early encoding of the PSA together with the sentence-initial verb. In Tzeltal, by contrast, both characters in the picture were looked at equally often for an extended duration (up to 2000ms after picture presentation). This gaze pattern indicates that both arguments, the subject and the object, are encoded early in the production process together with the verb. We attribute the difference in fixation patterns between Tzeltal and Tagalog to differences in agreement categories marked at the verb. Whereas in Tagalog only the semantic role of the PSA is marked on the verb, Tzeltal verbs mark both the subject and the object, meaning that information about both arguments is required for verb planning. Despite these differences, the fixation patterns of both languages indicate that extensive planning is necessary when preparing to utter verb-initial constructions: unlike English subject-initial sentences, more than just the first-mentioned syntactic element has to be prepared before speech onset.

Our results indicate that grammatical structure influences planning scope during sentence production. Uncovering such structurally driven differences in production processes may ultimately lead to new avenues in the study of the relationship between usage and grammar.

References

- [1] Hawkins (2004). *Efficiency and Complexity in Grammar*.
- [2] Christiansen & Chater (2008). *BBS*, 31, 489–509.
- [3] Jaeger & Tily (2010). *WIREs Cogn Sci*, 2, 323–335.
- [4] Bornkessel-Schlesewsky et al. (2008). Bridging the gap between processing preferences and typological distributions: Initial evidence from the online comprehension of control constructions in Hindi. In Richards & Malchukov, ed., *Scales*.
- [5] Evans & Levinson (2009). *BBS*, 32, 429–492
- [6] Griffin & Bock (2000). *Psych. Science*, 11, 274–279.
- [7] Griffin, (2004). Why Look? Reasons for Eye Movements Related to Language Production. In Henderson & Ferreira, ed., *The Interface of Language, Vision, and Action*.
- [8] Gleitman et al. (2007). *JML*, 57, 544–596.

Theme session: Quantitative linguistic typology

Bivalent verb classes: a quantitative-typological assessment

The aim of the study (which is part of an on-going collective project) is to reveal patterns of assigning bivalent verbs to valency classes and to assess the degree of cross-linguistic stability / variation of these classes with the help of quantitative methods. Based on a questionnaire that includes 130 predicative meanings (each is given in a particular context in order to avoid polysemy effects), the data have been (so far) gathered and statistically analyzed for 31 languages. In my talk, I am going to focus on those interim conclusions and hypotheses that can be achieved through comparing verb meanings to each other (the findings based on comparing languages is discussed elsewhere).

In accordance with previous assumptions, in most languages studied there is a clearly identifiable class of transitive verbs that stands apart from all other bivalent classes. The ratio of languages that employ transitive structures can be thus found for each meaning. The resultant distribution of transitivity-proneness is U-shaped, which means that verb meanings tend to cross-linguistically either favour transitivity or favour intransitivity. The data obtained allow us to refine some previous implicational hierarchies related to transitivity (cf. e.g. Tsunoda's work) and to propose new ones. Generally, implicational hierarchies **within** semantic domains (such as e.g. possession or directed motion) appear to be much more robust than cross-domain implications.

The central issue in the study is the way "less transitive" meanings are assigned to individual classes. Not surprisingly, apart from highly transitive verb meanings (in the Hopper and Thompson's sense), there are no other large semantic zones so that bivalent verb meanings belonging to these zones would cross-linguistically tend to have uniform argument realization.

There is a long-standing debate on whether minor valency patterns in individual languages are chiefly motivated by semantic (thematic) role structure or are largely idiosyncratic. In our project the dilemma is explored on quantitative-typological grounds; instead of abstract semantic schemata, we rely upon data from a sample of languages in order to study the degree of motivatedness of individual valency classes. In order to assess the degree of coding similarity between verb meanings we use Hamming-type measures (they can be then plotted in a Neighbor Net dendrogram).

There are semantic groups of meanings that were shown to indeed tend to fall into the same valency class in individual languages, e.g. "reciprocal / comitative verbs" ('agree', 'meet', 'fight' etc.); possession-related verbs; 'ablative verbs' (those meanings that can be construed as "motion from a source"). Even in these groups possible patterns of metaphor can be more important than actual similarity in terms of semantic roles (e.g. low degree of similarity between 'lose' and 'win').

In some areas the resultant groupings are very different from any classification that can be based on what is usually viewed as semantic roles. For example, 'forget', 'envy', 'look at', 'be surprised', 'enjoy', 'like' and 'be afraid' behave very dissimilarly cross-linguistically (although all of these verb meanings can be treated as encompassing an Experiencer and a Stimulus). Thus, either semantic roles generally are not good predictors for valency or, rather, the roles that are typologically relevant for argument coding are different from traditionally understood roles.

A special mathematical tool is proposed that can measure predictability of individual verb's valency class membership in a given language based on corresponding verbs' behaviour in other languages. The (not quite transitive) verb meaning are shown to form a hierarchy from those that tend to employ predictable coding devices (e.g. 'avoid') to those that tend to be more idiosyncratic (e.g. 'win').

Nominalization beyond the nominal domain – the case of Yakkha (Kiranti)

This paper describes the various functions of a nominalization pattern found in Yakkha, addressing both its role in syntax and in information structure. Tibeto-Burman languages are well-known for employing nominalizations beyond referential use; not just in relative clauses, complement clauses or auxiliary constructions; also independent finite clauses occur in nominalized form, typically with some effect on the information structure (Bickel 1999, Watters 1998, Yap et al. 2010). These structures may further develop into regular verbal inflection (DeLancey 2011), which is precisely what happens in Yakkha.

The markers discussed here are =na (singular) and =ha (nonsingular); etymologically related to a set of demonstratives. They may derive noun phrases, but also subordinate clauses with certain nominal features, such as relative clauses and finite complement clauses. Furthermore, and in line with findings from other Tibeto-Burman languages, nominalization in Yakkha extends its function into discourse, namely when it is found on finite main clauses. It may mark both backgrounded information and contested information, and thus, it is hard to find a neat label that subsumes its discourse functions. While others have described the nominalized clause as the marked structure, whether this is also the case in Yakkha cannot be answered straightforwardly. The frequency of nominalized main clauses depends highly on the respective text genre. In isolation, e.g. in elicitations, the inflected verbs always occur in the nominalized form. In narratives, nominalized clauses are rather rare; they are used to set the stage for further events, and they also mark the completion of a narrative episode. By omitting the nominalizer, the speaker conveys that something else is to come, i.e. that some fact is not yet fully established. In conversations, on the other hand, nominalizers are rather the norm, as conversations are generally about exchanging or negotiating facts and their truth value. Nominalizers are frequently found in questions and in the respective answers, and also in constructions expressing deontic modality. Notably, they are absent in nonassertive sentence types, such as imperatives and subjunctives.

As they are so frequent, and as they also show agreement with the verbal arguments, the nominalizers are developing into an integral part of the person inflection in Yakkha. They even have their own alignment pattern, a mixture of ergative and reference-based alignment. DeLancey has proposed that nominalization has played a major role for syntactic change in Tibeto-Burman (DeLancey 2011). Albeit showing many parallels and similarities, these constructions and their functions do not behave uniformly across the languages of the family. Yakkha provides another interesting piece to the puzzle of Tibeto-Burman nominalizations and their developments.

References

- Bickel, Balthasar. 1999. *Nominalization and focus constructions in some Kiranti languages*. In Yadava and Glover (eds.), *Topics in Nepalese linguistics*, 271–296. Kathmandu: Royal Nepal Academy.
- DeLancey, Scott 2011. *Finite structures from clausal nominalization in Tibeto-Burman*. In Yap et al. (eds.), *Nominalization in Asian Languages*, 343-359. Amsterdam: John Benjamins.
- Watters, David. 2002. *A Grammar of Kham*. Cambridge: CUP.
- Yap, Foong Ha and Karen Grunow-Hårsta. 2010. *Non-Referential Uses of Nominalization Constructions: Asian Perspectives*. In *Language and Linguistics Compass* 4 (12), 1154–1175.

Pre- and postverbal complement clauses: A simple ordering alternative?

The present paper examines the cross-linguistic properties of (object) complement clauses from the perspective of constituent-order typology. In contrast to relative clauses (e.g. Lehmann 1984, Dryer 2011), the position of complement clauses vis-à-vis their associated head element (here: the matrix verb) has not yet been considered sufficiently in the typological literature. Previous studies are based on limited or non-controlled samples (e.g. Dryer 1980, Ogihara 2009); they stop short of providing a comprehensive description of the positional patterns; and importantly, they do not investigate the extent to which clause order interacts with other morphosyntactic and functional properties of the complement clause. Conversely, studies that *are* concerned with the structure and semantics of complementation (e.g. Cristofaro 2003, among many others) do not typically make reference to the order of complement and matrix clause, suggesting that this parameter is extraneous to the coding of complement relations.

In this paper, by contrast, we argue that the position of the complement clause significantly affects its morphosyntactic structure, and that it can do so independently of semantic considerations. To begin with, a balanced sample of 102 languages firmly establishes an overall asymmetry in the position of complement clauses to the right of the matrix verb: Over 50% of the languages examined have *exclusively* postverbal complements, regardless of the number of complementation patterns distinguished. Exclusively preverbal-complement languages are found much less often (28%), as are languages with mixed or flexible ordering patterns (20%). As with recessive ordering patterns in general (cf. Greenberg 1966), the occurrence of preverbal complement clauses is typically contingent on harmonic alignment with other prevalent orders, especially with preverbal object NPs. Crucially, however, across both OV and VO languages, it also correlates with structural choices: In particular, preverbal complement clauses are conspicuously more prone to the reduction of clausal morphosyntactic properties (verb form, TAM choices, argument realisation) and are marked as subordinate units in very different ways from postverbal complements. All of these properties mirror the findings for relative clauses, and thus appear to be indicative of a more general pattern of left-branching constituents.

In complementation, however, it is also well-known that the structure of the complement clause interacts with the semantics of the matrix verb (Givón 1980, 2001): Certain predicate classes, notably phasal and desiderative verbs, are much less likely to be accompanied by full clauses than others, such as utterance or propositional-attitude verbs. This, too, is reflected in our data, and the combination of the above findings could thus plausibly suggest that the semantics of the matrix verb also interacts with positional choices: one would expect, for instance, that utterance complements are much more often postverbal than complements of other predicate classes. However, we show that this expectation is not borne out: There are no clear-cut implicational associations between semantics and position that would match the corresponding structural ones, and there are no significant differences in position between the predicate classes that are structurally most distinct from one another. By contrast, all predicate classes in our database show an increase in morphosyntactic reduction when they are coded by preverbal complements. This suggests that pre- and postverbal complement clauses are probabilistically associated with different grammatical structures, independently (to a large degree) of semantic considerations. We conclude that this intimate relation between linearization and syntactic structure needs to be recognised properly in the typological study of clause order, as it has important consequences for the discussion of constituent-order universals in general.

References

- Cristofaro, Sonia (2003). *Subordination*. Oxford: Oxford University Press.
- Dryer, Matthew S. (2011). Order of relative clause and noun. In: *The World Atlas of Language Structures Online*. 2nd ed. Eds. by Matthew Dryer and Martin Haspelmath. Munich: Max Planck Digital Library, chapters 90A-G.
- Dryer, Matthew S. (1980). The positional tendencies of sentential noun phrases in universal grammar. *The Canadian Journal of Linguistics* 25: 123–196.
- Givón, Talmy (1980). The binding hierarch and the typology of complements. *Studies in Language* 4.3: 333-377.
- Givón, Talmy (2001). *Syntax: An Introduction*. Amsterdam, Philadelphia: John Benjamins.
- Greenberg, Joseph H. (1966). Some universals of grammar, with particular reference to the order of meaningful elements. In: *Universals of Language*, 2nd ed., ed. Joseph H. Greenberg. Cambridge, Mass.: MIT Press. 73–113.
- Lehmann, Christian (1984). *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Narr.
- Ogihara, Saeko (2009). Verb-final typology. Unpublished PhD dissertation, University of Colorado.

A typological investigation of iconicity and preferred form in reduplication

This paper explores issues of REDuplication research which are basically pending since the programmatic typological sketch by Moravcsik (1978). RED has attracted the attention of linguistics for some time now because of a) being a morphological process without a phonologically constant exponent but rather deriving its reduplicant shape directly from the respective base form, and b) commonly expressing meanings of a relatively limited set across different languages, pertaining to various notions of plurality, intensity and, interestingly, diminution. An integrated account of phonological and semantic properties of RED focusing on typology is still largely fragmentary, though, not least because work on RED semantics has lagged behind phonologically oriented studies ever since the seminal dissertation by Wilbur (1973).

Deepening (by incorporating more detailed formal and functional analyses of RED patterns) and widening (by including partial RED as well) recent typological approaches to RED offered by Rubino (2005) and Stolz et al. (2011), the present paper investigates the intra- and cross-linguistic make-up of RED systems in the vein of a mono-constructural, non-holistic typology (Himmelmann 2000), based on a modified version of the 100 languages core sample underlying the maps in the WALS (Haspelmath et al. 2005). It is hypothesized that there is much less arbitrariness concerning RED form and meaning in language(s) than has normally been acknowledged and that this can be ascribed to the specific ways in which the general principles of structural preference (Vennemann 1988) and iconicity take effect in RED.

It is shown that formally the structure of reduplicants obeys the synchronic (and diachronic) maxim (Vennemann 1988: 2–3), languages employing partial RED always displaying CV reduplicants and often only these, frequently at the cost of exact base copying (additionally, languages in which the process is less or no more productive often exhibit relics of CV patterns only, preferentially with fixed segments). Moreover, the long-standing claim of partial RED strictly depending on full RED at least has to be reconsidered carefully in light of the data at hand. Functionally, a revised view of iconicity is able to capture the prevalence (perhaps exclusiveness) of certain meanings in RED, the essentially iconic but derived nature of diminution being reflected in the fact that a diminutive use seems to imply a pluralizing function of RED in a language. This argument elaborates on an idea developed by Kouwenberg and LaCharité (2005) and is supported by independent evidence (e.g. approximative plurals, diminutive semantics universals and echo-words). What emerges is thus a more systematic picture of the RED phenomenon which is characterized by typological implications and generalizations.

References

- Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
- Himmelmann, Nikolaus P. 2000. Towards a typology of typologies. *Sprachtypologie und Universalienforschung* 53(1). 5–12.
- Kouwenberg, Silvia & Darlene LaCharité. 2005. Less is more: Evidence from diminutive reduplication in Caribbean creole languages. In Bernhard Hurch (ed.), *Studies on reduplication*, 533–545. Berlin & New York: Mouton de Gruyter.
- Moravcsik, Edith A. 1978. Reduplicative constructions. In Joseph H. Greenberg (ed.), *Universals of human language*, vol. 3, 297–334. Stanford: Stanford University Press.
- Rubino, Carl. 2005. Reduplication. In Haspelmath et al. (eds.) 2005, 114–117.
- Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2011. *Total reduplication: The areal linguistics of a potential universal*. Berlin: Akademie Verlag.
- Vennemann, Theo. 1988. Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin. Berlin et al.: Mouton de Gruyter.
- Wilbur, Ronnie Bring. 1973. *The phonology of reduplication*. Bloomington: Indiana University Linguistics Club Publications.

If it would only rain! Predicate-centered focus, epistemicity and assertion in Secoya

It has long been noted that predicate-centered focus does not operate as a discrete focus subsystem within the grammar of a certain language. PCF rather tends to correlate with certain aspectual/temporal and polar grammatical categories, a fact that has been particularly discussed with respect to African languages (for two well-known accounts see Hyman and Watters 1984 regarding the effects of the intrinsic focus properties of negation, and Güldemann 2003 regarding the development of PCF into progressive aspect). The relevance that other grammatical categories closely linked to the predicate can have for PCF, notably modality, has received less attention so far (but see Lohnstein 2012 regarding verum focus in German).

