

Association for Linguistic Typology
6th Biennial Meeting

Hotel Bumi Minang
Padang, West Sumatra, Indonesia
21-25 July 2005

Abstracts

Abstracts are arranged alphabetically by author, following the appropriate system of alphabetization for each author's name. For example, Johan van der Auwera is under A, Bambang Kaswanti Purwo under B, Bernard Comrie under C, and so forth.

For the Workshops on the Languages of Sumatra and on Ditransitive Constructions, the abstracts can be found under the individual authors names. The Workshop on Ditransitive Constructions also has a general abstract under S(iewierska and Hollmann). For the Workshop on Sign Language Tyypology the abstracts are under Z(eshan). For the Workshop on the World Atlas of Language Structures there are no abstracts, except by Dahl.

**Vanishing Voices of the Andaman Islands:
The Case of Intriguing Sounds of
Great Andamanese, Jarawa and Onge
Anvita Abbi**

Various studies in the past, both linguistic and genetic, suggest that Andaman Islanders might be the last representatives of pre-Neolithic Southeast Asia (Hagelberg et al 2002), and (Endicott et al 2003). Anthropologists in the past have suggested a common origin to the people of the Andaman Islands, the Malay Peninsula, the Indonesia and the Philippines (Martin 1905). Such views have not gone unchallenged, however; Andaman Islanders still constitute a major area of research. Andamanese had little sustained contact with the outside world until the mid –19th century. Their languages, likewise, are linguistic isolates with Great Andamanese, a conglomeration of ten disparate but mutually intelligible dialects once spoken in the entire set of islands, presenting unique structures in morphological features and sound systems. It is, at present, not spoken by more than 25-28 people out of the total population of 37. The current paper is about the sound structures of three major surviving languages of the Island, such as Great Andamanese, Onge (94 in all), and Jarawa (<250). The fourth major language of the island, Sentinalese (approximately 150), is still inaccessible to outsiders.

Present sound system of Great Andamanese offers a large variety of sounds as no two people speak the same variety, and inter community and intra-community variation is substantial. Sounds such as voiced and voiceless bilabial fricatives, retroflex laterals, labialised laterals are some of the sounds that few speakers of Great Andamanese share with other members of the same community, nor do they share these sounds with the other two distinct linguistic groups of the area. Conversely, Onge and Jarawa do not share some of the vowel sounds with Great Andamanese. However, some similarities in the inventory and quality of diphthongs in the three languages suggest linguistic affinities of these languages with the extant languages of the South East Asia. Though history confirms the insular existence of the Andaman Islanders over the past 2 millennia (Cooper 1988), yet languages of the Islands are characterized by two significant Indian Areal phenomena, i.e., of retroflex and aspiration. All in all, the typology of the sound-system as a whole presents a very unique picture of the Andamanese language in general and of Great Andamanese in particular. The current study is based on the first-hand data collected during the pilot survey conducted in 2001-2002.

Video and audio files of the relevant languages accompany the presentation.

Is Great Andamanese Typologically Divergent from Standard Average Andamanese?

Anvita Abbi

Great Andamanese is a generic term used for the amalgam of ten different but mutually intelligible varieties of the same language once spoken in the entire set of Andaman Islands in the Bay of Bengal. At present, only 37 speakers [not all are fluent in the language] that are left in the tribe speak a kind of mixed language derivative of these varieties. The pilot survey conducted in 2002 of Great Andamanese (West Andamanese) and other accessible languages of the Andaman Islands, such as Jarawa (East Andamanese), and Onge (Southern Andamanese) gives us a fairly good cross linguistic comparison to raise a significant and probing question as to whether Great Andamanese is typologically divergent and genetically distinct language from Jarawa and Onge.

The paper discusses the typological issue by presenting non-shared areas such as (i) template morphology of verb complexes. Verbs in Great Andamanese, belong to different classes based on the nature of the initial consonant of the verbal ending, each identified by a specific but different consonant followed by a vowel that represents aspect or mood. Verb roots, thus, could have any of the following CV endings: –bV or –IV or –kV or –rV or –p!V or, –mV, which may be followed by a tense marker designated by the presence or absence of a final consonant; (ii) distinct classes of personal prefixes used for inalienable possessions that are constituted of two parts, pronominal clitic indicating the possessor and the body part classifying prefix, which serves as a host to the clitic, and finally, (iii) the inventory of intriguing consonant sounds in Great Andamanese. Shared vocabulary in the realm of basic word list, kinship and body parts terms suggest that Onge and Jarawa are genetically affiliated to each other while the Great Andamanese seems to belong to a separate and distinct family. Various studies in the past, both linguistic and genetic, suggest that Andamanese languages might be the last representatives of pre-Neolithic Southeast Asia. As they represent the initial settlement by modern humans (Hagelberg et al 2002), and Genetic and epigenetic data (Endicott et al 2003) suggest long-term isolation of the Andamanese for a substantial period of time, current linguistic results are very significant for population genetics.

Cheaper by the Dozen?
Reassessing Linguistic Diversity
in the Lampungic Language Cluster
Karl Anderbeck

The 15th edition of the *Ethnologue* lists nine Lampungic languages: *Abung*, *Kayu Agung*, *Komering*, *Krui*, *Lampung*, *Southern Pesisir*, *Pubian*, *Ranau*, and *Sungkai*. Are there really nine distinct, mutually unintelligible Lampungic languages in southern Sumatra? Recent research undertaken by a partnership between [Organization 1] and [Organization 2], as well as a re-examination of earlier publications upon which the *Ethnologue*'s language listing was based, indicate that the linguistic diversity in this region has been substantially overstated. Drawing on sociolinguistic interviews, ethnographic research, comparative evidence, lexicostatistics, and the results of intelligibility testing, this paper proposes a consolidation of Lampungic isolects into three clusters, namely *Lampung Pesisir*, *Abung* and *Komering*. Maps of the distribution of phonological and lexical features as well as a redrawn language map are included, as are substantially revised population figures based on the 2000 Indonesian census.

References

- Badan Pusat Statistik Indonesia. 2004. Indonesian National Census: 2000. Jakarta.
- Gordon, Ray, ed. 2004. *Ethnologue*. 15th edition. Dallas, TX: SIL International.
- Mitani, Yasuyuki. 1980. "Languages of South Sumatra." In Tsubouchi, ed., *South Sumatra: Man and Agriculture*, pp.1 - 16.
- Sofjan Abdurrahman & Yallop, Colin. 1979. "A brief outline of Komering phonology and morphology." In Amran Halim (ed.). *Miscellaneous studies in Indonesian and languages in Indonesia*, pp. 11-18 (NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia, 7). Jakarta: Lembaga Bahasa, Universitas Katolik Indonesia Atma Jaya.
- Walker, Dale F. 1975. "A Lexical Study of Lampung Dialects". In J.W.M Verhaar (ed.). *Miscellaneous studies in Indonesian and languages in Indonesia*, pp. 11-20 (NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia, 1). Jakarta: Lembaga Bahasa, Universitas Katolik Indonesia Atma Jaya.
- Walker, Dale F. 1976. A grammar of the Lampung language: the Pesisir dialect of Way Lima. Jakarta: Badan Penyelenggara Seri NUSA.

Perceptual Evidence for Missing Vowels in Defective Vowel Systems

Torsten Andreas

This paper examines the connection between gaps in vowel systems and the vowel-quality-dependent intelligibility in high pitch singing.

A five-vowel system is the most common language type worldwide. With /i, □, a, ~, u/ the vowel space is maximally distinct. In normal speech, this dispersion minimizes vowel confusions. “Defective vowel systems” are systems in which vowels are not evenly distributed in the available space. In the UCLA Phonological Segment Inventory Database (UPSID), these non-optimal systems typically have four or five elements. Missing vowels tend to have particular vowel-quality-dependent frequency: /a/ and /i/ are less likely to be missing than /e, o, u/.

The experiments conducted in the present study test the influence of pitch and tonal movement onto vowel quality.

They were focused on high pitched steady state tones as well as contour tones with usual (a quart) as well as extreme distances (a decime) between two tones. All tones were realized on the vowel of the syllable.

Within monosyllabic pseudowords, steady state vowels and contour tones with rising and falling pitch were produced by a soprano singer and have been used as auditory stimuli for a questionnaire experiments. There were a “low” (500 Hz), a “middle” (650 Hz) and a “high tone” (850 Hz) for steady state elements and a “small” (a quart), a “larger” (a seventh) and a “very large interval” (a tenth) as the contour conditions. The lowest contour tone was realized by 350 Hz, the “high” tone and the highest contour tone were realized by 850 Hz. All conditions were realized for the five tensed german vowels /a, e, i, o, u/.

After hearing each item, the subjects were required to write the perceived elements in a list. The analysed responses concerned the number of correct and false vowel identification.

Especially for /o/ and /u/, a change in F0 or interval differences of contour tones had a profound effect, where perceptual difficulty in vowel recognition increases as F0 is raised. Moreover, poor results for /o/ and /u/ were found for both the strongly rising and the slightly falling contour condition.

In contrast to /o/ and /u/, the data suggest few influence of pitch and interval differences on /e, a, i/. Very poor results have been found for /e/; /a/ and /i/ obtained very good results.

From these findings, it suggests that the best recognized vowels /a/ and /i/ are elements that are rarely absent in the vowel systems of the world’s languages.

Syntactic Information on Semantic Maps: The Case of the Comitative

Aleksandr Arkhipov

Semantic maps as a tool for typological research are usually taken to represent namely various *semantic* functions forming the domain under scrutiny. The syntactic information finds its way onto such maps only rarely and often implicitly (cf. [Haspelmath 2003: 219]). E. g., the only syntactic function regularly distinguished with regard to the comitative markers is that of coordinating NPs.

An example of semantic map dealing with the comitative domain is F1.

In this paper, arguments will be provided in favour of regular inclusion of the syntactic dimension on semantic maps. This is in line with the growing tendency to study the grammaticalization of whole constructions, not of isolated markers. Supporting evidence will be based upon the case of the comitative constructions and markers.

It will be shown that while constructing a typological inventory of functions related to the comitative it is necessary to distinguish not only between semantic types of uses of a particular marker but also between syntactic ones. In addition, different syntactic types of uses presumably yield different diachronic implications.

The function of markers labeled ‘comitative’ in the literature may vary. In European languages, it is most often similar to (1). In Australian languages, however, the markers so called function typically as adjective-deriving affixes adjoined to nominal stems (2) (the verbal comitative is mentioned below). Another type of use which is often mixed up with either of the former, is (3).

- | | |
|--|----------------------|
| (1) <i>John came later with Mary.</i>
Wargamay < PAMA-NYUNGA [Dixon 1981: 33] | COMITATIVE
PROPER |
| (2) <i>ŋaŋa ma:l-du wugar-giri-ŋgu ŋunday</i>
1SG.O man-ERG sleepiness-COM-ERG see.UNMKD
<i>The sleepy man saw me.</i> | MODIFIER |
| (3) <i>ŋuŋa ma:l duwara-bali baŋgay-giri</i>
1SG.S man stand-CONTIN.UNMKD spear-COM
<i>The man is standing with a spear (in his hand).</i> | DEPICTIVE |

For several reasons, I regard the type (1) as the core of the domain (*comitative proper*). Type (2) involves *modifier comitative*. I will call the uses like (3) *depictive comitative* following [Schulze-Berndt & Himmelmann 2004].

The difference between (2) and (3) is mostly syntactic. In both uses, the entity denoted by the *Com[itative]P[hrase]* semantically predicates on the entity denoted by one of the core NPs (= *the Core NP* hereafter). But in (2), *ComP* modifies the *Core NP* while in (3) it is an adjunct to the verb governing *Core NP*. As for (1), it is close to (3) syntactically, but not semantically: *ComP* is also an adjunct, but the entity it denotes has the same semantic role (thematic relation) in the situation as the referent of the *Core NP*.