In this talk I present an overview about the relation between the epistemic(-evidential) verb inflection in Secoya on the one hand and PCF notions and encoding on the other. Information-structural indications in Secoya rely on a large array of encoding means: constituent order, nominal and verbal inflectional options (case markers, etc.), semantically specific emphatic lexemes and prosodic support. The occurrence of particle *ti* carrying high pitch accent in (1), for instance, is particularly marked and serves the emphasis of the denotational semantics of the following constituent with regard to its degree or intensity. Here it targets the main verb (stem *ju'i* 'die') and underlines its sudden effectiveness.

- (1) saiona, i_s e'kë ta_n i ti ju_këña
 sai -o -na i -ë -se'e -kë ta_i -ni ti ju'i -kë -ña
 go -F -DS PRN -M -only -M fall -SS EMPH die -PFV.2/3SG.M -SH
 she went on and he fell down and died (on the spot)

Secoya is an Amazonian language (West Tucanoan) in which the speaker's epistemic stance (including some evidential considerations) is obligatorily expressed at the predicate by inflectional TAM suffixes that also cross-reference the subject. The modal inflectional choice contributes different degrees of illocutionary force. States of affairs regarded as certain (2) are encoded by inflectional suffixes that distinguish from those for rather speculative states of affairs (3). In the latter inflectional paradigm, speech act participants are completely ignored as subject cross-referencing controllers, so that only nominal suffixes get employed. Note that the nominal morphology does not compromise the predicative function of the target but just blocks its contribution of assertive force.

- (2) Ja'kë ti_t api.
 ja' -kë ti_t a -pi
 father -M reach_and_remain -PFV.3SG.M
 Father has come.
- (3) Ja'kë ti_t aë
 ja' -kë ti_t a -ë
 father -M reach_and_remain -PFV.2/3SG.M
 [- question intonation] Father might have come.
 [+ question intonation] Has Father come?

Secoya predicates with nominal inflection can also be extended for assertive reasons: the additional epistemic Probability marker *-ni* contributes moderate assertive force (as in biased questions, wishes,

beliefs), while the additional evidential Secondhand marker *-ña* (see 1) establishes full assertives for hearsay/reported information.

I will discuss the effects that these modal and assertive configurations have on PCF and particularities compared to non-PCF focalization. The study is based on a documentation corpus on Ecuadorian Secoya and some complementary experimental and elicitation data.

The semantic domain of dressing and undressing cross-linguistically

Although putting on and taking off clothes are among the most basic human activities, their linguistic expression has not been systematically studied so far. This paper approaches differences and regularities in how this semantic domain is crosslinguistically treated. Data were collected using a set of 32 specifically designed video clips as stimuli, involving a variety of pieces of clothing and accessories, such as socks, glasses, etc. Spontaneous descriptions of these events were recorded for a sample of 25 languages from a total of 75 speakers. Data were coded for identity vs. non-identity of the main lexical verbs of the response across stimuli (e.g. whether for putting on trousers and scarfs the same or a different verb was used) and the matrices thus obtained were analyzed using Multidimensional scaling (Figure 1).

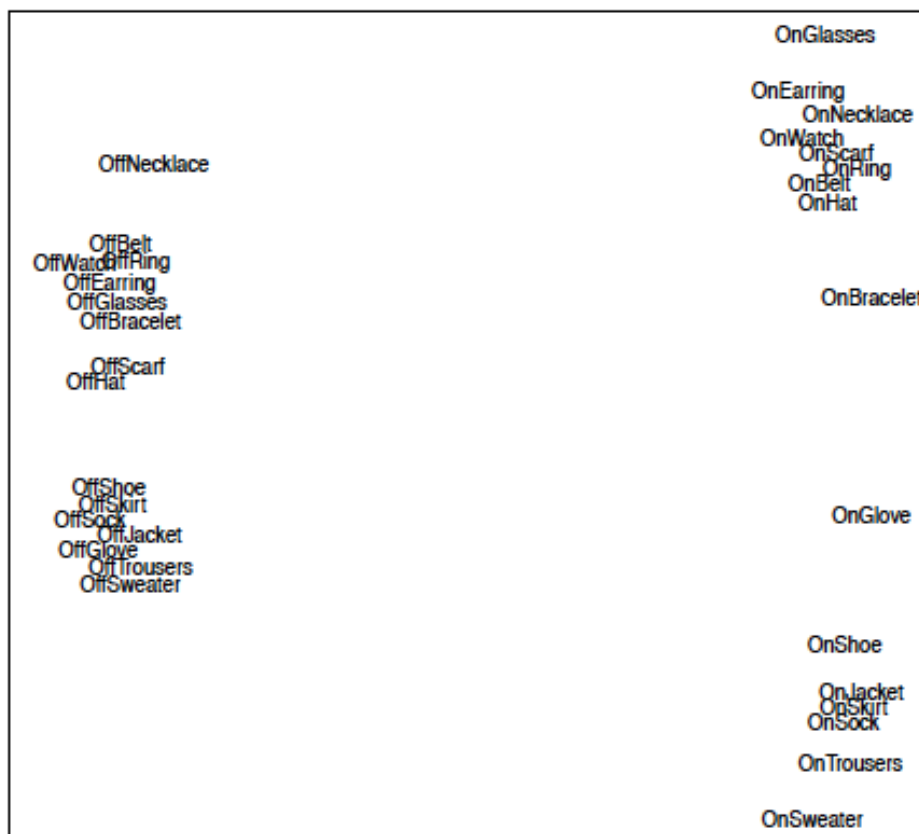


Figure 1: Multidimensional scaling plot of verbs for taking off and putting on clothes and accessories across 25 languages. The distance between the labels for events represents the number of languages in which the event are described by the same verb.

Here we report on two main results from this study. First, there are great differences in the number of verbs employed for events of putting on and taking off clothes. For instance, Haitian Creole speakers use only two verbs (one for putting on, one for taking off), while some speakers of German use up to 17 different verbs. These differences are related to the distinction between verb vs. satellite frame

¹ Sara Mitschke, Franziska von Bibra, Anne Wienholz, Mirjam Grauli, Tina Gregor, Suong-U Kim, Hanne Köhn, André Müller, Larissa Kröhnert, Alicja Stachura, Franziska Roß, Katarina Berndt, Luisa Oswald, Andreas Domberg, Stephanie Klapper, Bettina Klimek, Katrin Maiterth, Lisa Morgenroth, Susann Schildhauer, Hermann Sonntag, Christin Wienold, Saskia Wunder, Ingmar Brilmayer, Johannes Englisch, Hanna Thiele, Frank Wagner, Hans-Jörg Bibiko

languages: As suggested by Slobin (2003), satellite frame languages have rich manner-differentiated verbal lexica, while verb frame languages have fewer verbs. Second, our data clearly show that the subdomain of putting on clothes is cross-linguistically far more differentiated than the subdomain of taking off clothes, i.e. languages tend to have more verbs for putting on than for taking off clothes. (Figure 1: To the left events of taking off, to the right putting on). This can be interpreted as a consequence of the source-goal asymmetry, i.e. “a fundamental cognitive basis in spatial representation, with preferential attention paid to endpoints of motion rather than sources” (Narasimhan et al. 2012:3), which has been postulated based on more limited data of related semantic domains. Additionally, languages clearly differentiate between clothes in a narrow sense (towards the bottom in Figure 1) and accessories such as jewelry and glasses (towards the top in Figure 1).

References

- Majid, A., M. Bowerman, M. van Staden, and J. S. Boster. 2007. The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18.133–152.
- Narasimhan, B., A. Kopecka, M. Bowerman, M. Gullberg, and A. Majid. 2012. Putting and taking events. A crosslinguistic perspective. *Events of putting and taking. A crosslinguistic perspective*, ed. by A. Kopecka and B. Narasimhan, 1–18. Amsterdam, Philadelphia: John Benjamins.
- Slobin, D. I. 2003. Language and thought online: cognitive consequences of linguistic relativity. *Language in mind: advances in the study of language and cognition*, ed. by D. Gentner and S. Goldin-Meadow, 157–192. Cambridge, Mass.: MIT Press.

Grammatical properties which influence GNMCCs

This paper aims to extend or further circumscribe the typology of GNMCCs by surveying data from a wider range of languages than has been considered to date. Based on observations and hypotheses in the literature, we investigate three properties which may positively correlate with the presence of the GNMCC, and which are canonically attested in Japanese (Matsumoto 1997):

- (1) a. The noun modified by the clause is not syntactically represented in the clause (no relative pronoun).
- b. The modifying clause has no internal morphology which restricts the interpretation of the clause in relation to the noun.
- c. An extended array of semantic and grammatical relations can be represented by the head noun.

In regard to (a), languages with strong exemplars of GNMCCs like Japanese and Korean have no relative pronoun, or other nominal marker of a specific construction like a relative clause. We interpret this to mean that there should be a range of semantic relations between the head noun and modifying clause which goes even beyond the broadest conception of relative clauses. In other words, the semantic relations between a NMCC and a head noun cannot be reconfigured by simply reconstructing the noun back into the NMCC, either directly or due to the mediation of relative forms.

For (b) Korean does have a special set of verbal inflections which appear in relative clauses, but exactly these forms are also used in all (other) types of NMCCs (Kim and Sells 2008). As far as we are aware, languages which show different forms of verbs in different NMCC types are quite rare (if there are any at all), but we wish to investigate this property.

Strictly speaking, (a) and (b) are potentially enabling properties, and then (c) should be the manifestation showing the full GNMCC character. Languages which are otherwise quite similar do appear to differ along the dimension in (c). For instance, Korean shows a tighter range of semantic relations than Japanese, which would suggest that there are some factors that license the different range of interpretive relations in the two languages.

In order to extend our understanding, we consider comparative data on a number of languages, representing potential different strategies for relativization, and with different clause-internal morphological properties. Further, we investigate nouncomplement clauses and other adnominal clauses, to try to determine whether the language has a core GNMCC or not. Our planned sample of languages for which we believe enough data is available includes Diyari (Australian, Pama-Nyungen; source Austin 1981), Godoberi (Nakh-Daghestanian, Avar-Andic-Tsezic; source Kibrik 1996), Imonda (Border, Border; source Seiler 1985), Japanese (Japanese; Matsumoto 1997), Kannada (Dravidian, Southern Dravidian; source Sridhar 1990), Kombai (Trans-New Guinea, Awyu-Dumut; source de Vries 1993), Korean (Korean; Kim and Sells 2008), Mian (Trans-New Guinea, Ok family; source Fedden 2011), Mina (Afro-Asiatic, Chadic, Biu-Mandara; source Frajzyngier and Johnston with Adrian Edwards 2005), Russian (Indo-European, Slavic; various sources), Supyire (Niger-Congo, Gur; source Carlson 1994).

Towards a comparative concept of participle

Although participles are not universal in the sense that they are present in any language, the category seems to be universally applicable, since forms traditionally analyzed as participles are found in various genetically and geographically diverse languages. However, no cross-linguistic survey based on a representative language sample has been carried out so far. This paper aims to formulate a comparative concept of participle that could then be used for such typological study. It is claimed in Haspelmath (1994) that since participles are best defined as verbal adjectives, at least languages that lack primary adjectives will also lack participles. It turns out, though, that even languages without a distinct class of adjectives can sometimes have forms that are very similar to ‘prototypical’ participles in their behaviour.

Thus, West Greenlandic, though lacking the morphological category of adjectives, possesses verb forms that are used exclusively for modification, i.e. for relative clause formation, cf. (1a). These forms are clearly subordinate, since they cannot head independent clauses, and although tense and modality can in principle be expressed within them by separate affixes, this does not happen often, cf. van der Voort (1991). The agent in such relative clauses is expressed as a possessor, which makes West Greenlandic ‘participles’ altogether quite similar to contextually oriented participles of the Altaic type, cf. example (1b) from Kalmyk:

- (1a) *angut* [naapi-ta-ra]
 man.ABS meet-PTCP.PASS-POSS.1SG.ABS
- (1b) [*mini xary-sən*] *kün*
 1SG.GEN meet-PTCP.PST man.NOM
 ‘the man met by me’

These forms at the same time have much in common with passive participles typical of the highly inflectional Indo-European languages. They agree in case and number with the noun they modify and the agent can be marked by some non-core case, cf. examples (2a) from West Greenlandic and (2b) from Russian:

- (2a) *nanog* [Piita-mit toqu-taq-Ø]
 bear.ABS Peter-ABL kill-PTCP.PASS-ABS
- (2b) *mədved’* [*ubi-t-yj*] *Petr-om*
 bear.NOM.SG kill-PTCP.PASS.PST-NOM.SG.M Peter-INST
 ‘the bear killed by Peter’

Taking into account such significant constructional similarities in languages that differ a lot in their morphological structure, I suggest leaving morphology aside for a while and taking syntax as a starting point in creating a comparative concept of participle, cf. constructional approach in Creissels (2009). The common feature of the abovementioned languages is that they have relative clauses demonstrating some degree of subordination by means of the verb form itself or the coding of its arguments, so in the current paper I investigate the properties of such relative clauses in 40 languages representing all major language families and linguistic areas and explore how these properties tend to cluster together delineating the class of units that should be further studied together as participles.

References

- Creissels, Denis (2009) Participles and finiteness: The case of Akhvakh. *Linguistic Discovery* 7.1, 106–130.
- Haspelmath, Martin (1994) Passive participles across languages. In: Fox, Barbara & Hopper, Paul J. (eds.) *Voice: Form and Function*. (Typological Studies in Language, 27.) Amsterdam: Benjamins, 151–177.
- Van der Voort, Hein (1991) *Relative clauses in West Greenlandic*. Master's thesis. University of Amsterdam.

The Verbs of Oscillation in a typological perspective

This work presents a part of a lexico-typological project studying a group of verbs describing a type of motion : oscillation verbs. This semantic field was first considered in [Рахилина, Прокофьева 2005], where it was compared between two closely-related languages, Russian and Polish. We aim to continue this study from the typological point of view in order to reveal the relevant parameters and create a reliable semantic map¹ for the semantic field in question.

By the moment, we have achieved results in several Indo-European languages (English, German², French), several Finno-Ugric languages (Finnish, Komi-Zyrian and Tundra Nenets) and also in Korean³. Normally, closely related languages are not included in selections for typological studies, since the resemblance between them can be explained by their common origin and not by universal principles of language. However, they have proved to provide very significant data when it comes to lexical typology : for example, the studies of another group of verbs of motion, rotation verbs, have shown that Russian and Polish had considerably diverged and the same semantic field is not structured by the same parameters in both languages, despite the phonetical resemblance of the lexical items.

Our approach is based on comparison of the lexical combinatory of the items that form the semantic field of oscillation in each language. The comparison reveals the relevant parameters that oppose lexemes one to another and shows the way verbs combine different meanings in each system.

One of the essential features of the semantic field is whether the system is dominant in every given language. For example, according to the previous studies, in the system of Russian there exists a dominant verb that can describe different kinds of oscillation and several lexemes with narrower use, while in Polish the whole semantic field is divided by its lexemes in almost equal zones. However, the two systems do share some features : both oppose vertical surfaces to horizontal (e.g. a curtain vs. a waving meadow) and hanging objects to those attached by the lower end (e.g. a tree), both systems have means to emphasize that an object is unstable due to its loss of functionality (e.g. a rotten chair). Russian also opposes firm objects to the soft ones, and Polish verbs can be opposed by the amplitude of the oscillation they describe.

These parameters can be found in other languages as well. Thus, Indo-European languages also oppose firm objects and soft surfaces (e.g. tissue) and vertical surfaces to horizontal, mark out the objects unstable because of a damage. In addition, these languages also provide means to point out the pattern of the oscillation or its direction. Another parameter, relevant for German languages especially, is whether the object is animated. The Finno-Ugric system is human-oriented: the essential organizing factor is the extent, to which the situation is controlled by human. There are special lexemes for animated objects and for the objects that are supposed to oscillate by their function (e.g. a cradle, a swing). Korean provides a very rich system with one dominant verb and a wide range of peripheral lexemes, allowing to oppose objects by their topological type, their firmness/softness, the pattern and direction of the oscillation, its amplitude and cause. There is also a distinction between animated and inanimated objects.

The overview of the semantic field of oscillation verbs in languages given shows that, though the systems might be organized in different ways, there are some parameters that one can trace in every language and that might therefore turn out to be universal : soft surfaces, damaged objects and animated objects tend to be isolated by special lexemes. Several characteristics of the oscillation itself, such as its amplitude, its pattern, its direction also seem to be quite important for the semantic field overall. However, every system has its own special features, and the shared features are combined with each other in different ways : most often, a lexeme combines several features of different nature,

1 see [Haspelmath, 2003]

2 based on the data from [Велейшикова, 2010]

3 based on the data from [Рудницкая, 2012]

for example, a functional feature together with a physical characteristic of the situation. Nevertheless, it seems quite possible to create a semantic map based on our results, that could afterwards be supplemented by further investigations.

References

1. Haspelmath M. The geometry of grammatical meaning: Semantic map and cross-linguistic comparison // Tomasello M. (ed.) *The new psychology of language*, vol 2. Mahwah, NJ: Lawrence Erlbaum, 2003. P. 211-242.
2. Велейшикова, Т. В. Глаголы колебания: семантика и типология: (на материале германских и славянских языков) // *Вестник Томского государственного педагогического университета*, 2010, № 7 (97), 55–60.
3. Рахилина Е. В., И. А. Прокофьева. Русские и польские глаголы колебательного движения: семантика и типология // *Язык. Личность. Текст. Сб. ст. к 70-летию Т. М. Николаевой.* / Ред. В. Н. Топоров. М.: ЯСК, 2005, 304-314.
4. Рудницкая Е.Л. Глаголы колебательного движения в корейском языке // *Вестник российского корееведения*, 2012 (в печати).

'Take' serial verb constructions in Kwa: an intra-genetic typology

The paper presents an intra-genetic typology of 'take' serial verb constructions (SVCs) in Kwa. Methodologically, it deals with a phenomenon that is far from being universal crosslinguistically, but is attested throughout this specific language family. The data of ca. 20 Kwa languages (out of ca. 80 languages currently classified as Kwa) are accessible and were compared in the investigation.

'Take' SVCs are SVCs where the verb 'take' is the first one and is, to some extent, grammaticalized. There are three types of 'take' SVCs in Kwa: (a) lative 'take' SVCs where the verb 'take' is combined with a motion verb, and the whole construction is used to express a transference of an object, as in Gen example (1), (b) instrumental 'take' SVCs where the verb 'take' is used to express an instrumental valency of the second verb, as in Gen example (3), (c) objectal 'take' SVCs where the verb 'take' is used to introduce an object of the second verb, as in Gen example (3).

- (1) ayí só agban-a vá asíme
 Ayi take package-DEF come market
 'Ayi brought the package to market'. (Lewis 1992: 110)
- (2) amejró-á só klo gban kpé-á
 stranger-DEF take knee break stone-DEF
 'The stranger used his knee to break the rock'. (Lewis 1992: 138)
- (3) ayí só te ɔ́ zo-jí
 Ayi yake yam place fire-on
 'Ayi put a yam on the fire'. (Lewis 1992: 148)

Kwa languages differ significantly in the extent of elaboration of each of these types.

Lative 'take' SVCs are absent in some Kwa languages, are attested only with inanimate objects in most of them, and are also attested with animate objects in some of them.

Instrumental 'take' SVC are attested in most of the Kwa languages to express the Instrument proper. Other meanings are restricted to some languages: the meanings of extended Instrument (including, e.g., animate 'Instruments'), Manner and other adverbials, Comitative.

Objectal 'take' SVCs are attested in all examined Kwa languages to introduce the Theme of a verb that is ditransitive outside the SVC ('give', 'show' and some others). Less frequently objectal 'take' SVCs are attested to introduce the Theme of other verbs that have other valencies (oblique or locative). Finally, in many Kwa languages objectal 'take' SVCs are attested to introduce the object of monotransitive verbs; telic agentive verbs are attested in such constructions in most Kwa languages, other volitive verbs are also attested in some languages, and involitive verbs are quite rare.

Theme session: Linked Data in Linguistic Typology

Typological NNC-Database: Storage of Cross-Linguistic Data

The paper presents a cross-linguistic database of numeral-noun constructions (NNCs). The database under development will be freely accessible online and open-ended, allowing users to perform queries and to add their own research material. A special user interface will be provided for easiness of datamanagement.

The database is aimed to hold descriptions of a certain type of syntactic constructions in terms of a large number of grammatical features (morphological, morphosyntactic and morphosemantic). The use of the constructions in each language can be illustrated with linguistic examples supplied with interlinear glosses (following Leipzig glossing rules). These data and analyses will be described with basic metadata, such as information about languages, sources and contributors.

By now the majority of databases (not only those constructed for linguistic purposes) are developed as relational databases (RDB) and managed with some kind of SQL. It is the most popular and well-tried way to organize data storage and retrieval. However, in the last years data engineers give preference to Resource Description Framework (RDF) ([Lassila and Swick, 1999]) in combination with suitable RDF query languages, especially SPARQL. Using this approach for our purposes has a number of advantages over using RDB with SQL.

For each language, the database will hold an amount of highly structured data. An NNC usually consists of two or three elements: noun, numeral and (in some languages) classifier, and up to 20 features (depending on the language) are used to describe relations between them. Furthermore, each language can have several types of NNCs, differing in word order, syntactic relations, meaning, etc. To this we must add glossed examples and metadata. Multiplying by the number of languages, we get a highly complex data structure compared to a relatively moderate amount of data. Moreover, one might want to rethink this structure as new typological data come into consideration. RDF offers a way for uniform treatment of all relations as “subject-predicate-object” triplets. Also, in an RDF database we need not create additional tables or use foreign keys to create a new relation, as we do it in RDBs. One can merely add new triples by defining new subjects, objects or predicates, as shown in [Moran 2012]. This feature of RDF-storage is especially important for open-ended projects.

Second, RDF-storage is usually queried through web-interface, so the user need not install special tools to work with it, saving time and space.

Finally, the integration with other databases is much more seamless for an RDF storage (see e.g. [Chiarcos et al. 2012]). In addition to differences in data model between RDBs, there exist many kinds of SQL, which may impede integration or migration of data. RDF, on the contrary, has higher level of standardization, which makes changing storage a simple operation.

References

- Chiarcos, C., Hellmann, S., Nordhoff, S.: Towards a linguistic linked open data cloud: The Open Linguistics Working Group. *TAL* 52(3), 245–275 (2011).
- Lassila O, Swick RR (1999) Resource Description Framework (RDF): Model and syntax specification (recommendation). <http://www.w3.org/TR/REC-rdf-syntax>
- Moran S., “Using Linked Data to create a typological knowledge base”, in Chiarcos et al. (2012), p. 129-138, 2012. Companion volume of the Workshop on Linked Data in Linguistics 2012 (LDL-2012), held in conjunction with the 34th Annual Meeting of the German Linguistic Society (DGfS), March 2012, Frankfurt/M., Germany

Mongolic clausal complementation and information source

There are two major clause complementation types in the Mongolic languages: a non-finite complement (one of about ten participles/action nouns with a case marker) and a finite one with an introducing complementizer (Buryat *geže* / Kalmyk *giž* / Khalkha *gež*, an imperfective converb of an auxiliary quotation verb *ge-/gi-*). The third type, with governed postpositions such as e.g. Khalkha *tuxai* ‘about’, comprises only a few constructions.

We argue that the first two types distinguish meanings related to the information source. Our analysis based on a text corpus and fieldwork material shows that not every predicate taking a clausal complement can take both these types. In this paper we will show the distribution of complement types with verbs of different semantic domains.

The *geže/giž* complementizer (< ‘saying’) is grammaticalized on the base of a direct speech construction with verbs of speaking as main predicates. Due to grammaticalization its scope is also expanded to verbs of cognition, emotion and even perception.

With verbs of cognition, the *ge-/gi-*complement indicates that the speaker’s knowledge is indirect (the result of complex logical operations, common knowledge etc.); some verbs – like Kalmyk *sana-* ‘think’ - only allow the *ge-/gi-*complement. With verbs of emotion it signals that the stimulus is not an actual event, but a mental construct (cf. “emotions caused and emotions projected”, Bolinger 1984), such that it is mostly used with verbs like ‘hope’ and ‘fear’ e.g. Buryat *ai-* ‘fear’:

- (1) *Buuda-xa-ny geže ai-ba gü?*
 shoot-FUT-3sg CMPL fear-PST Q
 ‘(She) was afraid that (he) will shoot?’

Typical for verbs of perception are non-finite complements in the accusative, constructions with the *geže/giž* complementizer being quite seldom; the distinction here is between immediate perception and “mental perception” (Verhoeven 2007:293). With verbs of hearing it is auditory vs. hearsay, e.g. Kalmyk *soŋs-* ‘hear, listen’:

- (2) *Zal-d bää-sn uls Kugultinov šülg-üd-än*
 auditorium-DAT be-PP people K. poem-PL-REFL
umš-s-ig soŋs-v
 read-PP-ACC hear-PST
 ‘The people sitting in the auditorium listened to Kugultinov reading his poems.’

- (3) *Xalx Monŋh-yn poličišk boln olna üüldäč-nr dund*
 Khalkha Mongolia-GEN political and social leader-PL among
dörvd jas-ta uls oln bilä giž soŋs-la-v
 Dörbet bone-COM people many PTCL CMPL hear-PST.EVID-1sg
 ‘I heard that there were many Dörbet people among Khalkha-Mongolian political and social leaders’.

With Kalmyk verbs of seeing only non-finite complements with the accusative (immediate perception) are grammatical; in Buryat also finite complement clauses with *geže* are possible, signalling interpreted information (obtained e.g. through inference); Buryat *xara-* ‘look, see’:

- (1) *Butid Tagar-ai myaxa sabša-x-iye-ny xara-na*
 B. T.-GEN meat chop-PTCP-ACC-3sg look-PRS:3sg
 ‘Butid watches how Tagar chops meat.’

- (2) *yamar* *xemžee* *ab-aa-b* *geže* *xara-ža* *üze-xe*
what measure take-PST-PTCL CMPL look-CVB see-PTCP
xeregtei-l
necessary-PTCL
'It is necessary to take a look at what measures are taken'.

Therefore, the opposition of the two complementation types with mental and psychic verbs in Mongolic can be seen as an opposition between a firsthand and a non-firsthand information source in dependent clauses ("evidentiality strategy", Aikhenvald - Dixon 2003:18).

Theme session: Quantitative Linguistic Typology: State-of-the-Art and Beyond

Quantifying phonological complexity and its distribution

Contrary to an earlier orthodoxy that “all languages are equally complex” it is now generally agreed that spoken languages vary considerably in their complexity (Miestamo 2008, Sampson et al 2009, McWhorter 2011). Of course, either position presupposes that some measure or measures of complexity can be meaningfully calculated. With respect to phonological properties several well-developed notions have been proposed as appropriate indices of complexity (see, e.g. contributions to Pellegrino et al 2009). These include simple measures such as the size of consonant, vowel and tone inventories and the variety of permitted syllable structures, as well as somewhat more elaborate calculations taking into account the relative frequency with which contrasts are exploited, relative entropy of phonological information, or relative frequency of occurrence of simpler versus more complex elements. In this paper a number of these measures will be used, first of all, to support the view that phonological complexity in languages is not uniform but varies considerably. Rather, languages appear to be broadly distributed along any scale of phonological complexity (but see Pellegrino et al 2011). Two issues will be particularly explored. The first is primarily methodological and concerns how large geographically and genetically diverse samples of languages can be assembled in a standardized fashion so that comparisons across the languages are valid. Exemplification will include discussion of problems involved in harmonizing analyses of segment inventories and syllable structures, and calculating the length of phonological words. Inventories can be more reliably assessed if closer-to-surface rather than abstract forms are compared. Word-length is best compared when standardized text samples are available, rather than relying on lexical entries. The second issue concerns what factors might underlie the distribution of phonological complexity. Explanatory principles proposed include, among other factors, population size and isolation (e.g. Trudgill 2004, 2011, Hay & Bauer 2007), environmental setting (e.g. Munroe et al 1996 and later, Fought et al 2004 and later), social behavior patterns (e.g. Ember & Ember 2010), and decline of phonological diversity correlated with decline of genetic diversity as modern human populations dispersed from an African origin (Atkinson 2011). The last of these proposals can probably be dismissed since it is not motivated by any independent rationale, and fails to offer any explanation for the initial assumption of maximal complexity in Africa. The varied environmental and social factors suggested seem more promising but generally raise the same question: what time point in the history of the languages concerned is relevant? For example, catastrophic population crashes in Australia and the Americas following European colonization have no obvious effects on the phonological complexity of those languages which survived, nor has the phonological complexity of Spanish been clearly affected by its spread to tropical, sub-tropical and near-polar climates. Using maximally large sample sizes of languages may enable robust correlations to emerge from the noise contributed by such historical accidents. At the least, significant correlations can be demonstrated between some aspects of phonological typology and broad climatic/vegetational zones.

Polysemy of ‘again’-markers: evidence from cross-linguistic data

The paper deals with universal constraints on types of polysemy of ‘again’-markers (see Wälchli 2006). Uses of these markers cover the repetitive semantic domain that consists of the central meaning of one-time repetition of an event – repetitive proper (‘to do one more time’) and some other particular meanings such as reditive (‘to move back’), restitutive (‘to return to the same state’), reconstructive (‘to do over, to do better, than before’), responsive (‘to do in response’) and so on. Repetitive meanings are closely related to the idea of one-time repetition and often co-expressed by the same marker.

‘Again’-markers often reveal also polysemy with meanings from other semantic domains, such as verbal plurality meanings, valence-increasing meanings, phasal meanings and some others. The range of meanings that tend cross-linguistically to be co-expressed with repetitive ones and semantic links between them are discussed in the paper.

The study is based on a sample of languages of different genetic and areal affiliation (119 languages). The main data source is grammar descriptions and dictionaries.

The following meanings are particularly taken into consideration:

a) spatial meanings:

The repetitive domain links to the spatial one through the reditive meaning (‘to move back’) which belongs to both of them. In some languages (cf. *-irtne* in Aranda, *-err* in Yanesha) this meaning expand to non-spatial verbs, giving the meaning of associated motion (‘to do on the way back’). Repetitive meanings (especially reconstructive) can also develop from prolativ (cf. *pere-* in Russian, *gwo* in Cantonese).

b) verbal plurality meanings:

Rendering itself a single repetition of event, the repetitive meaning can be combined with meanings that indicate other types of repetition, such as multiplicative (‘an event represented as a series of repeated portions’): cf. *-si-* in Mongsen Ao (Tibeto-Burman) and distributive (‘an action with a plural participant’), cf. *-li-* in Azoyú Tlapanec (Otomanguean). The same marker in some languages can be used in repetitive contexts and in habitual ones (‘to do regularly’): cf. *té* in Nateni (Gur) or *n-* prefixes in Athapaskan languages. The polysemy of repetitive with meanings of many times repetition is however not so frequent as it could be expected.

c) reversive meaning:

There are also cases of polysemy of the repetitive meaning and the reversive one (‘undoing an action’), cf. *-t* in Fula.

d) aspectual meanings:

‘Again’-markers do not tend to have aspectual uses. An exceptional case is the phasal domain: the repetitive meaning is often co-expressed with the continuative one (‘still, keep on doing’): cf. *ancora* in Italian.

e) valency-changing meanings:

‘Again’-markers can be used in reflexive context (it is typical e.g. for Oceanic languages). Reciprocal uses of ‘again’-markers are also attested (cf. *taligu* in Lolovoli).