Now, there are quite a lot of languages in which the same comitative marker is used in contexts like (1) and (3) but not in contexts like (2); these languages typically treat adnominal modifiers as a separate class, like Mari (Uralic) or Basque (see (4)–(6)).

Given this, the map in F1 should be modified to distinguish between (1), (2) and (3); see F2. In addition to coordinating NPs we should introduce an *adjunct* syntactic function subsuming comitative proper and depictive comitative. The depictive function can further be split into several semantic subtypes like (in)alienable possession, container etc.

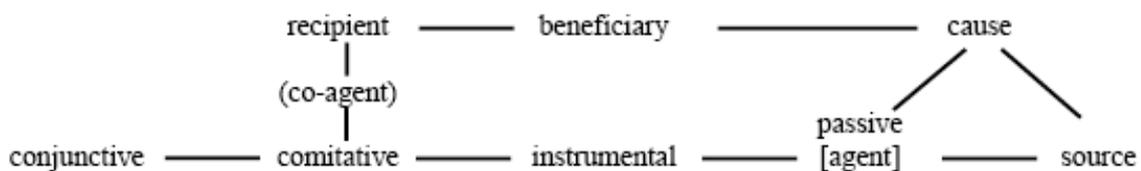
Several other syntactic functions suggest themselves, such as predicative, argument or sentential. The majority of the functions listed in F1 fall into either argument or adjunct syntactic type. Note that the comitative proper may also be expressed in an argument-type construction. This is the case in the languages using head-marking strategy, where the comitative marker attaches to the verb stem signaling the addition of an extra argument.

The resulting map would ideally have (an) extra dimension(s) devoted to syntactic parameters. In practice, however, to have a paired syntactic-semantic notation for all or some of the functions (see F3) seems to be the most convenient solution.

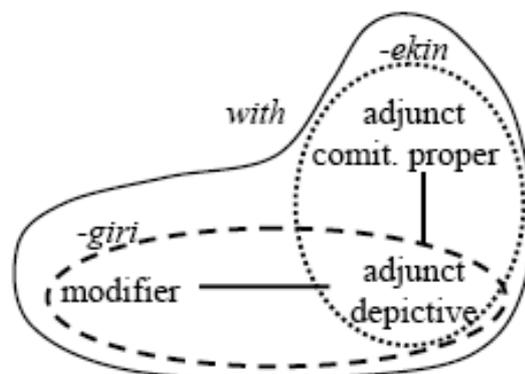
Basque (isolate, EUROPE) [elicited]

- | | | | | | | |
|-----|---|-------------------------------|----------------|---------|---------|----------------------|
| (4) | bere | emazte-a-rekin | etorri | da | | COMITATIVE
PROPER |
| | POSS | wife-DET-COM | come | AUX.3SG | | |
| | <i>He came with his wife.</i> | | | | | |
| (5) | poltsa | handi | bat-ekin | etorri | da | DEPICTIVE |
| | bag | big | DET-COM | come | AUX.3SG | |
| | <i>He came with a big bag.</i> | | | | | |
| (6) | [teilatu | {*gorri-(a-r)ekin/+gorri-dun} | etxe-a-n] | bizi | da | *MODIFIER |
| | [roof | {*red-(DET-)COM /+red-having} | house-DET-LOC] | live | AUX.3SG | |
| | <i>He lives in the house with the red roof.</i> | | | | | |

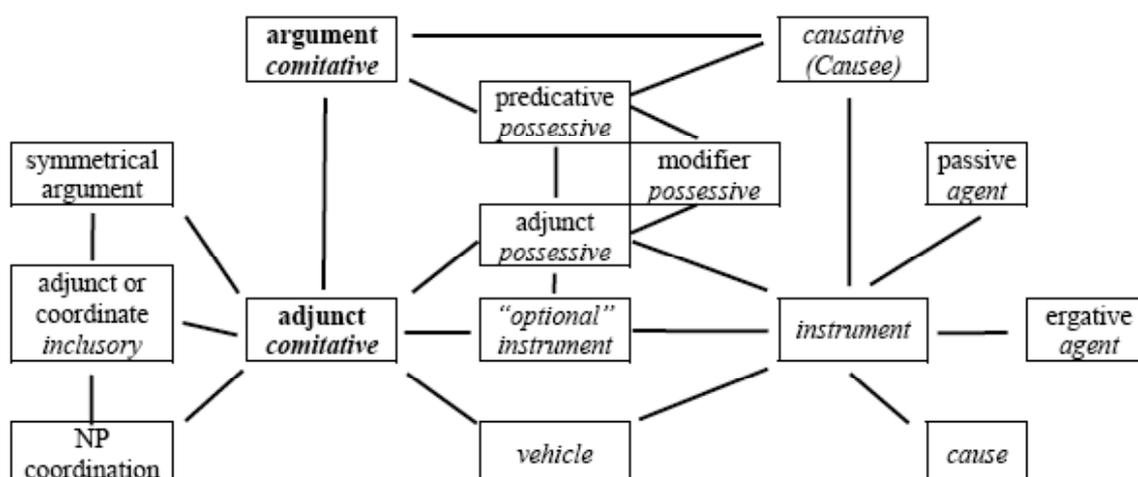
F1. Instrumental and related functions [Haspelmath 2003: FIG. 8.12, 229]



F2. A central fragment of the comitative domain (English *with*, Basque *-ekin*, Wargamay *-giri*)



F3. Comitative and several related syntactic-semantic functions



Notes on some functions listed in F3

Syntactic aspect is given in regular characters, semantic aspect is italicized. Note, however, that ‘comitative’ is defined on a complex morphosyntactic-and-semantic basis.

<i>symmetrical argument</i>	second argument occupying an obligatorily plural slot, e. g. in a reciprocal (cf. Haspelmath’s <i>co-agent</i> above).
<i>inclusory</i>	NP whose referent is included in the reference of a plural pronoun in an inclusory construction, e. g. in Russian <i>my s tobj</i> ‘you and I’, lit. ‘we with you’.
<i>adjunct possessive</i>	a kind of depictive expressing possessive relation (cf. (5)).
<i>“optional” instrument</i>	a kind of instrument optional in a given situation and not directly affecting the patient, e. g. in <i>translate with a dictionary</i> .

References

- Dixon, R. M. W. (1981) *Wargamay*. // R. M. W. Dixon & Barry J. Blake (eds.). *Handbook of Australian Languages*, vol. 2. — Canberra: The Australian National University Press, pp. 1–144.
- Haspelmath, Martin (2003) *The Geometry of Grammatical Meaning: Semantic Maps and Cross-Linguistic Comparison*. // Michael Tomasello (ed.). *The new psychology of language*, vol. 2. — Mahwah, NJ: Erlbaum, pp. 211–242.
- Schultze-Berndt, Eva; Himmelmann, Nikolaus P. (2004) *Depictive secondary predicates in cross-linguistic perspective*. // *Linguistic Typology*, v. 8, n. 1.

On the Typology of the Prohibitive

Johan van der Auwera

1. The problem

The ‘prohibitive’ is the modal category that crucially concerns an appeal of the speaker to the hearer *not* to do something (or *not* to be in a certain state). In English, the prohibitive is simply a negative Imperative - see (1). It seems that in the majority of the world’s languages, however, the Prohibitive is *not* simply the negation of the Imperative. Often, the negative marker is special, as in (2): Kaireru uses a prohibitive *sabin* instead of the ‘ordinary’ *ebai*. Sometimes the negative is ‘normal’, but the verb is not an Imperative. This is the case in Spanish (3), which uses a Subjunctive. Often also we obtain both a specialized negative marker and a non-Imperative verb. This is illustrated in (4). The general problem is why languages have at least four prohibitive strategies. The part of the problem addressed in the paper is why languages often have specialized prohibitive negations (§ 3) and why their Prohibitives often employ non-Imperative verbal forms (§ 4).

2. Data and framework

The database is a world-wide sample of about 500 languages. The approach allows reference to diachrony (more particularly, grammaticalization chains), meanings (as represented on a semantic map), and frequency of use (as explanatory of the number of semantic distinctions which a language keep apart with distinct markers).

3. The specialized negative marker

The following claims will be argued and illustrated:

- a. Specialized negative markers stem from at least 4 sources: (i) negative main verbs (e.g. ‘not want’), as in (4), (ii) positive main verbs (e.g. ‘see to it that’), (iii) either older or (iv) newer general negative markers -with the chronology making reference to the so-called Jespersen cycle—(5) illustrates how a Jespersen stage I negation survived longer in the prohibitive use.
- b. The reason why languages regularly grammaticalize special prohibitive markers has to do with an aspectual conflict. It has often been remarked (most recently also in Miestamo 2004) that the most frequent negation use, the assertoric one, has a stative character. Put in simple terms: the typical denial of a state remains a state, but the denial of an event is a state, too. The Prohibitive, however, is essentially dynamic: it is an appeal for action, and, one could argue, even for positive action, with the negativity characterizing only the desired result of the action (‘see to it that not’).

4. The non-imperative verb form

It will be argued and illustrated that there are at least three types of specialized non-Imperative Prohibitive verb forms: (i) univerbation of either the specialized or the ‘normal’ negative marker and the relevant verb form (as in (6)), (ii) the verb form is a function of the complementation pattern of the erstwhile verbal Prohibitive marker (as in (4)), or (iii) the Prohibitive is an implicature (‘Don’t do *p*’) of an assertion describing the state of affairs that the speaker desires (‘*p* will or should not be the case’) or warns against (‘unless you do something *p* will be the case’), and the morphology of the verb still betrays the assertoric origin. This, we will argue, is a partial explanation of the Subjunctive in Spanish (3).

Illustrations

- (1) Don't listen to me!
- (2) Kaireru (Austronesian; New Guinea; Ross 2002: 214)
Ur qa-qan sabin!
banana 2SG-eat PROH
'Don't eat the banana!'
- (3) Spanish
No cantes!
NEG sing.SUBJ.PRES.2SG
'Don't sing!'
- (4) Latin
Noli timere!
PROH fear.INF
'Don't fear!'
- (5) Jespersen cycle General negation Prohibitive negation
Latin I ne ne
 II ne + oenom ne
 II non ne
- (6) Djingili (Australian; Pensalfini 2003: 230)
Dunjumi-ji, mankiya-mi!
touch-PROH be,quiet-IMP
'Don't touch me, be quiet!'

Bitransitive Constructions in a Number of Languages

Bambang Kaswanti Purwo

The paper is an attempt to investigate how different and similar a number languages (English, Dutch, Italian, French, Hebrew, Indonesian, Javanese, Sundanese) handle the bitransitive constructions. Do they distinguish the ‘give’ from the ‘buy’ verb type, as exemplified in (1) through (3), where English allows the two verb types to have both constructions (a) and (b), while the Dutch ‘buy’ verb type rejects the (b) construction?

Dutch

- (1) a. Jan gaf een biertje aan Piet. ‘John gave a beer to Piet.’
b. Jan gaf Piet een biertje. ‘John gave Piet a beer.’
- (2) a. Jan kocht een biertje voor Piet. ‘John bought a beer for Piet.’
b. ?Jan kocht Piet een biertje. ‘John bought Piet a beer.’
- (3) *Jan maakt Piet een vlieger. ‘John made Piet a kite.’

The syntactic alternations of (a) and (b) as in (1) among the verbs belonging to the two verb types are further scrutinized, as exemplified below.

Among the ‘give’ verb type, i.e. the ‘explain’ subtype (like *describe*, *announce*, *introduce*, *narrate*), the (b) construction (or, the construction after “Dative Advancement”, to borrow the term from Relational Grammar) is not permissible in English and Indonesian, but is accepted in Dutch.

English

- (4) *John explained Peter the answer.

Indonesian

- (5) *Yanto menjelaskan saya jawaban. ‘Yanto-explained-me-the-answer.’

Dutch

- (6) Jan legt Piet een probleem uit. ‘John-explained-Piet-a-problem.’