The polysemy of ‘again’-markers reveals some features that are of interest in context of typology of grammar and grammaticalization theory. In this case one deals not with inflectional markers, but with derivational affixes or free adverbs and particles (less studied from this point of view) and it seems to leave its imprint. E.g. lots of cases can be interpreted as “ruins” of polysemy: the meanings are coexpressed synchronically, but diachronically they both link to a lost meaning and not to each other.

References

Wälchli, Bernhard. 2006. Typology of light and heavy ‘again’, or the eternal return of the same // *Studies in Language* 30 (1). P. 69–113.

Theme session: Lexical typology of qualitative concepts

Semantic domains 'full' and 'empty': a cross-linguistic study

This research is devoted to typological description of semantic domains 'empty' and 'full'. These domains possess unique features both on syntactic and semantic levels. Situations related to these domains include two semantic arguments (container and contents); it distinguishes them from the majority of other situations described by qualitative lexemes. These properties influence both syntactic features of the adjectives in question and semantic structures of the corresponding domains. The research was carried out on the material of 10 languages including languages of not closely related families (e.g., Chinese, English and Armenian) and closely related languages, such as Russian and Serbian. The latter revealed some important basic oppositions, which proves the relevance of closely related languages material for lexical typology.

The basic meaning of the domain '**empty**' is lack of contents in a container. The prototypical context is *empty X (of Y)*, where X is a noun, denoting container (e.g., a glass, a bottle, a box), and Y is the lacking contents. The basic opposition in this domain is the opposition of form and function. In the majority of the languages in our study this opposition is expressed lexically, i.e. the two zones are covered by two different lexemes (cf. Russian *по́лый* vs. *пустой*, English *hollow* vs. *empty* & *blank*, Armenian *բ'ւ՛ճ'* vs. *datark*, Spanish *hueco* vs. *vacío*, Greek *κούφιος* vs. *ἄδειος*). Note that the notion of 'empty' possesses higher cognitive relevance than the notion of 'hollow'. This reflects both in higher frequency and combinability of the corresponding lexemes and in larger amount of possible metaphoric extensions. The structure of the domain is rather complex and includes a number of oppositions, which are less frequent from the typological point of view, among them are the type of contents (cf. Serbian *пуста* 'without people' vs. *празан* 'without any objects or substances' [Tolstaya, 2008]) and the type of container (cf. Korean *thengpita* 'empty (prototypical container)' vs. *pita* 'empty (bracket or surface)' vs. *konghehata* 'empty (space)').

The basic meaning of the domain '**full**' is 'filled to utmost capacity'. The basic opposition in this domain is the opposition of qualitative and quantitative meanings (the adjective describes the state of the container or the quantity of contents). This opposition can be expressed both lexically (cf. Spanish qualitative *lleno* vs. quantitative *completo*), and syntactically by different constructions with one lexeme. In English we see this opposition not only on syntactic level, but also on the level of word formation (cf. two different models with suffix -ful (qualitative *beautiful* vs. quantitative *a spoonful of...*). Other oppositions relevant for this domain are: form vs. function, type of container and the degree of fullness.

The domains under study have a complex structure with a number of oppositions. In some languages none of the oppositions are expressed lexically (cf. Serbian *пун*, covering the whole domain of 'full'), in others, on the contrary, almost all the oppositions are realized (cf. Spanish *lleno*, *repleto*, *completo*, *pleno*, *integral* for 'full'). The systems of the first type are called poor; the systems of the second type are called rich (after [Maisak, Rakhilina eds. 2007]), but according to our data, the most frequent are average systems containing two or three lexemes (Russian, Armenian, Greek).

References

Majsak T., Rakhilina E. (eds) 2007. *Verbs of AQUA-motion: lexical typology*. Moscow, Indrik Tolstaya S., Prostranstvo slova. Lexicheskaya semantika v obschslavianskoy perspective. M, 2008

On the emergence of finite structures from non-finite constructions:

Evidence from ‘say’ constructions in Japanese

In this paper, we will examine how some ‘say’ constructions in Japanese develop finite structures from non-finite ones. Diachronic data for our analysis come from the Taikei Honbun Database based on texts from the 8th to 19th century. We will focus on the perfective *ihe-* and imperfective *ihi-* counterparts of the ‘say’ verb *ifu*. Our analysis reveals that the rise and fall of the *kakari musubi* system (see Ono 1993), which makes a distinction between attributive and conclusive forms, triggered a chain of events in which attributive (non-finite) quotative and evidential ‘say’ forms came to be used as conclusive (finite) structures as well. This extension contributed to the demise of the *kakari musubi* system. Crucially, from a typological perspective, this extension provides additional insight into strategies by which relativization and nominalization constructions develop into finite clauses (see DeLancey 2011).

Our findings show that between the 8th to 19th centuries there were three major waves of attributive/conclusive contrastive forms within the ‘say’ constructions in Japanese. The *to ihe-ru/to ihe-ri* forms, which emerged within the perfective *to ihe-* series, showed a clear attributive/conclusive contrast in the 8th century, but the attributive form soon extended via concessive uses, attested in the 10th century, to conclusive contexts by the 13th century. This gradual blurring of the attributive/conclusive distinction contributed to the disappearance of *to ihe-ru/to ihe-ri* forms in the 18th century, and ultimately to the demise of the *kakari musubi* system as a whole.

The other two attributive/conclusive contrastive forms emerged within the imperfective *to ihi-* series. Tokens of the *to ihi-keru/to ihi-keri* distinction were first attested in the early 10th century but the attributive *to ihi-keru* form had already developed conclusive uses as well, thus showing signs of an already blurred attributive/conclusive distinction, with the *to ihi-keru* form disappearing in the 17th century while the *to ihi-keri* form lingered on into the 18th century. A similar fate befell the *to ihi-taru/to ihi-tari* distinction first attested in the 10th century, with the *to ihi-taru* form disappearing in the 16th century and the *to ihi-tari* form surviving longer into the 18th century.

In this paper, we will also examine evidence from the early 10th century showing attributive *to ihi-taru* forms accompanied by particles frequently associated with nominal constructions. As illustrated in (1) with *ni*, these particles that typically mark nominal expressions were also being used as markers of subordinate clauses. The presence of these particles provides evidence of a link between the attributive and nominalization and relativization constructions, as well as subordinated clauses, and the drift from attributive-toconclusive uses seen in Japanese provide additional support for the Nominalist Hypothesis that non-finite nominalization constructions frequently develop into stand-alone finite clauses (see Starosta, Pawley & Reid 1982; Kaufman 2009; Yap, Grunow-Harsta & Wrona 2011).

- (1) *“Kore wa ikaga” to ihi-taru ni,*
 this NOM okay QT.ATTR PRT
 ‘When (X) said “How is this?”’,

tada, “Hayaku ochi ni keru.” to irahe tareba
 just early fall PRT PERF” COMP reply PERF.COND
 ‘(Y) simply replied “The flower had fallen early.”’

(*Makura no Sooshi*, 996 AD, pp.165)

References

- DeLancey, Scott. 2011. Finite structures from nominalization constructions in Tibeto-Burman. In *Nominalization in Asian Languages: Diachronic and Typological Perspectives*, Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds), pp. 343-359. Amsterdam: John Benjamins.
- Kaufman, Daniel. 2009. Austronesian typology and the nominalist hypothesis. In *Austronesian Historical Linguistics and Culture History: A Festschrift for Robert Blust*, Alexander Adelaar & Andrew Pawley (eds). Canberra: Pacific Linguistics.
- Ono, Susumu. 1993. *Kakari musubi no kenkyuu* (A study of Kakari musubi). Tokyo: Iwanami Shoten.
- Starosta, Stanley, Andrew K. Pawley, and Lawrence A. Reid. 1982. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics, Volume 2: Tracking the Travelers*, Amran Halim, Lois Carrington & S.A. Worm (eds), pp. 145-170. Canberra: Pacific Linguistics.
- Yap, Foong Ha, & Karen Grunow-Hårsta & Janick Wrona. (2011). Nominalization strategies in Asian languages. In *Nominalization in Asian Languages: Diachronic and Typological Perspectives*, F.H. Yap, K. Grunow-Hårsta & J. Wrona (eds), pp. 1-57. Amsterdam: John Benjamins.

Generic markers: a basic questionnaire and experimental toolkit

Introduction. The past decades have witnessed a surge of interdisciplinary interest in genericity disciplines: psychology, philosophy, and formal linguistics. However, languages that are the target of in-depth studies lack dedicated generic markers. My presentation sketches the state of the art in the study of generic markers in the languages of the world. I propose a basic questionnaire and an experimental toolkit for identifying and describing generic markers.

The phenomenon. Genericity covers two phenomena: a) reference to kinds, and b) propositions expressing general properties of kinds. Instances of generic sentences are given in (1a-3a), contrasted with episodic sentences (1b) and sentences that reveal some puzzling quantificational properties of generic sentences (2b, 3b).

- | | | |
|-----|------------------------------------|---|
| (1) | a. <i>Tigers have stripes</i> | (versus b. <i>This tiger is albino.</i>) |
| (2) | a. <i>Ducks lay eggs</i> | (versus b. ?? <i>Ducks are female.</i>) |
| (3) | a. <i>Ticks carry Lyme disease</i> | (versus b. ?? <i>Ticks do not carry Lyme disease.</i>) |

Generic sentences express general properties (1a-3a) as opposed to episodic facts (1b), and they typically convey the commonsense knowledge about the (“our”) world.

The interpretation of generic sentences has puzzled logicians and philosophers for decades. Philosophers of language have long realized that generic sentences cannot be analysed by standard logic approaches. Assigning truth values to generic sentences, as in (1a)-(3a) above, is problematic. For instance, there are less ducks that lay eggs (2a) than there are female ducks (2b). The sentence ‘Ticks do not carry Lyme disease’ in (3b) is false, while ‘Ticks carry Lyme disease’, as in (3a), is true, although only a small percentage of all ticks carry Lyme disease. Despite the problems of establishing the formal semantics of a possible “generic operator”, paradoxically, humans understand the generics without efforts and young children understand and produce generics earlier than they do explicitly expressed quantifiers.

Generic markers are hard to identify, because generic interpretations tend to be the default interpretations of grammatically unmarked sentences or, at least, generic sentences tend to show default or less complex marking as opposed to non-generic marking. In addition, it is hard to find dedicated generic markers that distinguish the sentences in (1-3a) from episodic sentences, because the generic versus specific distinction may cross-cut several grammatical categories. Even if the intuitive distinction between generic and non-generic is clear, generic forms covary with other grammatical categories. Genericity markers are found among definiteness, number, or quantifiers, impersonals, non-finites and nominalizations, aspect, tense, mood, evidentiality, and modal markers. They may mark nouns, verbs, or adjectives.

Why are there so few generic markers, is genericity a grammatical category at all?

The lack of special generic markers seems to be strongly socially cognitively motivated. Developmental psychologists have proposed that the capacity to generalize communicated information is innately given. Children are biased to receive general knowledge in non-linguistic communicative situations, expecting to be taught the knowledge that transcends individual experiences in a cultural community.

Methods for typology. Linguists have been hesitant to address “the generic category”, since there is too little evidence for one-to-one mapping between the generic form and meaning across languages. Generic markers are also frequently optional. Drawing upon the insights of recent interdisciplinary, concept-based and pragmatic approaches to generics, I present some methods – a questionnaire and a set of simple experiments – of capturing the form, meaning, and use of generics in particular languages.

Demonstratives in Cebuano: Referential and non-Referential Functions

This paper investigates the uses of the demonstrative forms in Cebuano (see Table 1), in particular their extension from referential to non-referential functions. Cleary-Kemp (2007) identified four basic functions of demonstratives in many Austronesian languages, one referential and the other three non-referential, namely, the discourse deictic use, the "tracking" use, and the "recognitional" use. In addition to these commonly attested non-referential functions, other grammaticized uses of Cebuano demonstratives are identified, namely, the placeholder function in repair situations and weak stance use in hesitations, the latter of which is not commonly attested or reported in many languages.

This paper thus has a twofold goal. First, to highlight the division of labor between the various forms of demonstratives in the acquisition of non-referential functions; that is, each of the demonstrative forms develops a distinct non-referential use. For example, the recognitional use is served by the Distal *kato*, while the placeholder function is taken up by the Near-Hearer *kana'*, which pairs with the dummy word *ku'an* in the organization of repair. Second, to demonstrate the weak stance function as a metaphoric extension of the placeholder function; that is, where *ku'an* is deleted and the demonstrative form *kanang* (< *kana'* – *ng* 'this-Linker') alone remains, the latter serves as a stance marker, especially in instances where the speaker is uncertain and hesitates or is obviously weakly committed to the proposition expressed in the main clause. Actual spoken data will be examined and used to illustrate the syntax, semantics, and pragmatics of these forms.

Table 1. Demonstratives in Cebuano (and their extended non-referential uses)

	nominative	oblique
Near Sp (and Hr)	<i>(ka)ni</i> <i>(ki)ni</i>	<i>(ni)'ini</i> <i>(ni)'ani</i>
Near Hr	<i>(ka)na'</i> → placeholder → hesitation marker	<i>(ni)'ana'</i> → extended argument in EICs
Far	<i>(ka)tu</i> <i>(kad)tu</i> → recognitional use	<i>(ni)'atu</i> <i>(ni)'adtu</i>

References

Cleary-Kemp, Jessica. 2007. Universal uses of demonstratives: Evidence from four Malayo-Polynesian languages. *Oceanic Linguistics* 46.2: 325-347.

Properties of possessive HAVE

Possessive predication typologies distinguish two classes of HAVE-type verbs: A(ction)-HAVE descends historically from action verbs meaning “hold”, “take”, etc. (e.g. English *have*, Latin *habeo*) and is assumed to be transitive. E(xistential)-HAVE shows an existential verb form (e.g. Mandarin *yǒu*, Malay *ada*), and is assumed to be intransitive (Heine 1997, Stassen 2009) This paper argues, in contrast, for a unified possessive HAVE across languages with such a verb, diachronic or synchronic associations notwithstanding, which (i) is two-place, describing an underspecified possessive relation, and (ii) performs a function of presentational focus, thus showing a definiteness effect (DE).

(1)-(3) provide initial evidence for (i). Possessive HAVE in English, Mandarin, and Malay encodes both alienable and inalienable possession. Actual interpretations depend on whether the possessee nominal is relational (inalienable e.g. *sister*: kinship, *thumb*:part-whole) or non-relational (alienable e.g. *car*:ownership/control), indicating HAVE is underspecified.

(1) John has a/#the/#every car/sister/crooked finger. English

(2) S⁻anm[´]ao yǒu li^ˇang/#n[`]a ge ji^ˇejie/b⁻ei-b⁻ao/m^ˇu-zh^ˇi
Sanmao have two/that CL elder.sister/back-pack/thumb
Sanmao has two/#those elder sisters/backpacks/thumbs. Mandarin

(3) a. Ali ada enam kereta/anak/jari Ali have six car/child/finger Ali has six cars/children/fingers.	b. .#Ali ada kereta/anak/jari itu Ali have car/child/finger that) Intended: Ali has that car/child/finger. Malay
---	--

(1)-(3) also demonstrate the DE of possessive HAVE (Partee 1999). For felicitous interpretation, definite or “strong” NP (Milsark 1974) complements as in *John has the sister* require a context e.g., of planning a VIP visit, where *John* is assigned to entertain the VIP’s sister. Indefinite complement HAVE, however, yields possessive readings both alone and in contexts licensing definite complements: In the VIP context, *John has a sister* could mean *John* will entertain a VIP sister, or be possessive (*John has a sister, so he’ll show the ladies around*). The same effects are found in Mandarin and Malay.

English *have* is disallowed in the existential construction (4a), unlike Mandarin *yǒu* (4b) and Malay *ada* (4c):

(4) a. There are/*have children nearby!	b. (zh`er) yǒu r`en! here have person There’s someone (here)!
---	---

c. Ada lipas di atas meja
have cockroach at top table
There’s a cockroach on the table.

Yet possessive HAVE sentences in all three languages exhibit the DE. It is possible to attribute the DE in Mandarin and Malay to the status of *yǒu* and *ada* as existential verbs, but this reasoning is not possible for English. The DE of English *have* supports the current proposal – that English, Mandarin, and Malay uniformly show possessive HAVE despite their differing affinities with action verbs or existential verbs.

I argue *yǒu* and *ada* are polysemous between possessive and existential senses, further supporting a possessive HAVE in Mandarin and Malay. For example, verb-initial *yǒu* sentences yield existential or

null subject (arbitrary reference) possessive interpretations, depending on the complement nominal (5). Relational nominals, e.g. sisters, allow only possessive interpretations. Non-relational NPs describing common possessions e.g. cars, yield ambiguity. Items not usually possessed by individuals, e.g. trains, yield an existential interpretation. Since the same surface form allows both interpretations, this indicates possessive *yǒu* and existential *yǒu* are distinct. I demonstrate the same for Malay. Thus in both Mandarin and Malay, a purported E-HAVE shows properties of a two-place possessive HAVE, supporting (i).

- (5) *yǒu* *jiěmei* / *chē* / *huǒchē* *zhēn* *hǎo*
 have sister / car /train true good
It's good [to have a sister/*that sisters exist]/ [to have a car/that cars exist] / [that trains exist/*to have a train]

Moreover, I show with anaphora facts that in Mandarin, the possessor (Psr) nominal in possessive *yǒu* sentences (2) is a grammatical subject, but in existential *yǒu* sentences, the optional pre-verbal location phrase (4b) is not a subject, further distinguishing existential and possessive *yǒu*. Freeze (1992) considers possessive sentences to be structurally identical to existentials, thus accounting for their DE, but this account cannot distinguish between the subjecthood properties of the preverbal nominal in Mandarin possessive and existential *yǒu* sentences. Stassen (2009) assumes that E-HAVE forms intransitive possessive sentences where the Psr is a topic, but argues that topic possessives may develop into transitive possessives where the Psr is a subject. The current proposal is compatible with this understanding, but does not require this to be the case. It predicts, however, that if such a development does occur, the resulting HAVE verb would show properties (i) and (ii).

Proposal (ii), that possessive HAVE is presentational, receives support from the DE of English A-HAVE, and makes predictions about possible developments of A-HAVE. The DE in existential sentences may be attributed to a presentational function of (re-)introducing a discourse participant (Abbott 1993). Treating possessive HAVE as also presentational not only accounts for the DE of English *have*, it also predicts that A-HAVE may develop an existential sense, exemplified by Serbo-Croatian *imati* (Creissels 2010).

In conclusion, possessive HAVE demonstrates both E-HAVE and A-HAVE features. This work refines existing typologies by showing that possessive HAVE need not be distinct from E-HAVE, which in turn need not always indicate a topic possessive.

A statistical association between head-complement orders and word-stress location

The correlation between morphosyntax and phonology has been one of the exciting but controversial topics in linguistic typology. Especially, it has been pointed out that languages with lefthand word-stress have head-final order (e.g. object-verb) while languages with righthand word-stress have head-initial order (e.g. verb-object) (Bally 1944, Lehmann 1973, Donegan and Stampe 1983). However, there has been no statistical analysis of the correlation with database in the world's languages. In this paper, we show our analysis of the word order data in Dryer (2008) and the word-stress data in Goedemans and van der Hulst (2008).

A statistical approach was taken to assess factors influencing on the head-complement orders in the language genera of the world. Using a dataset of 890 cases extracted from WALS Online (2008), we conducted a multinomial logistic regression analysis whose response variable was the observed head-complement order: head-initial, head-final, and either order being equally possible. The explanatory variables were the head-complement types (affix-stem, noun-genitive, adposition-NP, verb-object, and adverbial subordinator-clause) and three features of word-stress location: the fixedness (fixed/weight-sensitive), directionality (left/right), and the size of stress window (the maximal number of syllables to be searched from the relevant word-boundary).

The results of our statistical analysis show that head-complement type is the most important factor influencing on their order, as indicated by the smallest *p*-value of log-likelihood ratio test in Table 1.

Furthermore, we can see a clear increase of the proportion of head-initial order accordingly as the categorical hierarchy gets higher from affix-stem to adverbial subordinator-clause (Figure 1).

Our analysis confirms that two variables relating to word-stress placement are also significant determinants of head-complement order. As indicated in Table 1, head-complement order is strongly associated with the directionality of stress location: the genera with righthand word-stress show much higher proportion of head-initial order than those with lefthand one (Figure 2).

Weight-sensitive stress has turned out to be explanatory as well, having the effect of increasing the proportion of head-final order, compared with fixed stress (Figure 3).

The factor of stress window, although explanatory in isolation, did not remain explanatory in our main effect model.

These statistical findings are explained as follows. Constituents with head-final order are more tightly connected than constituent with head-initial order. Head-initial constituents may contain head-final constituents, but not vice versa (cf. Biberauer et al. 2008). This explains why the rate of head-initial orders against head-final orders increases as complement gets larger.

Assuming that stress universally falls on complement rather than head (Nespor and Vogel 1986 and Cinque 1993), head-final (complement-left) constituents must have stress on the left. Head-final constituents are (compound) word-like because of its strong juncture. Then, lefthand stress in head-final constituent must match lefthand word-stress. Weight-sensitive stress is more flexible in stress location than fixed stress. Thus, we can explain why word stress correlates with head-complement orders.

Table 1 The overall effect of the explanatory variables

Variable	LR chi ²	df	p-value
head-complement type	90.11	8	4.441e-16
lefthand/righthand word-stress	36.49	4	2.290e-7
fixed/weighted stress	9.63	2	0.008
stress window	2.70	2	0.259

Figure 1 Proportion of word order by head-complement type

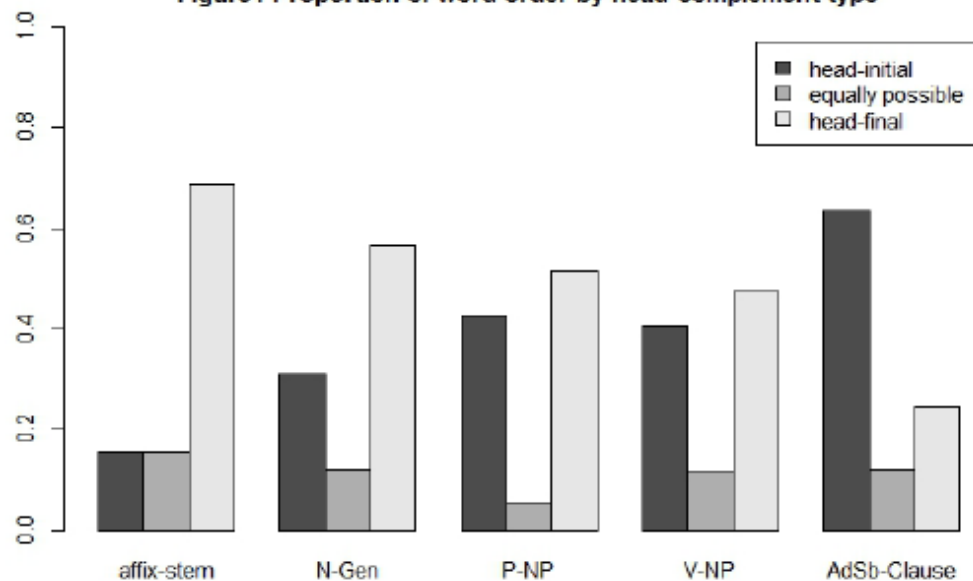
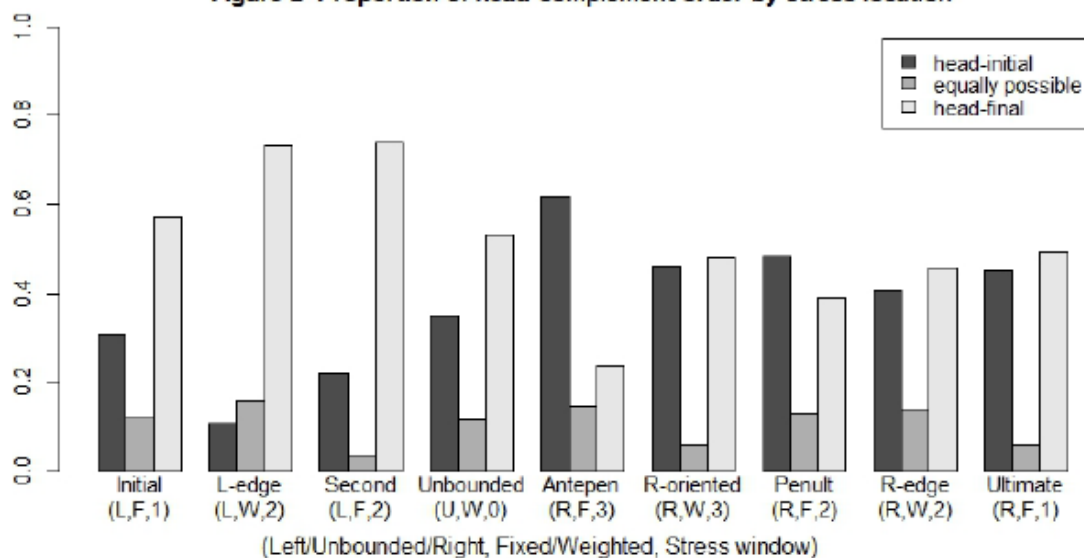


Figure 2 Proportion of head-complement order by stress location



Adjoined "complement" clause in Seediq?

Seediq is an Austronesian language spoken in Taiwan. Seediq has several types of complementation: verb serialization, VP embedding, clause-embedding, and clause-adjoining. Interestingly, a complement of predicate of knowledge cannot only be expressed by clause-embedding, as in (1), but also by clause-adjoining, as in (2).

- (1) me-kela=ku ka [m-usa sapah rubiq ka kumu].
 AV-know=1s.NOM CMP AV-go house Rubiq NOM Kumu.
 'I know that Kumu went to Rubiq's house.'
- (2) m-usa sapah rubiq ka kumu 'u, me-kela=ku ka yaku.
 AV-go house Rubiq NOM Kumu CNJ AV-know=1s.NOM NOM 1s
 'I know that Kumu went to Rubiq's house.'

Here I discuss only the above two types of clause combining.

Clause-embedding is used with predicates of knowledge and acquisition of knowledge, utterance predicates, and immediate perception predicates (see Noonan (2007: 120-145) for the classification of complement-taking predicates). The embedded clause subject is independent from the main clause subject. The predicate of the embedded clause may be Neutral, Perfect or Future. It is possible for the particle *ka* to appear before the embedded clause, as a complementizer.

Let us move on to clause-adjoining. Since Seediq is a verb-initial language, the preceding clause that is adjoined does not end in a verb. Between adjoined clauses, there is usually a sentence medial conjunction and a pause following it. Adjoining via juxtaposition with a non-final pause is also possible. There are four sentence medial conjunctions: *'u*, *de'u*, *ni* or *deni*. Their choice depends on the semantic context but it is not always clear-cut. Here I will focus on adjoining via *'u*. Clause-adjoining via *'u* is used in several contexts: contrast, overlap, conditionality, causation, concession, commentative, content of knowledge, and correlative. A complement of predicate of knowledge can also be expressed by clause-embedding, as we saw in the previous paragraph.

At least notionally, we can say that the preceding clause that is adjoined in example (2) is a 'complement' of the predicate verb of the following clause that is adjoined. It seems inadequate, however, to regard it as the syntactic 'complement' of *me-kela* "to know" in the following clause that is adjoined. Sentence (3) is another example of clauses adjoined by *'u*, which is used in a context of contrast.

- (3) t-em-egesa 'uyas kelemukan ka tiwaN 'u,
 AV-teach song Taiwanese NOM Ciwang CNJ
 t-em-egesa 'uyas nihuN ka daway 'uri.
 AV-teach song Japanese NOM Daway also
 'Ciwang taught Taiwanese songs, and Daway taught Japanese songs.'

As you can see from this example, adjoining via *'u* does not guarantee that the entity that precedes the conjunction *'u* be the 'complement' of something that comes after *'u*.

References

Noonan, Michael. 2007. 'Complementation'. In Shopen (ed.). 52-150.

Shopen, Timothy (ed.). 2007. *Language Typology and Syntactic Description. Vol.II: Complex Constructions*. Cambridge: Cambridge University Press.

An unusual path to evidentiality: evidence from Kakataibo

Evidentials grammaticalize from a number of paths including verbs, deictic and locative markers and copula constructions (Aikhenvald 2004:271). Where verbal sources are involved, reportative/indirect and visual/direct evidentials typically derive from verbs of saying and verbs of perception (to see), respectively, by the reanalysis of biclausal constructions that have one of these verbs with a complement clause (Aikhenvald 2004:272-73). Kakataibo, a Panoan language spoken in the Peruvian Amazon, shows a strikingly mirror-image pattern: Contrary to cross-linguistic expectations, in which direct evidentials are associated with visual evidentiality and indirect with reported, the verbs ‘to say’ *ka* and ‘to see’ *id* are apparently the historical sources of the direct and indirect evidentials (1), respectively. This unusual historical outcome is arguably the result of a reanalysis of performative constructions, in which ‘to say’ was associated with first person and ‘to see’ with third person (Shell 1978)¹.

(1a) Norua=∅=id-a kuan-dza
 Norua=S=IND.EV-3 go-PST.3
 ‘It is said that Norua left.’

(1b) Norua=∅=ka-a kuan-dza
 Norua=S=DIR.EV-3 go-PST.3
 ‘Norua left.’

This unusual path of grammaticalization, in which evidentials have arisen from performative constructions, has arguably given rise to the typologically rare synchronic behavior of these evidentials in Kakataibo (as documented in the author's fieldwork on this language). With respect to their morphosyntactic distribution, evidentials are second-position clitics that are obligatory in every main clause. Since =*ka* marks information from direct evidence as well as inferences, its use is pervasive in the language, making =*ka* the unmarked term in the evidentiality system. However, this behavior of =*ka* goes against the iconicity principle, which would predict leaving the direct evidential formally unmarked. With respect to function, the direct evidential in Kakataibo acts as a marker of quotations and inferences, whereas this task normally falls to reported/indirect evidentials cross-linguistically (Aikhenvald 2004).

This paper studies the evidential system in Kakataibo in both synchronic and diachronic terms. By doing so, it is shown that the Kakataibo evidential system does not fit current typologies (e.g. Aikhenvald: 2004) with regard to the functions encoded by evidentials; that is, the marking of inferences and quotations may be marked by evidentials other than indirect/reportative in a two-way evidential system. Consequently, the current typology of evidentials needs to be expanded to allow for more fine-grained distinctions such as those shown here. In addition, the Kakataibo data provides evidence for the more general implication that unusual paths of grammaticalization can lead to unusual synchronic behavior of categories.