Among the ‘buy’ verb type, with the verb ‘open’, English allows *Sam* in (7) to occupy the DO position because *Sam* is “the projected possessor of a beer” (Celce-Murcia and Larsen-Freeman 1999:368), but the use of *Sam* in (8) is out because the animate Object is not “the projected possessor of the door”. The equivalent constructions of both (7) and (8) are not accepted in Dutch, but they are in Indonesian, Javanese, and Sundanese.

- (7) I opened Sam a beer.
- (8) *I opened Sam the door.

References

- Celce-Murcia, Marianne and Diane Larsen-Freeman. 1999. *The Grammar Book: An ESL/EFL Teacher’s Course*. Second Edition. Heinle & Heinle.
- Kaswanti Purwo, Bambang. 1997. “The Direct Object in Bi-transitive Clauses in Indonesian”. In: *Grammatical Relations: A Functional Perspective*, T. Givon (ed.), 233–252. Amsterdam: John Benjamins.
- Kaswanti Purwo, Bambang. 2002. “Bitransitive Verbs in Some Austronesian Languages”, a paper for The 9th International Conference on Austronesian Linguistics (9ICAL), January 8–11, 2002, Australian National University, Canberra.

The History of the Ditransitive Construction in North Germanic

Jóhanna Barðdal

There are clear differences in the lexical scope, and thus in the constructional semantics, of the ditransitive construction across the modern-day North Germanic languages. One of the most prevalent differences is found between Icelandic, on the one hand, and Mainland Scandinavian, on the other, in that verbs of creation cannot be instantiated by the ditransitive construction in Icelandic. This difference is found in spite of the fact that Icelandic is the only modern Germanic language that has maintained the original five case constructions of the Germanic ditransitive in which the morphological case of the two objects can vary between Dat-Acc, Dat-Dat, Acc-Dat, Acc-Gen and Dat-Gen. Also, as opposed to English, verbs of ballistic motion cannot occur in the ditransitive construction in Germanic and verbs of obtaining only have a very limited access to it.

This raises the question whether the lexical and the semantic range of the ditransitive construction has contracted in the history of Icelandic but expanded in the other Germanic and North Germanic languages, or whether the facts of the present-day languages reflect differences existing already before the beginning of recorded history. A preliminary comparison between texts from two different periods of Icelandic reveals that there is a substantial reduction of 50% in the type and text frequency of the ditransitive construction from Old Norse-Icelandic to Modern Icelandic (cf. Barðdal, in prep: ch. 4). The question arises whether this decrease in the use of the ditransitive construction in Icelandic is also concomitant with a reduction in its lexical and semantic scope.

In this paper, a detailed analysis of the Icelandic facts will be presented, as well as the results of an ongoing investigation on material from Swedish and Norwegian dialects which have maintained the dative case. These dialects are almost extinct today but were documented in the Seventies by Reinhammar (1973). This material is available for linguistic research, as is Reinhammar's original discussion of how the dialectal facts tie in with the historical facts for this data set.

More generally, the aim of the present paper is to contribute to a discussion within CxG on whether and how the expansion and/or the constriction of a construction takes place over time in the history of individual languages. I will also discuss how such historical development links with new-sprung ideas of syntactic productivity (Barðdal, in prep), in which the extensibility of a syntactic construction is taken to be the function of the construction's type frequency and its semantic coherence and an inverse correlation between the two.

References

- Barðdal, Jóhanna. In Prep. *Productivity: Evidence from Case and Argument Structure in Icelandic*. To appear in John Benjamins' *Constructional Approaches to Language* series, edited by J.-O. Östman & M. Fried.
- Reinhammar, Maj. 1973. *Om dativ i svenska och norska dialekter*. Acta Academiae Regiae Gustavi Adolphi 53. Uppsala: Almqvist & Wiksell.

Origin and Control of the Initial Phase of the Process: How Does Kashmiri Cope with the Lack of Agent

Estella del Bon

The present paper examines the relationship between semantic properties of the role of agent (source/origin on one hand, control/volitionality on the other hand) and expression of the initial phase of the process in impersonal clauses of Kashmiri.

Broadly defined, the impersonal construction is characterized by the presence of a non-referential agent (NRA). Impersonal clauses are transitive¹ and the NRA appears on the finite verb as a third person singular marker which cannot be made explicit (ex.1: *-n*), though it usually has the form of an agent-agreement marker in personal clauses (ex.2: *-n* corefers *pujan* ‘butcher’).

Interestingly, impersonal clauses behave differently depending on whether or not they are constructed with the aspectual auxiliary *hyon* ‘start/begin’. When used with any other verb than the auxiliary *hyon*, the impersonal construction shows the following characteristics:

Impersonal clauses always occur in alternation with an intransitive personal clause of equivalent meaning (ex.3 is alternative to ex.1). The personal intransitive clause corresponds to the neutral description of the situation whereas the impersonal construction has a marked aspectual value; it indicates a change of situation² (ex.1).

Not all intransitive clauses accept impersonal clause as an alternative. The situations liable to accept the impersonal construction (from now onwards referred to as class A situations) may be roughly defined by the fact that they lack a potential agent, i.e. they lack a participant likely to assume agentive properties whatever the context may be. Typically, class A situations are: natural/meteorological phenomena and mental and body affections (diseases, growing old, etc.).

If we now consider the impersonal construction when used with the aspectual auxiliary *hyon*, (whose aspectual value, the inceptive, is distinct from that of change of situation³) we observe a somewhat different pattern:

For situations of class A, there is no personal alternative to the impersonal clause. It will be shown that this is to be explained in terms of agentive properties rather than Aktionsart proper. The introduction of the NRA only makes it possible to express the inceptive for situations that lack a participant which can be associated with the source/origin of the process.

¹ The verb belongs to the class of transitives and the clause shows (in accordance with the pattern of personal transitive clauses) accusative structure in the imperfective and ergative structure in the perfective.

² The speaker stresses on the change, or apparition of a (new) situation, or, said in other words, on the passage from a previous to a new situation (with implicit reference to prior situation). The aspectual category here involved is quite similar to that expressed by mandarin final particle *le* (*nǐ kàn ! xià yǔ le*, 2S look ! fall rain LE, “Look ! it is raining [now]”). So far, this aspectual category seems not to be very common among world languages.

³ Though both aspectual categories pertain to the sphere of the initial phase of a process, the former focusses on the very beginning of the process (i.e. initial boundary of the process) while the latter stresses on the fact that the initial boundary has been already crossed-over (it is therefore not concerned with the boundary itself). The auxiliary *hyon* behaves like a transitive verb (as an independent verb *hyon* ‘to take/grasp/catch’ is a transitive). Therefore, even when the main verb is an intransitive, the clause is of ergative structure in the perfective: the sole argument of the clause is ergative case marked, and the auxiliary, as well as the main verb (in the infinitive), show neutral gender/number agreement (i.e. masculine singular).

Even non-class A situations can accept the impersonal construction. Unlike with class A situations, the impersonal clause here occurs in alternation with a personal clause. The hypothesis to be demonstrated here is that the alternation impersonal vs personal clauses is to be interpreted neither in terms of Aktionsart nor in terms of source/origin of the process, but rather as corresponding to the semantic opposition [control -] vs [control +] of the main participant over the process.

Impersonal constructions analogous to that of Kashmiri have been reported in Burushaski, Shina, Kalasha and Limbu. Further research shall answer the question as to what extent these impersonal clauses can be related to the aspectual category found in Kashmiri, and how far it is relevant to raise the question of a typological feature common to the (western/central ?) himalayan area.

- | | | | | |
|----|------------------------|-----------------------------|----------------------|--|
| 1. | ru:d | lo:gu-n ⁴ | | <i>"It rained"</i>
<i>(the weather has changed, it has started raining, and now it is raining).</i> |
| | rain.MS-Ø | apply.PTMS-3S | | |
| 2. | (*puj-an) ⁵ | mo:ru-n | kaT | <i>"The butcher killed the goat."</i> |
| | <u>butcher.MS-ERG</u> | <u>kill.PTMS-3S</u> | goat.MS-Ø | |
| 3. | ru:d | pyov | | <i>"It rained."</i> |
| | rain.MS-Ø | fall.PTMS | | |
| 4. | ru:d | hyotu-n | py-on | <i>"It started raining."</i> |
| | rain.MS-Ø | begin/start.PTMS-3S | fall-INFMS | |
| 5. | * ru:d-an | hyot | py-on | <i>"It started raining."</i> |
| | rain.MS-ERG | begin/start.PTMS | fall-INFMS | |
| 6. | sɽ | hetsi-n' | pak-in' ⁶ | <i>"She started walking."</i> |
| | 3FS-Ø | begin/start.PTFS-3S | fall-INFFS | |
| 7. | tami | hyot | pak-un | <i>"She started walking."</i> |
| | 3FS.ERG | begin/start.PTMS | fall-INFMS | |

⁴ Abbreviations used in this paper are : M masculine; F feminine; S singular; P plural; ERG ergative ; INF infinitive; PT preterite. Gender/number agreement of the finite verb is shown in bold; the marker of the non-referential agent is shown in red bold; personal indexes (other than NRA) and their coreferent argument are underlined.

⁵ Cooccurrence of third person singular index -n with the corresponding nominal *pujan* 'butcher' is impossible. With deletion of nominal *pujan*, the unmarked word order of the sentence would rather be *kaT mo:run*, but for convenience of the present paper we are giving here the marked word order *mo:run kaT*, so that all examples in this paper show same word order pattern.

⁶ Palatalisation of final consonant is indicated by '. Retroflexes are in capitals.

Ergative Case in Nias

Lea Brown

Nias, an Austronesian language spoken on the island of Nias and its surrounding islands off the west coast of Sumatra, has an unusual case marking system in which nouns in the ergative case are formally unmarked while nouns in the absolutive case are marked by initial mutation (Brown 2001, 2005). While the labels 'ergative' and 'absolutive' are appropriate for the uses of these two case forms when they occur as core nominal arguments, there are various other uses of the two cases that do not fit within what is usually understood by these labels. (For this reason, I prefer to use the labels 'unmutated' and 'mutated' for the forms used to mark ergative and absolutive nominals respectively.)

First, there are uses of the unmutated forms for NPs which are clearly not ergative in the usual sense, and where these uses could be understood as involving the use of nominals not marked for case rather than occurring in the same inflectional case as that of ergative nominals. This includes (1) nominal predicates; (2) secondary nominal predicates (e.g. the complement of *fotōi* 'to call something X'); (3) arguments of the negative existential verb *lōna*; (4) complements of the verb *mae* 'be like'; (5) any preverbal argument (the unmarked position for core arguments is postverbal); (6) after a few prepositions; (7) with locative nouns in certain locative phrases; (8) with instrumental nouns and (9) as the third argument with certain verbs that take three semantic arguments (e.g. 'X [unmutated] throws Y [unmutated] at Z [mutated], 'X [unmutated] fills Y [mutated] with Z [unmutated]'). Unmutated forms are also used for non-initial NPs in conjoined NPs and in lists, although this could be analysed by treating initial mutation as a proclitic case marker that attaches to the first word of an NP consisting of conjoined or listed NPs. Unmutated ('ergative') forms are also used for grammatically absolutive NPs (both S's and P's) in relative clauses, and for P's in nominalizations, progressive clauses, immediate future clauses, and purposive clauses.

Some of the uses of the mutated forms are ones which correspond to uses of 'absolutive' in other languages and are unsurprising. These include the use of mutated forms for (1) the argument of a nominal predicate; (2) the argument of the existential verb (in contrast with the use of the unmutated form for the argument of the negative existential verb); and (3) NPs in apposition to the plural (absolutive) argument of a reciprocal verb. Other uses of mutated forms are less obviously ones that could be expected of absolutive case. These include (1) the possessor in the genitive construction; (2) the object of most prepositions; and (3) the argument of a numeral+classifier construction or the quantifier *fefu* 'all'. There are also a number of semantically transitive experiencer verbs both of whose arguments are mutated (though the two arguments behave differently in some respects). And while NPs in S and P functions are (unexpectedly) unmutated in relative clauses, NPs functioning as transitive subjects in clauses relativising P's or datives are mutated.