References

- Aikhenvald, Alexandra. 2003. ‘Mechanisms of change in areal diffusion: new morphology and language contact’, *Journal of Linguistics* 39: 1–29.
- Aikhenvald, Alexandra. 2004. *Evidentiality*. Cambridge, Cambridge University Press.
- Shell, Olive. 1978. ‘Los modos del Cashibo y el analisis del performativo’. In *Estudios Pano I*, p.23-62. SIL, Yarinacocha.
- Valenzuela, Pilar. 2003. ‘Evidentiality in Shipibo-Konibo, with a comparative overview of the category in Panoan’, in Aikhenvald and Dixon (eds.) (2003: 33–62).

Possessive marking in nominal/verbal contexts in Oceanic languages

Possessive marking is a prototypically ‘nominal’ feature, used in referential phrases headed by a thing- or person-denoting word. However, possessive marking also occurs in non-prototypical ‘nominal’ constructions like action nominalizations (referential phrases headed by an action-denoting word), as well as in main clauses (predicative phrases headed by an action-denoting word) instead of a distinct ‘verbal’ paradigm for dependent marking (Siewierska 1998; Croft 2001).

We study the distribution of possessive marking over the three construction types mentioned above, in 25 Oceanic languages of various sub-families. We focus on Oceanic languages for two reasons: First, they display multiple possessive marking strategies, associated with various degrees of (in)alienability (Lynch et al. 2002: 40; Lichtenberk 2009). Second, they are typologically remarkable in terms of lexical classes, specifically regarding the distinction between nouns and verbs: both referential phrases and predicative phrases accommodate thing/person-denoting words (semantic nouns) and action-denoting words (semantic verbs) without any difference in function-marking morpho-syntax. However, despite this so-called *lexical flexibility* (Rijkhoff & Van Lier 2013), we expect restrictions on the use of ‘nominal’ features like possessive markers in (functionally defined, formally unmarked) action nominalizations and in main clauses, compared to prototypical ‘nominal’ constructions.

The aim of our paper is therefore to determine how (i) the semantic class of the head (actions versus things/persons, but also finer sub-distinctions such as kin term/body part/other object; intransitive/transitive; active/stative) and (ii) the pragmatic function of the phrase (reference versus predication) contribute to the choice of (possessive) dependent marking strategy in Oceanic languages. In addition, we assess the influence of possessive splits and clausal alignment patterns as attested in individual languages.

We show that the possibilities for possessive dependent marking indeed decrease when the relevant constructions involve less prototypically ‘nominal’ combinations of semantic head and pragmatic function. Referential phrases, whether headed by a semantic noun or a semantic verb, display both alienable and inalienable markers, whereas main predicate phrases employ alienable markers only. The use of possessive (as opposed to ‘verbal’) morphology for subject agreement in main clauses – a cross-linguistically rare phenomenon – ties in with lexical flexibility: it developed out of action nominalizations, which are relatively frequently used in languages with a weak lexical noun/verb distinction (cf. Palmer 2011). Finally, we find that the choice of possessive strategy is largely lexically specified with semantic nouns, while with semantic verbs it is more strongly determined by syntactic considerations, independently of clausal alignment patterns.

References

- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Lichtenberk, Frantisek. 2009. Attributive possession constructions in Oceanic. In William Mc.Gregor (ed.), *The expression of possession*. Berlin/New York: Mouton de Gruyter. pp. 249-292.
- Lynch John, Malcolm Ross & Terry Crowley (eds.) 2002. *The Oceanic Languages*. Richmond: Curzon Press.
- Palmer, Bill. 2011. Subject-indexing and possessive morphology in Northwest Solomonian. *Linguistics* 49:4, 685-747.
- Rijkhoff, Jan & Eva van Lier (eds.) 2013. *Flexible word classes: a typological study of underspecified parts-of speech*. Oxford: Oxford University Press.
- Siewierska, Anna. 1998. On nominal and verbal person marking. *Linguistic Typology* 2:1, 1-56.

Ergativity splits and DSM in Cabécar (Chibcha)

The crucial question in differential subject marking (DSM) is where the differences in the coding of subjects come from (Woolford 2008). The present contribution deals with an ergative language (Cabécar, Chibchan, Costa Rica, see previous related studies, Quesada 1999, Margery Peña 2003). DSM in this language relates to (a) the presence/absence of the ergative suffix and (b) the choice among two different ergative suffixes. The conditions that determine DSM are highly complex (prominence of the argument, aspect, clause type). This talk presents a detailed description of these conditions (based on competence data elicited in threefieldwork periods and observational data from a corpus of 150 narrative texts) and develops a unifying account that sheds light on the determinants of DSM in this language and its relation to prominence asymmetries as established in previous research in argument structure, see Aissen 1999 and discussion with respect to DSM in De Hoop & De Swart 2008.

We first establish that ergativity is a syntactic phenomenon in Cabécar (based on raising facts and argument dropping facts). The root of the asymmetries between ergative and non-ergative arguments lies in a strong syntactic constraint in this language, according to which the lowest argument (i.e., the object of transitives or the single argument of either unergatives or unaccusatives) is strictly left adjacent to the verb (the only possible order permutations are SOV, OVS, SV). This is the absolutive argument, which is not case-marked; ergative marking only appears with the subjects of transitive verbs (and not with subjects of unergatives as in some other languages). The following instances of DSM are observed:

- (a) ergative marking is obligatory for postverbal subjects and optional for preverbal subjects (dropped with indefinites).
- (b) ergative marking is obligatory in the imperfective/habitual/future and optional in perfective past (dropped with indefinites).
- (c) ergative marking does not occur with reflexive and reciprocal constructions.
- (d) the ergative marker is *të/te* in affirmative and *wã* in negative contexts (see Margery Peña 2003:xii).

We claim that the sources of these instances are multiple: The properties (b) and (c) reflect the fact that negation and non-perfective aspects involve changes influencing the thematic properties of transitive subjects (see Blaszcak 2008 for a similar account on Polish). The property (c) reflects the fact that the subjects of reflexives/reciprocals are absolutives in this language (independent evidence comes from word order). The property (a) reflects the asymmetry between canonical and non-canonical orders, which is expected since the latter are more likely to give rise to ambiguous interpretations than the former.

References

- Aissen, Judith 1999, Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17, 673-711.
- Blaszcak, Joanna 2008, Differential subject marking in Polish. In De Hoop & Swart (eds.), *Differential subject marking*. Dordrecht: Springer, 113-149.
- De Hoop, Helen & Peter De Swart 2008, Cross-linguistic variation in differential subject marking. In De Hoop & Swart (eds.), *Differential subject marking*. Dordrecht: Springer, 17-40.
- Dixon, RMW, 1994. *Ergativity*. Cambridge: CUP.
- Margery Peña, Enrique 2003, *Diccionario cabécar – español, español – cabécar*. San José C.R.: Editorial de la Universidad de Costa Rica.

Quesada, Diego, 1999, Ergativity in Chibchan. *Sprachtypologie und Universalienforschung STUF* 52(1), 22–51.

Woolford, Ellen 2008, Differential subject marking at argument structure, syntax and PF. In De Hoop & Swart (eds.), *Differential subject marking*. Dordrecht: Springer, 17-40.

Manner verbs: a first attempt at a lexical typology

Lexical typology has become an increasingly important subfield of typology in the last two decades (for a recent overview see Koptjevskaja-Tamm et al. 2007). In this paper, we try to contribute to subfield by this presenting some initial results of a study on the lexicalization of manner in five different semantic verbal domains: motion (*run, fly*), placement (*shovel, wipe*), perception (*glare, squint*), eating and drinking (*gorge, nibble*), and communication (*screech, groan*).

The most well-studied of these semantic domains is that of motion verbs, from which we know that so-called satellite-framed languages, such as English, German, and Dutch, commonly use manner of motion verbs such as ‘walk’ and ‘crawl’ (Talmy 1991, Slobin 1996), while so-called verb-framed languages, such as French and Spanish, do not commonly use this type of verb. Some efforts have been made to compare the semantic domain of motion to that of perception (Slobin 2008), although the emphasis has not yet been on verbs.

In this paper, we present some preliminary results on the lexicalization of manner in these different semantic domains for English, Dutch, and French. The data come from a parallel corpus of *Harry Potter and the Philosophers Stone* and *Harry Potter and the Chamber of Secrets* (by J. K. Rowling). We focus on the following questions:

1. Do English and Dutch, which commonly use manner verbs to encode motion, commonly use manner verbs in semantic domains other than motion?
2. Does French, which does not commonly use manner verbs to encode motion, show the same tendency in the other semantic domains?
3. Are there any differences between the various semantic domains in terms of size, i.e. do some domains feature more manner verbs than others?
4. Are there any differences between the various semantic domains in terms of lexical structure, i.e. are there some concepts that are always lexicalized while others are only lexicalized in individual languages?

While the emphasis of the paper is on an across-domain comparison of manner verbs in these three different languages, efforts are made to address issues which will be relevant for the study of manner lexicalization on a much wider scale. Our ultimate goal is to study manner lexicalization in more languages, using specially designed questionnaires.

References:

- Koptjevskaja-Tamm, Maria, Martine Vanhove & Peter Koch (2007). Typological approaches to lexical semantics. *Linguistic Typology* 11: 159-185.
- Slobin, Dan I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In Masayoshi Shibitani & Sandra A. Thompson (eds.), *Grammatical constructions: Their form and meaning*. Oxford: Clarendon Press.
- Slobin, Dan I. (2008). Relations between paths of motion and paths of vision: A crosslinguistic and developmental exploration. In V. M. Gathercole (ed.), *Routes to language: Studies in honor of Melissa Bowerman*, 197-221. Mahwah, NJ: Lawrence Erlbaum Associates.
- Talmy, Leonard (1991). Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 480-519. Berkeley: University of California.

Lexicalized negative verbs: a cross-linguistic study

The goal of this project is to identify which negated concepts, connected mainly to states and events, are expressed lexically across languages, cf. (i) English *dunno* < I don't know or (ii) Tundra Nenets *jexerasj* 'not know'. Both (i) and (ii) can be semantically decomposed into a negative component and a positive sense. Following (Brinton and Traugott 2005, Moreno-Cabrera 1998), such forms are considered instances of lexicalization. The term is used here in a synchronic sense. Lexicalizations of negation are mentioned in numerous works, some examples include Jespersen (1917: 13, in passim), Croft (1991), Payne (1985), van Gelderen (2008). De Haan (1997), Palmer (1995), van der Auwera (2001) cover lexicalizations of modal senses such as „not be able to“, „need not“, etc. Eriksen (2011) discusses negation strategies in non-verbal sentences which also include lexicalized expressions meaning „not be“. However, a systematic cross-linguistic survey of lexicalized negative senses is missing both in the literature on negation as well as in work on lexical typology, cf. (Evans 2010, Goddard 2001, Koch 2001, Koptjevskaja-Tamm 2008). Zeshan (2004), on the other hand, offers a detailed discussion of irregular negative senses in sign languages. These expressions can be also viewed as lexicalizations and are compared with the results of the current study.

The data used here come from a sample of 97 geographically and genealogically distinct languages and also from three family-based samples which cover Slavic, Uralic and Polynesian.

Formally, lexicalizations of negation appear to be of two kinds. The first one is illustrated by (i) above; it represents a fused form where a phrase containing the negator *don't* has fused with the verb form *know*. The grammatical morpheme in the fused form is frequently, though not always, a negator. In the second kind, cf. (ii), the form *jexerasj* cannot be further segmented into separate morphemes and shows no formal relation to the negative morpheme in Tundra Nenets or the expression for the corresponding positive sense, cf. (1) below. In the current data set, lexicalizations of the second kind prevail. It should also be noted that these words are sometimes the only negators of the corresponding positive senses but there are also cases where they co-exist with a regularly negated positive verb cf. (2) below. There are also cases where they simply lack an affirmative correspondent. My database contains lexical expressions for 65 negative senses which can be grouped into broader semantic domains. There are also a handful of senses which are regularly lexicalized cf. (3) below. Perhaps not surprisingly, all of the semantic and grammatical domains identified by Zeshan (2004: 50) as being coded by irregular negatives in sign languages are also observed as lexicalizations in spoken languages. These domains are: COGNITION (not know, not understand), EMOTIONAL ATTITUDE (not want, not like), MODALS (cannot, need not), POSSESSION/EXISTENTIAL (not have, not exist), TENSE/ASPECT (did not, not finished), EVALUATIVE JUDGEMENT (not right, not enough). One domain that tends to be lexicalized in spoken languages but is not reported by Zeshan for sign languages is labeled here NON-UTTERANCE; it is represented by senses such as „not talk“, „not tell/inform“, „not mention“, cf. (4) below for an example.

Lexicalized expressions for „not exist“ are so common, that it is easier to identify areas where they do not occur. In the current sample such areas are Western Europe, South East Asia and southern and central parts of South America. Lexicalized expressions of the remaining senses appear to occur less commonly in Africa than in the other macro-areas. As regards the micro-samples, Slavic and Polynesian languages show a preference for lexicalization of „not want“ while „not know“ is commonly lexicalized in Uralic. Further on, within each family, these lexicalizations can be correlated with smaller genealogical and areal clusters.

The cross-linguistic data collected here do not allow for the postulation of an implicational hierarchy; that is, it is currently not possible to predict the order of lexicalization of negative senses. However, it is clear that negative lexicalizations are organized around a limited number of cognitively salient categories. As Zeshan (2004: 51) points out “events and states such as not liking, not knowing,

not having are all identifiable human experiences". This is why these concepts are often expressed by lexicalized expressions cross-linguistically regardless of language medium.