The case system for pronouns functions in a similar way to that of nouns, but differs in certain respects. Most importantly, there is a distinct genitive case form for pronouns, so that the use of mutated forms for noun genitives could be treated as syncretism of absolutive and genitive cases.

as a way of referring to the two men. (This example comes from a translation of a children's story from English into Walman; however, there is nothing in the process of translation that would explain this, and such usage is also found in original texts.)

Nor is it clear how such an approach would account for examples like (10), in which we have two occurrences of 'and', namely *maro* 'me and her' and *naro* 'you and her', functioning like pronouns which themselves serve as conjuncts of a third occurrence of 'and', namely *kcha* 'we and you'.

Examples

- (3) kum m-ch-a chi k-ranychwa.
 1SG 1SG.SUBJ-2.OBJ-and 2SG 1PL.SUBJ-play
 'Let's you and me play!'
- (4) kum m-ch-ekrien chi.
 1SG 1SG.SUBJ-2.OBJ-see 2SG
 'I saw you.'
- (5) m-ch-a k-ranychwa.
 1SG.SUBJ-2.OBJ-and 1PL.SUBJ-play
 'Let's you and me play!'
- (6) runon n-aro-n Willie.
 3SG.MASC 3SG.MASC.SUBJ-and-3SG.MASC.OBJ Willie
 'He is with Willie.'
- (7) Steve n-aro-n John y-ara, korue
 Steve 3SG.MASC.SUBJ-and-3SG.MASC.OBJ John 3PL.SUBJ-come but
 Steve n-ara eni, John n-ara amtre.
 Steve 3SG.MASC.SUBJ-come today John 3SG.MASC.SUBJ-come
 yesterday
 'Steve and John came, but Steve came today and John came yesterday'
- (8) John n-aro-n Simon y-o ro<y>rani.
 John 3SG.MASC.SUBJ-and-3SG.MASC.OBJ Simon 3PL.SUBJ-be tall<PL>
 'John and Simon are tall'
- (9) Kamtei wiei pa ri yiripil. Rim y-etere wuel.
 men two PRT 3PL 3PL.SUBJ-wake.up 3PL 3PL.SUBJ pig
 'The men woke up. They saw the pig.'
 Tu w-esi chalien. Wuel won no kisiel
 MIRAT 3SG.FEM-go.out net pig heart be angry
 'The pig was out of the net.' 'The pig was very angry with the men.'
- runon n-aro-n. Willie*
 3SG.MASC 3SG.MASC.SUBJ-and-3SG.MASC.OBJ pig
 'The pig
- w-rachre-y *runon n-aro-n.*
 3SG.FEM.SUBJ-chase-3PL.OBJ 3SG.MASC 3SG.MASC.SUBJ-and-3SG.MASC.OBJ
 chased the men.'
- Runon n-aro-n y-arul kisiel.*
 3SG.MASC 3SG.MASC.SUBJ-and-3SG.MASC.OBJ 3PL.SUBJ-run.away fast
 'The men ran away as fast as they could.'

The Nature of Labialvelar Obstruents

Michael Cahill

Labialvelar obstruents ($\widehat{k}p$, $\widehat{g}b$, $\widehat{\eta}m$, generically “KP” here) occur in at least 7% of the world’s languages, mostly in Africa, but in a few dozen languages in Papua New Guinea and Oceania as well. One of the unsettled questions regarding these is whether they should be regarded as having one primary place of articulation, be it labial or velar, or rather be regarded as having two coequal places of articulation. Chomsky and Halle (1968) as well as Anderson (1976) favored the one-primary-place view, although it was a language-specific matter which place was primary. Ohala & Lorentz (1977) insisted that labialvelars had two co-equal places of articulation, but had no formal theory available at the time for doing so. Under a different theoretical modal, Sagey (1990) and Clements and Hume (1995) were able to formulate labialvelars as having either one or two primary places of articulation. Currently, then, investigators are at theoretical liberty to propose either. However, any such proposals must be informed by a wide variety of cross-linguistic data, and such data must be examined in the light of the asymmetric gestures of labialvelars to avoid spurious claims. That is the aim of this paper, which includes data from several African languages and two from Papua New Guinea.

First we examine what evidence should *not* count in considering primary place of KP. Labialvelars do not have absolutely simultaneous gestures; the velar gesture slightly precedes the labial one. This phonetic fact must be taken into account when examining phonological patterns, since assimilations and other phonology to the left of KP could well be sensitive to the velar articulation, and those to the right of KP would be more sensitive to the labial one. One example of this is nasal place assimilation. In some languages, a nasal assimilates completely to a following KP as [$\widehat{\eta}mKP$], but many languages it partially assimilates, as [ηKP] (see (1), (2)). This might be thought of as evidence for the primarity of velar place. However, in the languages in which a nasal follows KP, a partial assimilation yields [KPm] (3). Thus nasal place assimilation evidence cannot be taken as decisive in the question of primarity of KP place.

There does exist data that is more decisive, from a variety of phenomena. One pattern is consonant co-occurrence restrictions. In at least the Kaanse, Ngbaka, Kukœ, and Bari languages, KP cannot co-occur with a labial in a word, though it can with a velar. This restriction is non-directional. Another pattern is that in Dagbani and Nzema, KP has an allophone TP before front vowels, preserving the labial place while replacing the velar place. A third pattern is the neutralization of KP with labials word-finally in Amele, Ndyuka, and possibly Efik (4). Again the labial portion of KP is preserved. A fourth is the blocking of round vowel harmony in Nawuri, in which KP patterns with labials. The vowel of the prefix /gI-/ surfaces as round if the stem vowel is round. But if the initial stem consonant is labial (including labial-velars), the process is blocked (5).

When the data that is influenced by the phonetic edges of KP is filtered out, the positive patterns above give cross-linguistic evidence that labial is the primary place of articulation for labialvelar obstruents.

- (1) Nasal assimilation as [ŋmKP] in
- a. GÈdÈr, (Kleiner 1989)

□ m -gbunto□o□	‘carried thing’	□ m -kpa□	‘life’
-----------------------	-----------------	------------------	--------
 - b. Yeletnye (Papua New Guinea) (Henderson 1975)

□ mgb ’	‘frigate bird’	□ mgb □□	‘overripe’
----------------	----------------	-----------------	------------
- (2) Nasal assimilation as [ŋKP] in Dagaari (Kennedy 1966, personal data):
- | | | | |
|--------------|-------------|-------------|----------|
| kpa\$Nkpa\$N | ‘upper arm’ | N@ kpa@no\$ | ‘I lock’ |
|--------------|-------------|-------------|----------|
- (3) Nasal assimilation as [KPM] Senufo (Mills 1984)
- | | | | |
|----------------|-----------|-----------------------|---------------|
| kpm ~Œ: | ‘to beat’ | n7...- gbm ~Œ: | ‘herb doctor’ |
|----------------|-----------|-----------------------|---------------|
- (4) Neutralization of /gb/ and /b/word-finally in Amele (PNG) (Roberts 1987).
Word-medial ‘we’ is [gb]
- | | |
|---------------------|-------------------|
| [»ho- gb -’] | ‘we came (today)’ |
| [hç-»lç- p] | ‘we used to come’ |
- (5) Blocking of round harmony: Nawuri (Casali 1995)
- a. before [-round] stem vowel:

[g̃,̄i]	‘tooth’	[g̃,ke:li]	‘kapok tree’
---------	---------	------------	--------------
 - b. before [+round] stem vowel - round spread:

[gu-jo]	‘yam’	[gu-ku:]	‘digging’
---------	-------	----------	-----------
 - c. before [+round] stem vowel and labial stem consonant - blocking of spread

[g̃,pula]	‘burial’	[g̃-kpo:]	(type of dance)
-----------	----------	-----------	-----------------

Variable Word order in Toba Batak

Peter Cole and Gabriella Hermon

Toba Batak has been described in Schachter (1984) and Clark (1984) as a VOS language. The VOS order is described in Greenberg (1966) as a rare dominant order. Moreover Greenberg notes that VOS languages pattern with VSO and SVO rather than with SOV languages, in terms of his word order correlations.

We shall address these issues within the framework of generative grammar. We shall discuss the question of what the phrase structure and derivations are for VOS clauses and how this order is related to the other observed word orders in Toba Batak (SVO, and somewhat less frequently, VSO). We shall argue that the simplest account for Toba Batak supports deriving all word orders from underlying SVO structure via leftward movement, rather than treating VOS clauses as clauses which have a base generated specifier to the right.

An SVO basic word order and the movement to the left will explain a variety of facts about the order of the DO and IO in the VP, the order of VP adverbials, and the so-called 'freezing' effect on the NPs following both active and passive verbs. It will also explain the seemingly strange facts related to reflexives, namely the fact that in the passive voice the agent can antecede a reflexive in subject position.

Similar analyses have been proposed for other subject final Austronesian languages. (See Pearson 2001 for Malagasy and Rackowski and Richards 2004 for Tagalog). Our results raise the question of whether all VOS languages are best described as derived from an underlying SVO order.

References

- Clark, Robin. 1984. "The Syntactic Nature of Logical Form: Evidence from Toba Batak", in P. Schachter (ed.) *Studies in the Structure of Toba Batak*. Dept. of Linguistics, UCLA.
- Greenberg, J. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg, ed., *Universals of Language*, MIT Press, pp. 73-113
- Pearson, M. (2001). *The Clause Structure of Malagasy: A Minimalist Approach*, UCLA Dissertations in Linguistics 21. UCLA Linguistics Department
- Rackowski, A. and N. Richards. (2004). Phase edge and extraction: a Tagalog case study. Ms, Linguistics Department, MIT, Cambridge, Mass.
- Schachter, Paul. 1984. "Semantic-Role-Based Syntax in Toba Batak", in P. Schachter, (ed.) *Studies in the Structure of Toba Batak*. Dept. of Linguistics, UCLA

Some Argument-Structure Properties of ‘Give’ in the Languages of Europe and Northern and Central Asia

Bernard Comrie

Haspelmath (forthcoming) provides a typology of the expression of the object arguments of the verb ‘give’ using three basic types, according to whether the recipient, the gift, or neither is overtly marked differently from patients of monotransitive verbs:

indirect object, e.g. ‘Mary gives the book to John’
primary object, e.g. ‘Mary gives John with the book’
double object, e.g. ‘Mary gives John the book’

In addition, he allows a mixed type for languages that permit more than one of these types, such as English, which has both indirect object and double object constructions. A quick glance at the map accompanying his text shows that Europe and Northern and Central Asia are dominated by the indirect object type, with exceptions being largely geographically peripheral or, as in the case of Ket, in a language that is otherwise typologically unusual for its location: For instance, the primary object type is found in Nivkh, the double object construction in Ainu and Ket, and the mixed type in English (and in some other neighboring Germanic languages not included in Haspelmath’s sample).

While not challenging Haspelmath’s basic finding, and indeed in part following up on passing observations made by Haspelmath, this paper aims to show that consideration of a broader range of morphosyntactic features than just the overt marking of recipient and gift leads to a more differentiated picture of argument-structure properties of ‘give’ in what otherwise seems a very homogeneous area. In particular, many languages show features that indicate a less marked, morphosyntactically more central status of the recipient. Properties to be discussed include the following:

- a) Differences between full noun phrases and pronouns as object arguments of ‘give’, as when Maltese has the double object construction with pronoun objects, alongside the indirect object construction elsewhere.
- b) Verb-object agreement, including the use of resumptive pronouns, which often gives preference to resumptive dative pronouns agreeing with the recipient over resumptive accusative pronouns agreeing with the gift, as in Spanish.
- c) Passive verb forms, which may give preference to promotion-to-subject of the recipient rather than of the theme, in the indirect object construction (e.g. Japanese) and in the double object construction (e.g. English).
- d) Constituent order, with the possibility of having the recipient rather than the gift closer to the verb in the double object construction (e.g. English, Mandarin Chinese) and perhaps even in the indirect object construction.
- e) Stem suppletion according to features of the recipient rather than of the gift, as in Old Basque, Tsez (East Caucasian family), Yukaghir, and Japanese.

Reference

Haspelmath, Martin. Forthcoming. Ditransitive constructions: the verb ‘give’. In Matthew Dryer, Martin Haspelmath, David Gil, and Bernard Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Did Colonial Valley Zapotec Numerals Violate a Linguistic Universal (or Two)? Bernard Comrie & Pamela Munro

A number of languages of the world construct some of their numerals by means of subtraction, as with Latin, where 19 is expressed as *un-de-viginti* (i.e. 1-from-20) and 18 as *duo-de-viginti* (i.e. 2-from-20). A reasonable question for a typologist to pose with regard to such use of subtraction concerns the set of constraints to which subtraction might be subject cross-linguistically. Two such constraints are proposed by Greenberg (1978: 260–261). The first says (with correction of a typographical error): “If a number n is expressed by subtraction as $y - x$, then every number z ($y > z > n$) is also expressed subtractively and with y as the minuend.” The second says: “Every minuend is a base of the system or a multiple of the base.”

Colonial Valley Zapotec (CVZ), the variety of Zapotec (Oto-Manguean family) described in Juan de Córdova’s *Arte del idioma Zapoteco* (1578), is identified as a “minor exception” to the first of these universals, and as a clear exception to the second. CVZ has a basically vigesimal numeral system, and uses, or at least can use, subtraction for the numerals 55–59, 75–79, 95–99, and 115–119; relevant higher numerals are not cited by Córdova, and this way of forming the relevant numerals is not found in modern Zapotecan languages. The forms for 55–59 are given on the attached data page. While the precise way in which these forms come to mean what they mean is not entirely clear, the longer form for 55 seems to be something like ‘[it] is [with] another five [that it] will go to sixty’, which, while not explicitly subtractive in form, can be interpreted functionally as an instruction to subtract 5 from 60. Given that 55 is expressed by subtraction (from 60), the first universal would predict that 56–59 should also be expressed as subtraction, and more specifically as subtraction from 60. The second, longer variant in each case is as predicted, e.g. 56 is something like ‘[with] another four [it] will go to sixty’, i.e. an instruction to subtract 4 from 60. Greenberg analyzes the shorter variants of 56–59 as 5 from $(60 + 1)$ through 5 from $(60 + 4)$, which thus violate the first universal, since the minuend for numbers between 56 and 59 is not 60, the minuend for 55, and also the second universal, since none of 61–64 is a base of the system or a multiple of the base.

We argue that Greenberg’s analysis of the apparent exceptions is unjustified, with a plausible alternative having a different bracketing, namely $(5 \text{ from } 60) + 1$ through $(5 \text{ from } 60) + 4$, i.e. simply involving addition to the already established item for 55, a process that is attested elsewhere in the CVZ numeral system, e.g. 16 can be expressed as $15 + 1$. This leads us to more general reflections on the nature of possible minuends in subtractive numeral formations, a plausible hypothesis being that such minuends must be “salient” numerals within the system, thus including not only bases but also low numbers. This proposal encompasses other exceptions noted by Greenberg: in (one variety of) Montagnais, 7 is $10 - 3$ or $8 - 1$, 9 is $10 - 1$, but 8 is 4×2 , in a decimal system; in (one variety of) Arikara, 7 is $8 - 1$, 9 is $10 - 1$, and 11 is $12 - 1$, in a decimal system. We assume that 8 and 12 are, or can be, salient even in a decimal system; none of the numerals 61–64 can be.

Data: Selected Colonial Valley Zapotec Numerals

Note: Not all orthographic and minor morphological variants are listed, and the precise interpretation of some morphemes, as well as the use of commas, is uncertain. The definite aspect prefix is used on numbers to mean ‘another n’.

55 ce-caa quiona
DEF-five sixty

ce-caayo zaa, qui-zaha, chaa c-aca cayona
DEF-five ? IRR-walk and IRR-be sixty

56 ce-caayo quiona-bi-tobi
DEF-five sixty-and-one

ce-tapa c-aca, qui-zaha, chaa c-aca cayona
DEF-four IRR-be IRR-walk and IRR-be sixty

57 ce-caa quiona-bi-topa
DEF-five sixty-and-two

ce-chona c-aca qui-zaha chaa cayona
DEF-three IRR-be IRR-walk and sixty

58 ce-caa quiona-bi-chona
DEF-five sixty-and-three

ce-topa c-aca, qui-zaha chaa cayona
DEF-two IRR-be IRR-walk and sixty

59 ce-caa quiona-bi-tapa
DEF-five sixty-and-four

ce-tobi c-aca, qui-zaha, chaa cayona
DEF-one IRR-be IRR-walk and sixty

Abbreviations:

DEF definite

IRR irrealis

Reference

Greenberg, Joseph H. 1978. Generalizations about numeral systems. In Joseph H. Greenberg et al., eds., *Universals of Human Language*, vol. 3: *Word Structure*, Stanford: Stanford University Press, 249–295.

The Canonical Approach in Typology: Suppletion and Features

Greville Corbett

Suppletion (*go* ~ *went*) represents the extreme of inflectional morphology, since lexemes can have inflected forms sharing no phonological material. This is therefore a challenging area for typologists. Investigations since Ostoff (1899) provide sufficient data for an outline typology, which is of an unusual kind, since its domain is individual lexemes. We adopt a ‘canonical’ approach to our typology, locating a point in theoretical space from which real examples of suppletion are calibrated. We start from Mel’čuk’s definition (1994:358): ‘For the signs X and Y to be suppletive their semantic correlation should be maximally regular, while their formal correlation is maximally irregular.’ From this we establish dimensions along which suppletive lexemes vary. Canonical instances combine maximal semantic clarity with maximal formal opacity. Our criteria for canonicity follow from Mel’čuk’s definition, and we illustrate each. (‘>’ means ‘is more canonically suppletive than’.)

A. Lexeme-internal criteria

- 1: fused exponence > stem-only suppletion. This criterion relates to formal opacity; canonical suppletion has stem and affix fused. Thus *bad* ~ *worse* is a more canonical (better!) instance than *good* ~ *bett-er*.
- 2: full > partial. Russian *idti* ‘go’ ~ *šel* ‘went’ (full suppletion) > English *bring* ~ *brought* (partial).
- 3: zero > phonological material. Russian ‘be’: with Ø (present) ~ *byl* (past) > ‘go’ *idu* ~ *šel*.
- 4: more variants > fewer variants. Georgian ‘come’, with four suppletive stems > Russian ‘go’ with two.
- 5: morphological (Aronoff’s ‘morphomic’) distribution of stems > morphosyntactic. Purely morphological patterning (French *aller* ‘go’) > morphosyntactic patterning (Russian *idti* ‘go’).
- 6: non-alternating > alternating. Canonically, suppletive forms are categorical in use. Less canonical are alternating forms, like Russian *čelovek* ‘person’, genitive plural *čelovek* **or** *ljudej*.
- 7: less relevant feature > more relevant (cf: Bybee 1985:92)
- 8: contextual feature > inherent (Booij 1996). Both criteria can be illustrated from East Norwegian, where the adjective *liten* ~ *vesle* ‘little’ suppletes in part according to definiteness (a contextual, non-relevant feature).

B. Lexeme-external criteria

- 9: complementary stems > overlapping with other lexemes; French *aller* ‘go’ > Spanish *ir* ‘go’, preterit *fuí*, overlapping with *ser* ‘be’ (Juge 1999).
- 10: no remainders > remainders. Russian *rebenok* ‘child’ ~ *deti* ‘children’ is non-canonical, since lexically related ‘remainders’ exist, like *rebjata* ‘chaps’.
- 11: unique > non-unique. The existence of derived forms (e.g. *undergo* ~ *underwent*) makes *go* less canonical than an equivalent without derivatives.
- 12: no outside condition > outside condition. Russian genitive plural options *čelovek* / *ljudej* ‘people’ (see 6) are conditioned, hence they are even less canonical.

13: syntactic effects > no syntactic effects. Canonical instances are morphologyinternal. A non-canonical instance is (partially suppletive) Serbian-Croatian- Bosnian *d(ij)ete ~ d(j)eca* ‘child(ren)’, which has dramatic effects on agreement.

Conclusion

Our criteria converge on the canonical instances (there is no ranking of criteria). The typology thus identifies the indisputable and the most interesting examples of suppletion, some of which are quite remarkable. Our results give support to the canonical approach to typology, while points 5-8 represent a contribution to the typology of features.

Uncommon Patterns of Core Term Marking: Antiaccusative and Antiergative

Denis Creissels

Alignment typology compares the coding properties of NPs representing the two arguments of prototypical action verbs, agent (A) and patient (P), in constructions in which both are treated as core terms, with those of the term S that shows a maximum of properties typical for core terms in intransitive constructions (that is, in constructions that do not include a couple <A,P>). The coding properties of S may coincide with those of A (accusative alignment), they may coincide with those of P (ergative alignment), they may show a mixture of A-like and P-like coding properties (mixed alignment), and finally, they may coincide, neither with those of A, nor with those of P (neutral alignment).

This paper examines a particular aspect of this typology, namely the manifestations of the contrast between A, P and S at the level of the NPs assuming these roles: *core term marking* includes both the attachment of case marks to nouns or NPs assuming a core syntactic role, and the combination of NPs with adpositions used as core syntactic role markers.

In this paper, I argue that the most basic notion in a typology of core term marking is the contrast between syntactically marked forms of the noun, which must be licensed by a syntactic context in which the noun assumes a particular role, and a syntactically unmarked form, used for quotation or in a function of pure designation, and also, as a default form, in syntactic contexts that do not require the use of a syntactically marked form. This notion of syntactic markedness must be distinguished from morphological markedness, since the form of the noun recognized as syntactically unmarked according to this definition does not necessarily coincide with the basis to which morphological operations giving rise to the other forms of the noun apply. The characterization of core term marking systems put forward in this paper primarily relies on syntactic markedness, but for each type, a further distinction can be introduced between a more typical variant, which follows the general tendency towards coincidence between syntactic and morphological markedness, and a less common variant, which contradicts this tendency.

Descriptions of core term marking generally use *nominative* (the form used for S and A) *vs accusative* (the form used for P) for systems following accusative alignment, and *absolutive* (the form used for S and P) *vs ergative* (the form used for A) for systems following ergative alignment. The terms *nominative* and *absolutive* strongly suggest that the form of the noun used in an extra-syntactic function of pure designation necessarily coincides, either with the form used in S and A roles (in systems following accusative alignment), or with the form used in S and P roles (in systems following ergative alignment). But this is not always the case, and current terminology must be revised in order to develop a consistent typology of atypical systems of core term marking (i.e., systems that depart from the tendency to use of the extra-syntactic form of the noun, either for S and A, or for S and P).

The proposals I put forward in this paper can be summarized as follows:

1. By virtue of their etymology, *nominative* as well as *absolutive* are possible labels for nouns forms used in the extra-syntactic function of pure designation, irrespective of their distribution in syntactic contexts. However, in current practice, each of them has specialized to systems in which a syntactically unmarked form of the noun assumes a particular range of syntactic uses. Therefore, the proposal to give them back their original meaning would be a

source of misunderstandings. Consequently, I propose *absolute form* as a universal label for the form taken by nouns in the extrasyntactic function of pure designation, without any hint at its possible syntactic uses.

2. *Accusative* is maintained as a label for syntactically marked forms of nouns used for P, but not for A or S, and *ergative* is maintained as a label for syntactically marked forms of nouns used for A, but not for P or S.
3. New terms are introduced for syntactically marked forms involved in less common patterns of core term marking: a syntactically marked form common to nouns in S and A roles is labelled *antiaccusative*, and a syntactically marked form common to nouns in S and P roles will be labelled *antiergative*.