Examples

(1) Tundra Nenets (Uralic, Samoyed), (Wagner-Nagy 2011: 129-131)

- a. jexasj „not know"
- b. ténewasj „know"
- c. n'i- negative auxiliary for standard negation

(2) Central Alaskan Yup'ik (Eskimo- Aleut) (Jacobson and Jacobson 1995: 26)

- a. ner-yu-nrit-ua
eat-want.to.V-NEG-1SG
„I don't want to eat"
- b. ner-yu-llru-nrit-ua
eat-want.to.V-PST-NEG-1SG
„I didn't want to eat"
- c. ner-yuumiit-llru-unga
eat-not.want.to.V-PST-1SG
„I didn't want to eat"

(3) Most frequently lexicalized negative senses

EXAMPLE	SENSE	Number of languages with this lexicalization
Turkish <i>yok</i>	not exist	63
Ecuadorian Siona <i>wehsə</i>	not know	32
Lezgian <i>ruš</i>	not be of identity	30
Ojibwe <i>gwiinawi-</i>	cannot, be unable	29
Kwaza <i>heʔai</i>	not want	21
Anejom̐ <i>emwin</i>	not talk	10
Tetum <i>lalika</i>	need not	8

(4) Mele-Fila (Austronesian, Malayo-Polynesian, Central-Eastern Oceanic, [...]Polynesian, Nuclear, Samoic-Outlier, Futunic) (Clark 2002: 692)

In Mele-Fila SN is expressed discontinuously, by means of obligatory postverbal particle *kee* and an optional prefix *s(e)-*

- a. *au (s)-taae-a kee*
1SG NEG-know-TRANSITIVIZER NEG
'I don't know'
- b. *kaijpunu* 'not speak, say nothing'
- c. *vanaga* 'speak'

References

- Brinton, Laurel J., and Traugott, Elizabeth Closs. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Clark, Ross. 2002. Ifira-Mele. In *The Oceanic Languages*, eds. John Lynch, Malcolm Ross and Terry Crowley, 681-693. Richmond, Surrey: Curzon.
- Croft, William. 1991. The Evolution of Negation. *Journal of Linguistics* 27:1-39.
- de Haan, Ferdinand. 1997. *The Interaction of Negation and Modality: A Typological Study*. Outstanding Dissertations in Linguistics Series: Garland Publishers.
- Eriksen, Pål Kristian. 2011. 'To Not Be' or not 'to not be': The typology of negation of non-verbal predicates. *Studies in Language* 35:275-310.
- Evans, Nicholas D. 2010. Semantic Typology. In *The Oxford Handbook of Linguistic Typology*, ed. Jae-Jung Song, 504-533. Oxford: Oxford University Press.
- Goddard, Cliff. 2001. Lexico-semantic universals: A critical overview. *Linguistic Typology* 5:1-65.
- Jacobson, Steven A., and Jacobson, Anna W. 1995. *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language*. Fairbanks: Alaska Native Language Center and Program, University of Alaska Fairbanks.
- Jespersen, Otto. 1917. *Negation in English and other languages*. København: Hovedkommissionær: Andr, Fred, Høst & Søn, KGL. Hof-boghandel, Bianco Lunos Bogtrykkeri.
- Koch, Peter. 2001. Lexical typology. In *Language Typology and Language Universals*, eds. Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible, 1142-1179. Berlin/New York: Walter de Gruyter.
- Koptjevskaja-Tamm, Maria. 2008. Approaching lexical typology. In *From Polysemy to Semantic Change*, ed. Martine Vanhove, 3-52. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Moreno-Cabrera, Juan C. 1998. On the relationships between grammaticalization and lexicalization. In *The Limits of Grammaticalization*, eds. Anna Giacalone Ramat and Paul Hopper, 211-227. Amsterdam: John Benjamins.
- Palmer, F. R. 1995. Negation and the Modals of Possibility and Necessity. In *Modality in Grammar and Discourse*, eds. J. Bybee and S. Fleischman, 453-471. Amsterdam: John Benjamins Publishing Company.
- Payne, John R. 1985. Negation. In *Language Typology and Syntactic Description, vol I: Clause Structure*, ed. Timothy Shopen, 197-242. Cambridge: Cambridge University Press.
- Wagner-Nagy, Beáta. 2011. *On the typology of negation in Ob-Ugric and Samoyedic Languages*. Helsinki: Suomalais-Ugrilainen Seura -- Société Finno-Ougrienne -- Finno-Ugrian Society.
- van der Auwera, Johan. 2001. On the typology of negative modals. In *Perspectives on Negation and Polarity Items*, eds. Jack Hoeksema, Hotze Rullmann, Victor Sanchez-Valencia and Ton van der Wouden, 23-48. Amsterdam: John Benjamins Publishing Company.
- van Gelderen, Elly. 2008. Negative Cycles. *Linguistic Typology* 12:195-243.
- Zeshan, Ulrike. 2004. Hand, head, and face: Negative constructions in sign languages. *Linguistic Typology* 8:1-58.

On the universality of reflexive strategies

It is a long-standing observation that languages employ special strategies to express reflexivity (see, for instance, Jespersen 1933). And despite occasional prima facie exceptions, there appears to be a common view among language typologists that the use of special strategies is the ‘norm’ (Schladt 2000, Moyse-Faurie 2008, Heine & Miyashita 2008). That is, as a rule, one doesn’t find the configuration in (1) with a reflexive interpretation.

(1) *Subject verb pronominal

There is, however, an influential view that, nevertheless, the employment of special strategies is just a tendency, reflecting pragmatic preferences, where some languages simply have not yet developed the tools to express these preferences (Levinson 2000, Evans and Levinson 2009). The alternative is to derive the need for special marking of reflexivity from fundamental properties of the computational system (Reuland 2011). The latter type of approach raises a question of whether the employment of special strategies to express reflexivity is truly universal. To resolve this issue, it is important to focus attention on those languages that – at least prima facie – don’t have dedicated reflexives, and determine whether they actually do, or don’t employ special strategies to express reflexivity.

One such language is Khanty (Nikolaeva 1995, 1999). According to Nikolaeva (1995), Khanty has no dedicated reflexive pronouns; instead, personal pronouns are used.

- (2) a. UtliteXoi łuveŋi/k išʔʔk-s-ʔʔɦe.
 teacher he.ACC praise-PST-SG.3SG
The teacher praised him(self).
- b. NemXojati łuveŋi/k änt išʔʔk-s-ʔʔɦe.
 no.one he.ACC NEG praise-PST-SG.3SG
No one praised him(self).

łuveŋ in object position can be bound by a co-argument subject. It can also receive a value from discourse, showing that *łuveŋ* is a true pronominal (2a). (2b) with a quantificational antecedent shows that the local dependency is one of binding, not coreference.

The question is, then, how Khanty uses its pronouns to express reflexivity, just by ‘brute force’ binding (which could support the ‘tendency’ view), or does it have structural properties that independently license reflexivity? In this talk we review data collected on a field trip in July 2012, and show that these support the latter option.

Khanty has two types of verbal agreement: obligatory subject agreement and optional object agreement (OAg), as illustrated in (3).

- (3) UtliteXo poXlen’ki išʔʔk-s-ʔʔɦe / išʔʔk-s.
 teacher boy praise-PST-SG.3SG / praise-PST.3SG
The teacher praised the boy.

The following condition applies: a personal pronoun can be locally bound – yielding a reflexive predicate – only if the verb carries object agreement, cf. the ill-formedness of (4).

- (4) *UtliteXoi łuveli išḡk-s.
 teacher he.ACC praise-PST.3SG
*The teacher praised him / *himself.*

The presence of object agreement facilitates object drop, as in (5).

- (5) TāmXätł ma c'ăta van-s-em.
 today I there see-PST-SG.1SG
 {LC: Yesterday my son went to Beryozovo.} *Today I saw (him / *myself) there.*

But a zero object is incompatible with local binding. The predicate in (5) cannot be interpreted as reflexive. In order to avoid the configuration in (1), the object argument should be complex. It is, since OAgr licenses a null object. Overt *łuveł* forms a constituent with the null object. This analysis is further supported since *łuveł* is also used as an intensifier (note that in this capacity it cannot be null).

These facts provide an argument against the 'tendency' approach to reflexivity and add new data to the typology of reflexive strategies.

Wanting universals

This paper is a corpus-typological contribution to the dispute whether **'want'** (also called **desiderative**) is a cross-linguistic universal (Khanina 2008 vs. Goddard & Wierzbicka 2010). Unlike Khanina (2008) and Goddard & Wierzbicka (2010) but like Haspelmath (2005) I assume that 'want' is always expressed by a **construction**. Rather than discovering structural universals of 'want', the aim is to identify how 'want' constructions can be identified given comparable texts and a functional domain. Hereby the task is to find out by which combination out of the large number of all marker candidates in a corpus (both words and morphemes) a meaning is encoded (if it is encoded at all). It is argued that there is a uniform solution to this task: a **procedural universal**.

'Want' constructions are automatically extracted from **parallel texts** (NT) in a world-wide sample of 100 languages texts using a universal extraction algorithm. The sample is biased on purpose toward Eurasia and New Guinea (an area distant from Europe) in order to verify to what extent a SAE-based extensional definition of the domain 'want' affects the quality of extraction. All extracted constructions are manually evaluated with dictionaries and reference grammars. The approach differs thus from most approaches to **quantitative typology** in that quantitative methods are applied at the very beginning before conventional qualitative methods are used.

The automatically extracted constructions are highly similar to Tomasello's (2003) **itembased constructions** in language acquisition. Like item-based constructions they are **constructional islands** (independent of other domains and constructions). Finding an itembased construction requires **local semantic decomposition** ('want' meaning as opposed to all other meanings) rather than the global semantic decomposition of Natural Semantic Metalanguage. Goddard & Wierzbicka (2010) claim that a semantically primitive meaning such as 'want' will always be expounded by means of a **segmental sign**. This paper largely confirms the claim but associates it with the **local cue validity** of such markers (Tomasello 2003: 136 based on work by Dan Slobin).

It is shown that 'want' constructions with considerable **paradigmatic and syntagmatic complexity** can be extracted without language-specific expert knowledge (expert knowledge is a precondition for Goddard & Wierzbicka's 2010 polysemy analysis) with a universal algorithm even though desiderative markers exhibit a wide range of **polysemy patterns** (Khanina 2008). A major finding is that there is more than one dimension of variability of construction types (Haspelmath's [2005] typology being one of them) and that lexical polysemy patterns (Khanina 2008) do not determine construction types. The study also confirms previous findings that **egocentricity** is an important ingredient of 'want'.

References

- Goddard, Cliff & Wierzbicka, Anna. 2010. 'Want' is a lexical and conceptual universal. *Studies in Language* 34(1):108–123
- Haspelmath, Martin. 2005. 'Want' complement subjects. In Haspelmath, Martin & Dryer, Matthew & Gil, David & Comrie, Bernard (eds.) 2005. *The World Atlas of Language Structures* 124. Oxford: OUP.
- Khanina, Olesya. 2008. How universal is 'wanting'? *Studies in Language* 32(4):818–865.
- Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Why is ergativity often restricted to certain environments? A look at the diachrony of Differential A Marking

The present paper aims at explaining why ergative case alignment often only occurs in a part of the grammar together with other instances of Differential A Marking (DAM), which here refers to a variation in the case-marking of the more agent-like argument of two- and three-argument constructions. (I follow Bickel's 2011 approach to grammatical relations.) DAM can be conditioned by the following factors:

- Referential properties of the A argument, such as person, number, animacy
- Predicate classes (or valence classes having different case frames)
- Clause properties: TAM values, polarity, clause type (main vs. non-main clauses), scenario (nature of co-arguments)
- Semantic and pragmatic function (including esp. agentivity and focus)

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

Languages having different kinds of A splits are generally also more likely to have an ergative case alignment pattern in some parts of the grammar, when some A arguments are morphologically marked, while S (single arguments of one-argument constructions) and P arguments (the more patient-like arguments of two-argument constructions) are unmarked, as are other A arguments. The present study shows that split ergativity is often the result of the diachronic emergence of DAM, which involves, among others, the following contexts and mechanisms:

- Reanalysis of complex NPs (or nominalized clauses) as full (main) clauses
- Emergence or extension of predicate classes
- Reanalysis of adjuncts as core arguments
- Reanalysis of focus markers as A markers

On the other hand, since DAM is often a precondition for the emergence of ergative patterns, languages lacking DAM usually also lack ergative patterns.

Some methodological prerequisites should also be mentioned here:

- Case marking is defined here in fairly broad terms, including any element of dependent marking on the clause level irrespective of their morphological nature (affixes, clitics and separate words), since the properties (and definitions) of words vary widely across languages (see Dixon and Aikhenvald 2002).
- Arguments (and valence) are defined in purely semantic terms (following the approach by Bickel 2011) since the application of syntactic criteria of argumenthood poses problems for the crosslinguistic investigation of arguments (cf. Witzlack-Makarevich 2011: 41-47).

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

References:

- Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 399-444.
- Dixon, R. M. W. and Alexandra Y. Aikhenvald. 2002. Word: a typological framework. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Word: a Cross-Linguistic Typology*. Cambridge: Cambridge University Press, 1-41.
- Witzlack-Makarevich, Alena. 2011. *Typological Variation in Grammatical Relations*. Ph.D thesis, Universität Leipzig.

Theme session: Quantitative Linguistic Typology: State-of-the-Art and Beyond

Simulating language, family, and feature evolution: a review of the state of the art

In recent years simulation models have become current as a complement to empirical inquiries into language dynamics (understood as a broad field encompassing not only language change—the domain of traditional historical linguistics—but also "external factors" such as interaction among languages and their speakers). Simulation models have been designed to test hypotheses concerning the development of linguistic diversity (Nettle 1999, Holman et al. 2008), language competition (e.g., Abrams and Strogatz 2003, Patriarca and Leppänen 2004, Mira and Paredes, Stauffer et al. 2006, Schulze et al. 2008, Murilo et al. 2008), the measurement of differential stabilities of abstract typological features (Wichmann and Holman 2009), the relationship between language change and population structures (Nettle 1999b, Wichmann et al. 2008), and the performance of different phylogenetic methods (Barbançon et al. 2013). (In addition, numerous studies have applied computer simulations as an approach to language evolution, but this field is not considered here).

Among agent-based models, which are suitable for studying the interaction among languages and their speakers, some operate with a simulated space (a lattice) and no internal language structure (de Oliveira et al. 2006a,b, Patriarca & Leppänen 2004, Pinasco and Romanelli 2006), whereas others similarly study language dynamics in a simulated space but also has some simple way of representing language structure—typically as a string of binary features (bitstrings) (Schulze and Stauffer 2005, Kosmidis et al. 2005, Stauffer et al. 2006, Murilo et al. 2008, Teşileanu and Meyer-Ortmanns 2006). The latter class of models is obviously of most immediate interest from a linguistic point of view.

In the present paper we review existing literature on simulation models suitable for the investigation of topics relevant for historical-typological linguistics in a broad sense, propose evaluation criteria, and set forward a model representing a consensus of the group of co-authors. With the proliferation of work in this area it is important to distinguish between less and more adequate models. Since the simulation of the dynamics of linguistic diversity is a field of great potential we feel that it is important to assess previous results as well as to outline a generic simulation model that strikes a balance between the two opposed desiderata of maximal realism and minimal complexity. Our currently evolving general-purpose simulation model is intended primarily as a tool for evaluating specific methods for estimating diachronic typological developments, not as a tool for modeling everything we know about language history. It is simple, containing in its basic structure mainly parameters for birth, death, and change, but will allow for extensions via plug-ins (e.g. different kinds of networks for modeling causes or effects of change). Languages are modeled as vectors of properties (values of variables) as its basic units (the units whose behavior is simulated). A distinction is made between cognation variables (extremely low probability of recurrence, but defining genealogies) vs. typological variables. Finally, it is tuned to fit reality to the extent that it replicates known global patterns of languages such as family sizes, language sizes, and quantitative features of tree topologies.

Transitivity of Resultative Verbs and Word Order Typology

Resultative Serial Verb Constructions (RSVCs) can be found in many languages. RSVCs generally consist of an action verb (V1) and a resultative verb (V2). One key syntactic characteristic is the sharing of internal argument by the two verbs (Collins, 1997). This paper focuses on the relationship between the transitivity of resultative verbs(V2) and word order typology.

As shown in (1), both Mandarin Chinese and Yoruba are VO languages where V1 is followed by V2. (1a) differentiate itself from (1b) by the position of internal arguments. Both V2s are unaccusative but V2 in (1a) is argued to be shelled by a causative *vP*(R. Sybesma, 1999; R Sybesma & Shen, 2006). Comparatively, V2 in the Jingpo language of (2a) is affixed by a causative morpheme *ja-*. The dropping of the morpheme leads to ungrammaticality (2b). Besides of Jingpo, other OV languages such as Korean also select a transitive verb as their V2 (Lee, 1996).

- (1) a. Zhangsan tui dao le Lisi.
Zhangsan push(V1)fall(V2) ASP Lisi
Zhangsan pushed Lisi down. (Mandarin Chinese)
- b. Femi ti Akin subu.
Femi push(V1) Akin fall
Femi pushed Akin down. (Yoruba, Lord, 1974)
- (2) a. Palong hkrut ja-hpro kau sai.
clothes wash(V1) CAUSE-be.white(V2) AUX 3SG(Subj)PERF
He made the clothes white by washing.
- b. *Palong hkrut hpro kau sai.
clothes wash(V1) be.white(V2) AUX 3SG(Subj)PERF
He made the clothes white by washing. (Jingpo, Peng & Gu, 2006)

In this paper, we argue for Principles of Resultative Verbs (PRV) shown in (3). (3a) shows that the syntactic order of V1 and V2 in VO/OV languages remain the same. (3b) rules out the intransitivity of resultative verbs in OV languages. We further argue that V2 being transitive is a remedy to keep syntactic derivation away from crashing.