On the basis of these definitions, illustrations of the following types of atypical systems of core term marking will be presented; (a) and (b) are atypical manifestations of the accusative type of alignment, and (c) and (d) are atypical manifestations of the ergative type of alignment:

- (a) antiaccusative S & A vs absolute P (so-called 'marked-nominative', common among African languages)
- (b) antiaccusative S & A vs accusative P (Japanese, Korean)
- (c) antiergative S & P vs absolute A (Nias).
- (d) antiergative S & P vs ergative A (several Polynesian languages).

References

- Aikhenvald, A., R. Dixon & M. Onishi (eds). 2001. *Non-Canonical Marking of Subjects and Objects*. Amsterdam: J. Benjamins.
- Brown, L. 2001. *A grammar of Nias Selatan*. PhD thesis. University of Sydney.
- Comrie, B. 1989. *Language universals and linguistic typology*, 2nd ed. Oxford: Blackwell.
- Dixon, R. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- König, C. Forthcoming. The marked-nominative languages of eastern-Africa. In B. Heine & D. Nurse (eds) *Africa as a linguistic area*.
- Lazard, G. 1994. *L'actance*. Paris : Presses Universitaires de France.
- Moyse-Faurie, C. & F. Ozanne-Rivierre. 1983. Subject case markers and word order in New Caledonia and Loyalty Islands Languages. In A. Halipm, L. Carrington & S. Wurm (eds) *Papers from the Third International Conference on Austronesian Linguistics*, vol. 4: *Thematic variation*, 113-152. *Pacific Linguistics*, C-77.

On the Typology of Content Interrogatives

Michael Cysouw

The seven modern standard English question words (*who, what, which, when, where, how, why*) are often considered to represent some basic concepts of linguistic structure. However, when studying the actual cross-linguistic diversity of content interrogatives, these concepts are clearly not universal in any sense. For example, many languages do not have a monomorphemic expressional equivalent to 'when' or 'why'. In contrast, many languages have a monomorphemic interrogative asking for quantity (English compound *how much/many*).

I will present some first results from a typological study into the diversity of content interrogatives. I am working towards a semantic map showing the possible links between ontological categories as established by the range of usages of content interrogatives among the world's languages. One of the many interesting findings is that the often claimed universal distinction between 'who' and 'what' (e.g. Weinreich 1963: 121-2, Ultan 1978: 228-9, Sadock & Zwicky 1985: 184-5) has more counterexamples than expected. Most languages indeed have an animacy distinction in their content interrogatives, though a few percent of the world's languages do not (many of which are found in South America).

It turns out to be rather difficult to find accurate descriptions of the usage of content interrogatives for many languages. The problem is that interrogatives will normally be mentioned in grammars or dictionaries, but most of the time only as translational equivalents of the English question words (or any other European language used by the author of the description). Such rough descriptions do not allow for any analysis of the finer grained differences between the range of the usage of individual interrogatives. For example, the question to the extent (e.g. English *how far*) is in most European languages based on the manner interrogative (i.e. English *how*). However, in many other languages the question word for quantity is used instead (e.g. Tagalog *gaano kalayo* 'how.much far', Schachter & Otones 1972:515).

To solve this issue of data availability, I am currently working on a more contextually based comparison of content interrogatives among the world's languages using parallel texts (e.g. the Bible). For example, there appears to be a linear semantic map as shown in (1) as a generalisation over the cross-linguistic diversity, with different languages using different interrogatives for part of the map. However, it turns out not to be possible to establish how languages deal with all these finer grained concepts on the basis of grammars or dictionaries alone.

Studying the cross-linguistic diversity of content interrogatives is not only of interest to typology, but also for a wider cognitive-philosophical debate about human ontological categories. For example, Heine *et al.* (1991:159) and Jackendoff (1983:52-3) partly base their proposals for general human cognitive principles on an analysis of content interrogatives, arguing that interrogatives show the basic categorisation of human thinking. Unfortunately, their analysis of content interrogatives does not stand the test of the typological diversity. I even doubt whether there are any strictly universals to be discerned in content interrogatives, though there are very many interesting (statistical) typological generalisations.

Examples

(1) THING - INSTRUMENT - MEANS - MANNER - EXTENT - QUANTITY

THING:	<i>what is that?</i>
INSTRUMENT:	<i>with what have you done it?</i>
MEANS:	<i>how have you achieved it?</i>
MANNER:	<i>how did he perform?</i>
EXTENT:	<i>how large is it?</i>
QUANTITY:	<i>how much is it?</i>

References

- Heine, Bernd, Claudi, Ulrike, and Hünemeyer, Friederike. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, Mass.: MIT Press.
- Sadock, Jerrold M., and Zwicky, Arnold M. 1985. Speech act distinctions in syntax. In *Language Typology and Syntactic Description*, ed. Timothy Shopen, 155-196. Cambridge: Cambridge University Press.
- Schachter, P., and Otnes, F. T. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.
- Ulan, Russell. 1978. Some general characteristics of interrogative systems. In *Universals of Human Language*, ed. Joseph H. Greenberg, 211-248. Stanford: Stanford University Press.
- Weinreich, Uriel. 1963. On the semantic structure of language. In *Universals of Language*, ed. Joseph H. Greenberg, 114-171. Cambridge, Mass.: MIT Press.

Measuring Typological Distance in the World Atlas of Linguistic Structures Östen Dahl

The advent of the World Atlas of Linguistic Structures (WALS) makes the application of quantitative methods to typology significantly easier, although it must be kept in mind that the data contained in the WALS database are still limited and patchy in many ways. I have tried to establish a global measure of the typological distance between the languages represented in the WALS material. The crudest way of doing so is by just counting the WALS parameters (one for each of 140 maps) that are different for each pair of languages. I have refined this measure by treating parameters as numeric whenever possible and reanalyzing multivalued features into binary when that makes sense. There is great variation in the WALS database with respect to the number of parameters for which there is information for each language. This means that the typological distance measure is most reliable for those languages which have the number of shared parameters.

One way in which the typological distance measure is useful is for sampling of languages. Earlier sampling methods have in general tried to reduce areal and genetic bias by choosing languages evenly over areas and families. These methods are problematic in that they presuppose assumptions about areal and genetic classifications of languages and of the impact of areal and genetic affinities. With the WALS database, it is possible to maximize the typological diversity in a sample on the basis of actual typological distances between languages. This can be done by removing from an initial sample one member of every pair of languages whose distance does not reach a certain threshold. The threshold can be manipulated to obtain the desired sample size. The number of languages from a certain family or area retained in such a sample is also a measure of its internal diversity. A preliminary comparison with earlier sampling methods such as that of Rijkhoff et al. suggests that the diversity of areas such as the Americas and phyla such as Nilo-Saharan has been underestimated. The method suggested is however limited by the quality of the initial sample, and oversampling is generally easier to detect than undersampling.

I have also compared the typological distance measure with earlier attempts to measure the lexical distance between languages in the sense of number of retained cognates in pairs of genetically related languages, e.g. in the work by Dyen et al. In a sample of 13 Indo-European languages, I obtained correlations as high as 0.8 between typological distance and lexical distance as measured by the scholars mentioned. This gives good hope for the usefulness of the kind of distance measure between language proposed here.

The Category of Destinative in Nganasan, North-Samoyedic and Typology of Prospective Possession

Michael Daniel

Nganasan and other North-Samoyedic (Uralic) feature a typologically rare category of destinative, a means used to mark ‘prospective possession’ relation between the prospective possessor (recipient or beneficiary of creation verbs) and the prospective possessum, the object being transferred or created. The destinative marker is attached to the stem naming the transferee and is typically followed by a regular possessive suffix; cf. ex. 1.

In this usage, the destinative marker introduces a position for possessive marking on the transferee, also specifying the prospective aspect of the possessive relation. However, destinative may also be used without the possessive suffix. This is obligatory when the prospective possessor is expressed by a nominal dependent on the possessum (ex. 2); and is also possible when the prospective possessor is specified by the general context (ex. 3). These morphosyntactic contexts of destinative in ex. 1 through 3 parallel the main types of the Samoyedic possessive construction. The regular possessor is also expressed either by a full nominal or by a possessive marker; co-occurrence of the nominal and pronominal possessor is also ungrammatical.

The important point is that in North-Samoyedic the destinative construction is the main morphosyntactic pattern of the ‘give’ verbs. There is no case form for the recipient (the European dative); not even a separate argument position – in most cases the prospective possessor is expressed together with the prospective possessum, i.e. as a possessive marker on the possessum, as in ex. 1, or as its nominal dependent, as in ex. 2. (An alternative construction, similar to the European pattern, has limited usage and is interpreted by some speakers as conveying prospective temporary possession meaning; cf. ex. 4).

This suggests an interesting perspective of the typology of ditransitivity in terms of internal vs. external possession. When an action affects an object belonging to another person (possessor), many languages tend to express this possessor separately from the possessum, providing the possessor with a separate argument position to underline his / her secondary affection by the action. This is called external possession construction; cf. [Payne and Barshi 1999]. The fact that the basic ‘give’ scenario in North-Samoyedic is bivalent, with the recipient morphosyntactically dependent on the transferee, suggests that the crosslinguistically dominating ditransitive pattern may be considered as being in the same relation to the destinative construction as the external possessor construction to the internal possessor construction. Within this approach, the destinative construction may be called internal prospective possessor construction. Cf. The following scheme (where * marks constructions ungrammatical in English but available in other languages)

external possession <i>*I broke him his gun</i>	<->	internal possession I broke his gun
trivalent ‘give’ scenario I gave him bread	<->	destinative construction <i>*I gave his (prospective) bread</i>

Madurese and a Typology of Raising

William Davies

This paper focuses on the grammatical properties of a Madurese structure in which an argument of a complement clause appears to occur in a non-thematic position in its dominating clause (1). Raising-to-Object (or its analogue in non-derivational theories) has been proposed over the past 30 years or so for the corresponding construction in the closely related Austronesian languages of Balinese (Wechsler & Arka 1998), Indonesian/Malay (Chung 1976, Kana 1986) and Javanese (Davies 1990). However, I argue here that a Raising analysis is not supported in Madurese. The Madurese construction includes a number of properties that distinguish it from classic Raising constructions:

- i. A pronoun coindexed with the ‘raised’ NP can occur in the complement clause (2).
- ii. The ‘raised’ and ‘unraised’ structures are not cognitively synonymous. Thus, (3a) is a proof about an event, whereas (3b) is a proof about the NP Hasan.
- iii. The ‘raised’ NP can occur as a prepositional object (4).
- iv. Unlike standard cases of raising, the raised NP need not be the embedded subject, but in fact may bear any grammatical relation in the embedded clause, including possessor of an object (5).
- v. Idiomatic expressions cannot occur in the Madurese construction. Thus, (6b), in which *nase* ‘rice’ occurs in the matrix clause, cannot have the idiomatic interpretation available to it in (6a).
- vi. The Madurese construction is not sensitive to syntactic islands. In (7), the matrix NP *Bambang* appears to have raised from an coordinate structure in the embedded clause, leaving a pronominal copy.

No one of these characteristics is sufficient to rule out a Raising analysis; most have been claimed to occur in constructions described as Raising in the literature. However, taken together with the fact that virtually any non-control predicate that takes a complement can occur in this construction, these characteristics support an analysis in which the matrix NP is base-generated as an argument of the matrix clause and is coindexed with a null or overt pronoun in the embedded clause, as in (8). Thus, the Madurese data reveal that a proleptic NP analysis proves superior to the raising analysis and shares virtually all the same properties as the parallel English construction (*I believe about Herman that he is capable of outrageous behavior*). Enumeration of these properties and comparison with both Raising and Copy Raising constitute the initial step in identifying the hallmarks of each construction and how they might differ typologically. Finally, I suggest that some constructions in other languages previously identified as instances of Raising should perhaps be reassessed.

Data

- (1) Siti ngera Hasan [bari' melle motor].
S AV. think H yesterday AV.buy car
'Yesterday Siti thought Hasan bought a car.'
- (2) Hasan e-kera Siti bari' [(ja') aba'engi melle motor].
H OV-think S yesterday COMP he AV.buy car
'Siti thought yesterday (that) Hasan bought a car.'
'Hasan was thought by Siti yesterday to have bought a car.'
- (3) a. Ita a-bukteagi ja' Hasan ngeco' sapedha motor.
I AV-prove COMP H AV.steal motorcycle
'Ita proved that Hasan stole the motorcycle.'
b. Ita a-bukteagi Hasan ja' ngeco' sapedha motor.
I AV-prove H COMP AV.steal motorcycle
'Ita proved Hasan stole the motorcycle.'
- (4) Siti ngera parkara Hasan ja' e-pareksa dokter juwa
Siti AV.think about H COMP OV-examine doctor that
'Siti thinks about Hasan that that doctor examined him.'
- (5) Siti ngera Hasan ja' dokter juwa mareksa ana'-eng.
S AV.think H COMP doctor that AV.examine child-DEF
'Siti thinks that the doctor examined Hasan's child.'
- (6) a. Siti ngera bari' ja' nase' la daddi tajjin.
S AV.think yesterday COMP rice already become porridge
'Siti thought yesterday that it is too late to do anything about it.'
lit. 'Siti thought yesterday that the rice had become porridge.'
b. Siti ngera nase' bari' ja' la daddi tajjin.
S AV.think rice yesterday COMP already become porridge
'Siti thought about the rice yesterday that it had become porridge.'
- (7) Hasan ngera Bambang; ja' Marlana ngerem paket dha' Alibi' aba'engi.
H AV.think B COMP M AV.send package to A and him
'Hasan thought about Bambang that Marlana sent a package to Ali and him.'
- (8) Siti ngera Hasan; bari' [pro; melle motor]

Turkic Forms in *-*gaj* :
an Essay on Evolution of a Modal Category
N. Dobrushina

Crosslinguistic investigation of combinations of different categories in formally identical markers or constructions is a generally adopted instrument to conclude what meanings are semantically related to each other, to hypothesize what are typical diachronic paths of semantic development of morphological categories and, after having accumulated enough data, to build a semantic map.

In this talk, I use a different approach. I trace the development of one morphological marker in as many as possible Turkic languages (presently data of varying reliability is available for 19 languages of the family). Proceeding from the same original source (future participle in *-γaj/-qaj*, sometimes called optative [Baskakov & al. 1988: 330-333]), this marker, attested in most contemporary Turkic languages, is reported to express a wide range of modal meanings. Combination of meanings in one marker are thus not considered in a given language but in a group of related languages, the identity of the markers being assumed on the basis of diachronic evidence. The study leads to establishing meaning connections that lie in the domain of diachrony and are not present at the synchronic level.

Modern Turkic forms going back to the *-γaj/-qaj* have four most frequent meanings.

1. The speaker's wish that a certain situation P took place (**optative proper**): Karaim, Karakalpak, Kazakh, ex. 1.
2. The category of mild wish (manifested in Balkar, Kumyk, Mishar Tatar, Azerbaijani and other) is different from the optative proper in (1) not in that the speaker's desire is less strong but in that it is expressed more mildly (ex. 2 and 3). As compared to the optative proper, this variety includes an additional epistemic component 'the speaker wants that P took place but is not sure it will'.
3. In many Turkic languages the *-*γaj/-qaj* marker is used when the speaker considers realization of P as probable (Yakut, Shor, Altai, Khakas; ex. 4 and 5).
4. In North-Eastern Turkic, the form under investigation is used to convey the speaker's consent that a P take place or his permission to carry out a certain action (Khakas, Shor, Altai, Black Tatar; ex. 6, 7, 8, 9).

Next, we isolate two areal (and group-level genetic) clusters of these four elementary meanings of the *-γaj/-qaj* form across Turkic languages. In the first cluster, covering South-Western, North-Western and South-Eastern Turkic, the form expresses various types of the speaker's wish. In the second cluster, consisting of North-Eastern Turkic only, the form conveys the speaker's consent, permission, or epistemic possibility.

Note that, at the level of individual languages, the form never combines (strong) wish (optative proper) with mild wish, with epistemic possibility, or with consent/permission (deontic possibility), making this connection invisible for the traditional crosslinguistic approach. On the other hand, epistemic possibility and consent/permission combine in many North-Eastern Turkic languages. Some of the Turkic languages also combine epistemic possibility with mild wish.

In addition to the four most frequent meanings listed above, the form may also mark the following categories: future (Chulym-Tatar), necessity (Tuva), apprehension (Yakut, Tatar), purpose (Uygur, Karaim, Azerbaijani, Gagauz, Balkar), protasis

(Gagauz), and sentential actants with some verbs (Azerbaijani). In Tuva the marker conveys deontic and epistemic necessity but has none of the four more frequent functions listed above; for Tuva, connection between necessity and other categories may only be traced diachronically.

The Turkic evidence for the connection between various modal meanings will be compared with the conclusions made on the basis of previous typological studies (Bybee, Perkins & Pagliuca 1994; van der Auwera & Plungian 1998).

- (1) **Karaim** (Musajev 1964: 287)
bur-gej-s *baš-yj-ny da* *kör-m'a-gej-s* *jaxšyj-ny!*
 bend-GAJ-2Sg head-2Sg-Acc and see-Neg-GAJ-2Sg good-Acc
 I wish you'd wring your neck and see no good!
- (2) **Balkar** (Mundart of Verkhnjaja Balkaria, personal field data)
zawun *zaw-γγ* *edi*
 rain rain-GAJ AUX
 I wish it would rain
- (3) **Kumyk** (Gadžiakħmedov 2000: 246)
γali *sama* *biraz* *juxla-γaj* *edi*
 now only a.little sleep-GAJ AUX
 It would be good if he'd sleep a bit now
- (4) **Yakut** (Korkina 1970: 235)
manan *buollaħyna, any* *tüün biħigi muus-put* *xams-aaraj*
 this.INSTR if now night we ice-POSS.1PL move-GAJ
 If it goes on like that, it looks like (our) ice will break this night
- (5) **Khakas** (Sagaj dialect; personal field data)
aya: *pre:* *xyryx-š'e* *š'as* *pol-γaj*
 he.DAT about forty-PROL yearbe-GAJ
 He is about forty
- (6) **Khakas** (Baskakov 1975: 198)
Pis *tee* *pol-ys-xaj-bys*
 we well be-Rec-GAJ-1Pl
 Ok, we'll help (if nobody else can)
- (7) **Shor** (Dyrenkova 1941: 176)
andig *pol-gaj*
 so be-GAJ
 Ok, let it be so
- (8) **Black Tatar** (Baskakov 1966: 83)
d'je, *kožo* *bar-gaj-zaar*
 well together go-GAJ-2Pl
 Ok, come with me
- (9) **Altaj** (Dyrenkova 1940: 165)
d'je, *men* *ajt-kaj-yn*
 well I say-GAJ-1Sg
 Ok, I'll do the talking

References:

- Baskakov A.N. 1966. Dialekt chernevykh tatar (tuba-kiži). [The dialect of the Black Tatars (Tuba-Kiži)]. M.
- Baskakov A.N., ed. Grammatika khkasskogo jazyka. [Grammar of Khakas]. 1975. M.
- Baskakov N.A. & al., eds. 1988. Sravnitelno-istoricheskaja grammatika tjurkskikh jazykov [A comparative historical grammar of the Turkic languages]. M.
- Bybee, Joan & Revere Perkins & William Pagliuca. 1994. The evolution of grammar: Tense, aspect and modality in the languages of the world. University of Chicago Press.
- Dyrenkova N.P. 1940. Grammatika ojrotskogo jazyka. [Grammar of Oyrot]. M.-L.
- Dyrenkova N.P. 1941. Grammatika shorskogo jazyka [Grammar of Shor]. M.-L.
- Gadziakhmedov N.E. 2000. Slovoizmenitelnye kategorii imeni i glagola v kumyjskom jazyke. [Verbal and nominal grammatical categories in Kumyk]. Makhachkala.
- Korkina E.I. 1970. Naklonenija glagola v jakutskom jazyke [Verbal moods in Yakut]. M., Nauka.
- Musajev K.M. 1964. Grammatika karaimskogo jazyka [Grammar of Karaim]. M.
- Sovremennyj kazakhskij jazyk. Fonetika i morfologija [Modern Kazakh. Phonetics and morphology]. 1962. Alma-Aty.
- van der Auwera, Johan & Vladimir Plungian. 1998. Modality's semantic map. *Linguistic typology*, Volume 2-1. Mouton de Gruyter.

Functions in Passives, Functions of Passives in Tukang Besi

Mark Donahue

Models of passives all involve the empirical observation that the subject of the active clause is demoted in a passive. The subject is recoded as an adjunct (or, in some languages, omitted altogether; rarely, it is an oblique), and the object (/internal argument) typically assumes subject ‘properties’.

In Tukang Besi, an Austronesian language of central Indonesia, the agentive argument of a passive is removed from the clause (though still implied), and the patient-like ‘object’ advances to acquire subject properties, as can be demonstrated through case marking, positional possibilities with respect to adjuncts, and (optionally) verbal agreement. This can be seen in (1) - (3).

- (1) [VP **No-tinti**] di ito **na** kalambe.
 3R-run OBL:PST there(up) NOM girl
 ‘The girls ran up there.’
- (2) [VP **no -’ita** te ‘obu] di ito **na** kalambe.
 3R-see CORE dog OBL:PST there(up) NOM girl
 ‘The girls saw the dogs up there.’
- (3) [VP **No-to-’ita**] di ito **na** ‘obu.
 3R-PASS-see OBL:PST there(up) NOM dog
 ‘The dogs were seen up there.’

The patient does not, however, change in terms of syntactic behaviour. The properties that are exclusively associated with subjects are not shared with the patient of a passive clause. This is illustrated with data from floating quantifiers, shown in (4) - (6), in which we can see that the argument of a passive clause is not able to be the restriction of a floated quantifier.

- (4) Saba’ane no-tinti na kalambe.
 all 3R-run NOM girl
 ‘All of the girls ran.’
- (5) Saba’ane no-’ita te ‘obu na kalambe.
 all 3R-see CORE dog NOM girl
 ‘All of the girls saw the dogs.’
 * ‘The girls saw all of the boys.’
- (6) * saba’ane no-to-’ita na ‘obu
 all 3R-PASS-see NOM dog
 For: ‘All of the dogs were seen.’

In many ways, then, this passive resembles what has been described as a ‘demoting’ passive, in which while the agent changes status, the patient does not. This absence of a nominal with subject properties in a clause is a challenge for models of language that require subjects, as has been firmly demonstrated by Dubinsky and Nzwanga (1994) and documented for other languages. In Tukang Besi data while there is no argument that is syntactically a subject, there is one that has all the morphological and *phrase structural* trappings of a subject.

Observing that the passive is used to end discussion of topics in discourse, we consider a semantic/discourse analysis of the passive.

References

- Donohue, Mark. 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.
- Dubinsky, Stanley, and Mazmba Nzwanga. 1994. A challenge to Burzio's generalization: Impersonal transitives in western Bantu. *Linguistics* 32: 47-64.
- Keenan, Edward L. 1985. Passive in the world's languages. In Timothy Shopen, ed, *Language Typology and syntactic description: Volume I, clause structure*, 243-281. Cambridge University Press
- Mohanan, Tara. 1995. *The notion logical object in Hindi*. MS, National University of Singapore.

Languages of Sumatra from a Geographical Perspective

Matthew Dryer

In this talk, I discuss the ways in which properties of the languages of Sumatra compare to those of languages of the surrounding geographical area. The similarities to surrounding languages reflect two factors: genealogical and areal. The languages of Sumatra resemble other languages of Indonesia because they are Austronesian. But there are also interesting similarities between languages of Sumatra (and languages of Indonesia in general, and often Austronesian languages in general) and languages of mainland southeast Asia. For example, the languages of mainland southeast Asia are VO (and more specifically SVO), and so are languages of Sumatra. On the other hand, a number of languages of Sumatra are verb-initial, which is not a common order in mainland southeast Asia, although it is found in Nicobarese (which is not Austronesian). While some languages of Sumatra are relatively isolating (e.g. Minangkabau), like languages in mainland southeast Asia, others (e.g. Nias) have more morphology. While languages of mainland southeast Asia are generally tone languages, languages of Sumatra (and elsewhere in Indonesia) are not.