(3) Transitivity Constraint of Resultative Verbs (TCRV)

- a. **Iconicity Condition:** the resultative verb (V2) always follows the action verb (V1) in VO and OV languages;
- b. **Transitivity Constraint:** The resultative verbs can be unaccusative or transitive in VO languages, but only transitive in OV languages.

References

- Collins, C. (1997). Argument sharing in serial verb constructions. *Linguistic Inquiry*, 461-497.
- Lee, S. (1996). Resultative Serial Verbs: The Interaction between Event Structure and Headedness. *Korean Journal of Linguistics*, 23(3), 873-888.
- Lord, C. (1974). Causative constructions in Yoruba. *Studies in African Linguistics*, 195-204.
- Peng, G., & Gu, Y. (2006). *Research on 'V1+V2causative' SVC in Jingpo*. Unpublished manuscript.
- Sybesma, R. (1999). *The Mandarin VP*. Dordrecht, [Netherlands] ; Boston: Kluwer Academic Publishers.
- Sybesma, R., & Shen, Y. (2006). Small Clause Results and the Internal Structure of the Small Clause. *Journal of Huazhong University of Science and Technology (Social Science Edition)*(4), 40-46.

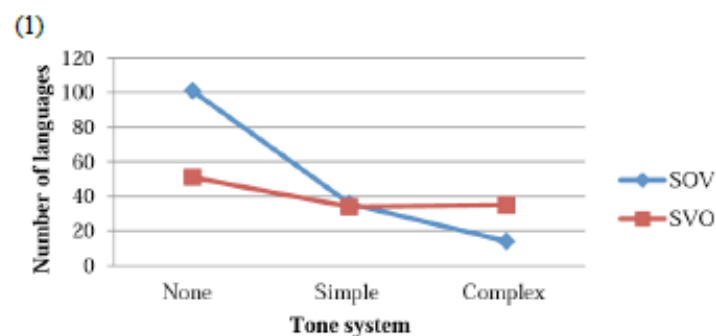
Correlations between Tonality and Word Order Type

Exploring the prosodic typology of language, Gil (1986) argues for extending the typology for metered verse to ordinary language based on 170 languages. Among the results observed is an indirect correlation between word order type and the presence of lexical tone: iambic languages tend to be VO and tonal, while trochaic languages tend to be OV and non-tonal. Gil's hypothesis that the most basic distinction is between iambic and trochaic feet, however, cannot be tested using the World Atlas of Language Structures online (WALS) due to insufficient data; many languages with complex tone systems are arguably iambic (Thai, Chaozhou) or cannot be categorised as either iambic or trochaic (Cantonese). More explanatory factors are thus needed.

This paper reexamines the correlation of tonality and word order typology, with evidence from a larger and more updated database which provides relevant data from 527 languages (WALS, Maddieson 2011). It is shown that:

- a. Overall, there is a significant relationship between word order and lexical tone: 57% of SVO languages are tonal, vs. 33% of SOV languages;
- b. Among tonal languages, 51% of SVO languages have complex tone systems (contrasting more than two tones), compared with 28% of SOV languages.

These differences are significant based on χ^2 tests. Figure (1) shows that SOV languages are around twice as likely to be non-tonal as tonal, whereas SVO languages are around twice more likely to be tonal. SVO languages are twice as likely to have complex tonal systems as their SOV counterparts.



Tonal complexity appears to be related to word order type via morphological typology: SVO order favours isolating morphology whereas SOV favours agglutinative morphology. This distribution is illustrated within the Sino-Tibetan family, where genetic factors may be assumed to be held constant. Among Tibeto-Burman SOV languages, the most complex tone systems occur in isolating languages such as Lahu, while the most morphologically complex languages such as Limbu are non-tonal. In the Sinitic branch, the geographical distribution displays a continuum with northern areas having fewer tones and more SOV structures, and southern varieties more complex tone systems and more SVO structures (Hashimoto 1976). These relationships will be illustrated with examples from Sinitic languages.

References

- Gil, David. 1986. A Prosodic Typology of Language. *Folia Linguistica* 20, 165-231.
- Hashimoto, Mantaro. 1976. Language Diffusion on the Asian Continent: Problems of Typological Diversity in Sino-Tibetan. *Computational Analyses of Asian and African Languages* 3, Tokyo, 49-66.
- Maddieson, Ian. 2011. Tone. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 13. Available online at <http://wals.info/chapter/13>. Accessed on 2013-01-15.

The category of addressee's perspective in Kashibo-Kakataibo

Kashibo-Kakataibo is a Panoan language spoken in Peru by around 3,500 people. It exhibits a predominantly postpositional morphology; a split ergative case marking system; a relatively free and pragmatically-oriented constituent order (but with a tendency toward verb-final sentences); a rich system of switch-reference used in clause-chaining; and a pervasive use of nominalizations for several functions such as relativization and complementation. In this talk, I would like to present a fascinating grammatical category that plays an important role in Kakataibo grammar: the morphological marking of the presumptions and expectations of the speaker about the hearer's access to the information being presented in an utterance. This grammatical category has been called *addressee's perspective* in Zariquiey (2011: 428).

The category of hearer's perspective in Kashibo-Kakataibo revolves around the functions of at least three morphemes that belong to three different paradigms. (i) The verbal suffix *-ín*, which indicates that, in the speaker's conceptualization of the speech act, the propositional content of the utterance is accessible to the hearer. (ii) The second position enclitic *=pa*, which is used to explicitly indicate that the speaker assumes that the information he is presenting is not accessible to the hearer. This enclitic only appears in combination with the enclitic *=ri*, which indicates that the event expressed by a sentence is strongly contextual (i.e. the event is happening close to the speech act participants; was previously introduced by a question; or is the topic of the conversation). Therefore, *=pa* is only used in very specific situations where, for any reason, the information is assumed to be available to the speaker but not to the hearer. Finally, (iii) the verbal marker *-ie;*, which is used in what Zariquiey (2011: 451) call *accusatory constructions*, to indicate that the hearer is far from the event (and, therefore, does not have perceptual access to it). In this talk, I will describe the functions of these three morphological markers, which establish interesting interactions with other grammatical categories, particularly with the distinction between narrative and conversational registers (Zariquiey 2011: 480-527). Besides those interactions, the category of hearer's perspective seems to be primarily focused on the accessibility of the information to this specific speech act participant and, therefore, in typological terms, it could be understood as a complement to evidentiality, which has to do with the speaker's access to the information and has been argued to be analyzable as addressee-oriented proposition deixis by De Haan (2005). Deixis seems to be a primary aspect of the category of hearer's perspective in KK and, therefore, following De Haan, it may be seen as a case of addressee-oriented proposition deixis (Anderson and Keenan 1985).

References

- De Haan, F. 2005. "Encoding speaker's perspective: Evidentials". In: Frayzinger, Zygmunt, D. R., Hodges, A. (Eds.), *Linguistic Diversity and Language Theories*. John Benjamins Publishing Company, pp. 379-397.
- Anderson, S. and E. Keenan. 1985. "Deixis". In: Shopen, Timothy. *Language typology and syntactic description*. Volumen III. Cambridge: Cambridge University Press, 259-308.
- Zariquiey, Roberto. 2011. A grammar of Kashibo-Kakataibo. PhD dissertation. Melbourne, Australia: La Trobe University.

Cross-modal typology

Cross-modal typology focuses on the relationship between typological variation and the effects of language modality when comparing signed languages (SL) and spoken languages (SpL).

Sign Language Typology (e.g. Zeshan 2006, Zeshan & Perniss 2008) has demonstrated that cross-linguistic variation is comparable in SL and SpL. The observations here use two of the largest SL data collections:

- Negatives, with 38 SL (Zeshan 2004)
- Cardinal numerals, with 29 SL (Zeshan & Sagara, in prep.)

Information from SpL is based on secondary sources from typological literature e.g. Dryer (1988, 2011) for negation, Barriga Puente (1988) and Comrie (2011) for numerals.

Figures 1 and 2 represent structural aspects of numerals and negation. The intersection of the circles is a typological space occupied by both SL and SpL, while the extreme right and left are unique to each modality. Where a label spans across lines, this indicates occurrence predominantly, but not exclusively, in one modality. Thus in Figure 1, digital numerals are common across SL, but marginal in SpL. As the representation focuses on modality differences, absolute frequencies are disregarded. In Figure 1, "Spatial" and "Iconic" occupy the same section of the diagram as they are both unique to SL, but using iconicity is universal in the SL sample, while a spatial numeral construction occurs only once. The results show two kinds of modality effects:

Absolute modality effects:

Here a feature occurs in one of the modalities only and is not found in the other modality. This may be due to physical articulatory characteristics, for example spatial morphology in SL numerals, or cognitive-perceptive aspects of the linguistic signal, such as ready availability of iconicity in SL numerals and negatives. Sometimes there is no straightforward explanation. For instance, conjunctions ("and", "with") are unattested in SL numerals, although there is nothing in the modality itself that would prevent this. Conversely, in SpL, a digital numeral strategy derived from writing (like saying "one zero zero" for 100) is marginal and never used as the only strategy, though nothing in the spoken modality prevents its use.

Relative modality effects:

This concerns features found in both SL and SpL, but with a very different distribution. An example is the use of non-manual suprasegmentals in SL negation, most commonly a headshake co-occurring with manual signs. The SpL equivalent, intonational features marking negation, is attested, but very rare. Morphological negation is common in SpL, but restricted in SL. Negative affixes in SL do not apply to an entire word class, and there are fewer options, as only suffixes and enclitics, but not prefixes and proclitics, are attested in SL.

In some cases, typological variation seems to have nothing to do with language modality. For instance, both SL and SpL use addition and multiplication in cardinal numerals, while subtraction is rare in both modalities. There is, presumably, no modality-specific pressure towards the type of arithmetic operation.

Each cross-modal typological space exhibits complex patterns of intra-modal and inter-modal variation, and cross-modal typology is needed for further grammatical domains in the future.

Figure 1: Cross-modal typological space: Numerals

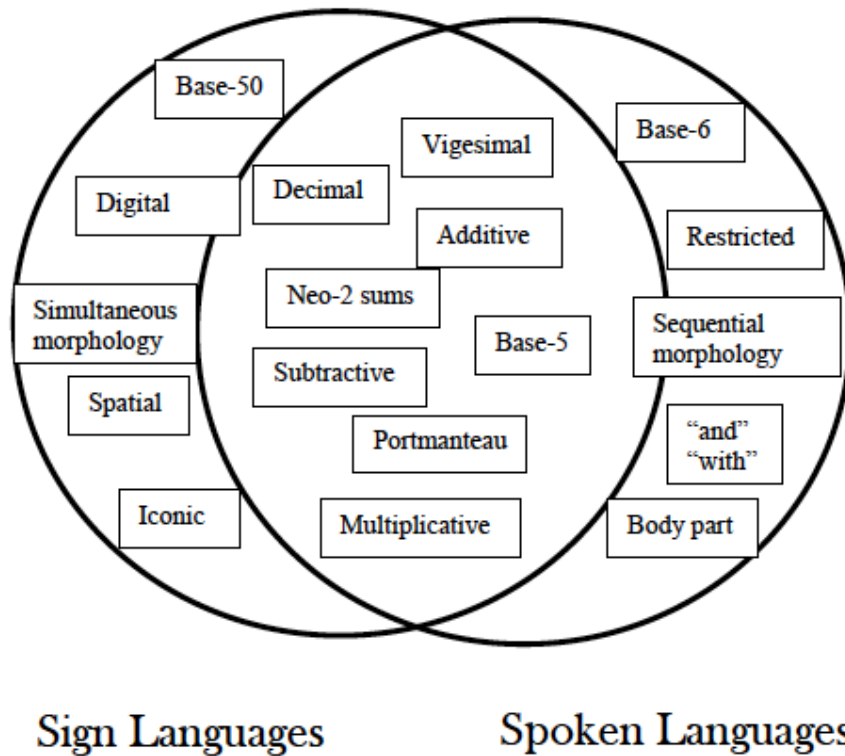
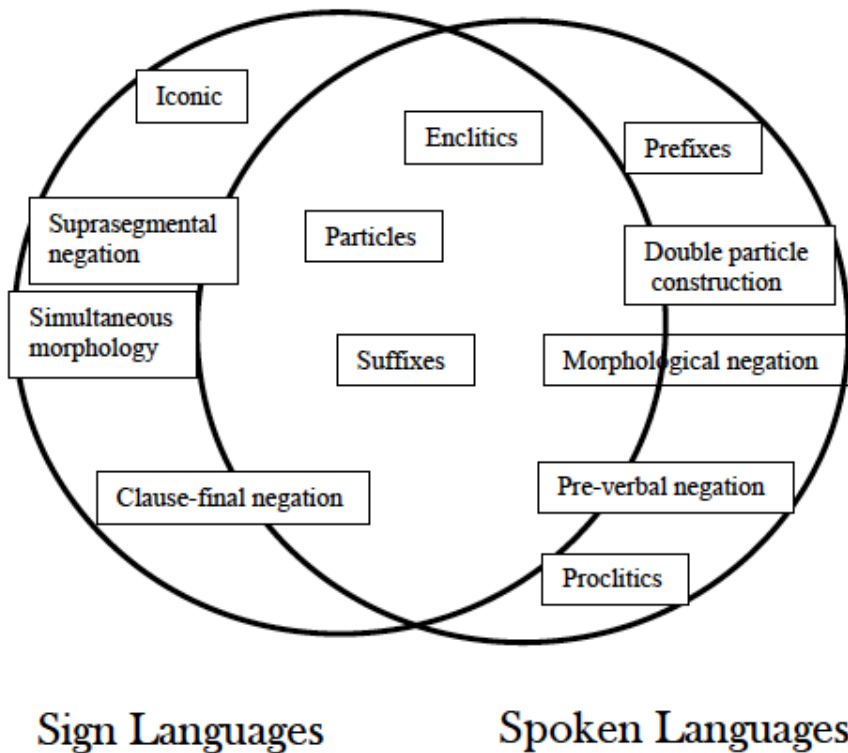


Figure 2: Cross-modal typological space: Negation



References

- Barriga Puente, Francisco (1998). *Los sistemas de numeración indoamericanos: un enfoque areotipológico*. Mexico: Universidad Nacional Autónoma de Mexico.
- Comrie, Bernard (2011): Numeral bases. In: Dryer, Matthew S. & Haspelmath, Martin (eds.): *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 131. Available online at <http://wals.info/chapter/131>. Accessed on 2013-15-01
- Dryer, Matthew (1988). Universals of negative position. In Michael Hammond, Edith Moravcsik, & Jessica Wirth (eds.), *Studies in Syntactic Typology*, 93–124. Amsterdam: Benjamins.
- Dryer, Matthew (2011): Order of negative morpheme and verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.): *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 143. Available online at <http://wals.info/chapter/143>. Accessed on 2013-15-01
- Zeshan, Ulrike (2004). Hand, head and face: negative constructions in sign languages. *Linguistic Typology* 8(1): 1-58.
- Zeshan, Ulrike, ed. (2006). *Interrogative and Negative Constructions in Sign Languages*. Sign Language Typology Series No. 1. Nijmegen: Ishara Press.
- Zeshan, Ulrike & Perniss, Pamela, Eds. (2008) *Possessive and existential constructions in sign languages*. Nijmegen: Ishara Press.
- Zeshan, Ulrike & Sagara, Keiko, Eds. (in prep.) *Semantic domains in sign languages: Number, colour and kinship*.