Vacuous Appeals to the Number Hierarchy

Matthew Dryer

There are various linguistic phenomena which are claimed to reflect the number hierarchy given in (1) that are often described as involving markedness.

(1) Number Hierarchy: singular > plural > dual > trial I assume that these so-called markedness phenomena are all ultimately to be explained in terms of frequency: each number value on this hierarchy is more frequent than number values to the right of it. The Number Hierarchy is also claimed to be relevant to defining possible number systems. According to Greenberg 1963 (among others), if a language has a particular number value on this hierarchy, then it will also have the number values to the left of it on the hierarchy. In this paper, I argue that:

- a. In so far as the Number Hierarchy correctly describes possible number systems, this is purely coincidental, that the explanations for these principles governing possible number systems do not involve frequency.
- b. In some cases, the Number Hierarchy makes the wrong predictions regarding possible number systems.
- c. In those cases where the Number Hierarchy makes the wrong predictions, the explanations for these patterns ARE (ironically) due to frequency.

In short, in so far as frequency plays a role in explaining possible number systems, it has the opposite role from what is often assumed.

I consider various logically possible number systems and consider principles that would explain why such systems probably do not exist. Consider the fact that there are (apparently) no languages with a dual but no plural. What might a language look like that did have a dual but no plural? Assuming that there must be SOME way to refer to more than two of something, what would that form be like? If the form is distinct from both the singular and the dual, then the language would have a plural, so we must assume that is either identical to the singular or identical to the dual. If it is identical to the dual, then we would call the form that is used for 2 or for more than 2 a plural, so again this would not be a language with a dual but no plural. Hence it would have to have a form that was identical to the singular. What would seem to rule out such a possibility would be that a given form would have the meaning 'one or more than two', which is probably ruled out as an unlikely meaning. Note that none of the explanations for the nonexistence of the various types makes any reference to markedness or the relative frequency of referring to two compared to referring to more than two.

It is often claimed that a language cannot have a plural without having a singular. But it is difficult to reconcile this claim with languages which have a plural which is purely optional, where the nonplural form is obligatory when referring to one and optional when referring to more than one. It is tempting from an English perspective to call such a nonplural form a singular, but strictly speaking such a form is completely vague as to number, since its use does not imply anything about number. Hence such languages (of which there are many) are in fact instances of languages with a plural but no singular. Hence the number hierarchy does not accurately describe what is a possible number system.

Consider the mirror image of the language type described in the preceding paragraph. This would be a language with an optional affix which when added to a noun would mark it as singular, but where the bare noun stem is obligatory when referring to more than one and optional when referring to one. By the reasoning of the preceding

paragraph this would be a language with a singular but no plural. However, such a language does not appear to be possible, and it is ruled out by markedness principles: a zero form can only be used for the most common meaning. But note that the impossibility of this language shows that, again contrary to the Number Hierarchy, one cannot have a language with a singular but no plural. And the explanation for why there can be a language with a plural but no singular but not a language with a singular but no plural is that singular is more frequent (less marked) than plural.

While the conclusions of this paper are not inconsistent with the approach of Corbett (2000), his approach does not make clear the extent to which markedness has the opposite effect on possible number systems from that proposed by Greenberg, and his principles fail to account for number systems which have a plural but lack a singular.

The Typology of Voice in Gayo

Domenyk Eades

Some of the recent debate on the alignment systems of WAN languages has focused on whether these languages show a pattern of ergative alignment, or rather constitute a type of symmetrical system, where neither alignment can be considered the more basic or unmarked (cf. Himmelmann 2004). Gayo would appear to be a good candidate for a symmetrical voice system, as both Actor- and Undergoer-oriented predicates (marked by *mun-* and *i-* respectively) can take two arguments, and both orientations involve affixation to the verb. However, in contrast with more familiar voice systems such as English, Actor and Undergoer-orientation are employed relatively infrequently in coding events involving two participants. Rather, a range of other strategies involving various semantic affixes are employed to code such events, for example those that are uncontrolled or involve middle-type semantics. Furthermore, the Actor and Undergoer prefixes are not indicative of the number of arguments in the clause, functioning relatively independently of syntax. Thus, AO and UO constitute only a subset of the possible strategies by which events involving two participants are coded. Verbal affixation is partly determined by the semantics of the event in a given instance in which a verb is used. Accordingly, it is argued that the alignment system of Gayo is semantically motivated. This is in contrast with asymmetrical systems, whose grammatical patterning is syntactically-motivated, i.e. grammatical patterning is sensitive to the number of arguments in the clause regardless of the semantics of the event specified by the predicate (cf. Dixon 1994).

References

- Dixon, R.M.W. 1994, *Ergativity*, Cambridge, New York: Cambridge University Press.
- Himmelmann, Nikolaus, 2004, "The Austronesian languages of Asia and Madagascar: Typological characteristics", In K. Alexander Adelaar and P. Nikolaus Himmelmann, *The Austronesian languages of Asia and Madagascar*, 110-181. London: Routledge

Negation and the Typology of Non-Verbal Predication

Pål Kristian Eriksen

In a substantial number of languages the expression of negation is dependent on the part of speech of the predicate it negates. Bahasa Indonesian is a good example: Verbal and adjectival predicates⁷ are negated with the particle *tidak* and nominal predicates with the particle *bukan*:

Mereka tidak menolong kami
3.PL NEG help 1.PL.EXCL
“They didn’t help us.”

Dia bukan guru.
3.SG.FEM NEG teacher
“She is not a teacher.”

In the typological literature on non-verbal predication (e.g. Verhaar (ed.), Hengeveld and Stassen) it is pointed out as one of the main differences between verbal and non-verbal predication, but no explanation is given to the phenomenon, and its typological distribution has not been mapped.

I will argue that divergent non-verbal negation forms like the Indonesian *bukan* are instances of a phenomenon which I will refer to as **direct negation avoidance**. Direct negation avoidance implies that predicates of a non-verbal part of speech cannot be the main object of the negation’s scope. Non-verbal negation expressions can be attributed to different negation strategies, the common feature of which is that negation only indirectly affects the non-verbal predicate. These negation strategies may be divided into a set of formally and semantically well-defined types, and I will give a brief overview of this typology.

I will furthermore show that the typology of non-verbal negation strategies is interdependent with the typologies of related grammatical phenomena – e.g. the typology of non-verbal predication and the typology of negation as such:

The extension of the use of direct negation avoidance is itself determined by the typology of non-verbal predication. If a language has “verby” adjectives⁸, only nominal predicates require direct negation avoidance; if adjectives are “nouny”, direct negation avoidance extends to cover all non-verbal predicates.

The distribution of negation systems with a divergent non-verbal negation strategy as opposed to systems with a uniform negation strategy for all parts of speech, is dependent on the pragmatic and semantic markedness of the non-verbal negation strategies, in interaction with the predication typology mentioned under pt. 1.

In languages with “verby” adjectives the typology of direct negation avoidance strategies may be used to predict whether nominal predicates will require special modifications (like copulas) in positive clauses as well.

Finally I will make some short remarks in the direction of a semantic explanation to the direct negation avoidance universal.

⁷ I use the terms *verbal* and *adjectival* (and *verb* and *adjective*) to refer to universal semantic categories, in the sense that the concept “tall” will have the semantic features of an adjective in all languages, despite the formal features of the word class it is assigned to in the individual language.

⁸ I use the terms “*verby*” and “*nouny*” to refer to the formal characteristics which the above mentioned universal semantic categories obtain in a given language. If an adjective is “verby”, it shares (most of its) formal characteristics with verbs, whereas “nouny” adjectives are formally more similar to nouns.

References

- Hengeveld, Kees. 1992. *Non-Verbal Predication – Theory, Typology, Diachrony*. Functional Grammar Series 15. Mouton de Gruyter. Berlin / New York.
- Stassen, Leon. 1997. *Intransitive Predication*. Clarendon Press. Oxford.
- Verhaar, John W. M.. (ed.). 1967-1973. *The Verb “Be” and its Synonyms, vol. 1-6*. D. Reidel Publishing Company. Dordrecht.

Core and Extended Meanings of the Reciprocal: a Study in Semantic Typology

Nicholas Evans, A. Gaby, S. Levinson, A. Majid, S. Robinson, Ulrike Zeshan, N. Enfield, P. Brown, N. Burenhult, C. Wegener, M. Dunn & C. Rapold

It is easy to define the meaning of ‘core reciprocals’ like ‘John and Mary love each other’, along lines like those used by Nedjalkov (1991:276-7): ‘Strict reciprocals are those designating symmetrical relations between actants having distinct semantic roles: each of two symmetrical actants fulfils two symmetrical roles.’ [our translation]

However, if we try to extend the definition to other uses of reciprocal expressions, such as English *each other* (cf Langendoen 1978), we encounter a number of complexities, among others:

- can we discard simultaneity as a requirement? (*John and Mary massaged each other in front of the fire*)
- how do we deal with cases where the relation does not hold symmetrically between all members of the group? Consider ‘chained reciprocals’ (Lichtenberk 1985) like *The students followed each other onto the stage* or ‘pairwise reciprocals’ like *Everyone at the party was married to each other*)
- how do we deal with ‘asymmetric’ uses of reciprocals, where the event is not wholly symmetric? Consider the asymmetric interpretation (available for many though not all speakers) of English *The woman and the burglar chased each other down the street*

Can these all be subsumed under a single statement of meaning (Plank, forthcoming), or should they be treated as polysemically-related senses? Dalrymple et al (1998) propose a solution where a number of formally-characterised types are linked by the ‘strongest meaning hypothesis’, which predicts which meaning a reciprocal expression will have in a given context.

Clearly there is a role for typological investigation here, to determine whether all languages apply their reciprocal expressions across the same semantic range and, consequently, whether the semantic range of English *each other* should be taken as a cross-linguistically invariant category, or an arbitrary (albeit common) set of linked semantic extensions. This paper reports on a cross-linguistic study based on new field data obtained from a series of 64 video clips designed to obtain event descriptions of short scenes that permute the various variables discussed in the reciprocal literature (number of reciprocants, simultaneity vs sequentiality, degree of ‘reciprocity saturation’ across the members of a large group, and various types of asymmetry) across a number of different event types. Our sample pools data from fifteen languages, most of them previously uninvestigated, spanning five continents, seven spoken language families, and including material from two sign languages.

Despite considerable convergence in the treatment of reciprocals across the languages in our sample, we identify significant cross-linguistic differences in the treatment of some semantic subtypes, particularly those involving chaining and asymmetric situations, as well as the existence of constructional differences for coding some of the parameters mentioned above.

References

- Dalrymple, Mary, Kanazawa, Makoto, Kim, Yookyung, Mchombo, Sam and Peters, Stanley 1998, Reciprocal Expressions and the Concept of Reciprocity. *Linguistics & Philosophy* 21.2: 159-210.
- Langendoen, D. Terence. 1978. The Logic of Reciprocity. *Linguistic Inquiry* 9(2): 177-197.
- Lichtenberk, Frantisek. 1985. Multiple Uses of Reciprocal Constructions. *Australian Journal of Linguistics* 5, 19-41.
- Nedjalkov, Vladimir P. 1991. Tipologija vzaimnyx konstrukcij. In A. V. Bondarko (ed.) *Teorija funkcional'noj gramatiki: Zalogovost', Personal'nost'*, pp. 276-312. Leningrad: Nauka
- Plank, Frans. To appear. Thoughts on the origin, progress and pro status of reciprocal forms in Germanic, occasioned by those of Bavarian. MS.

The Iwaidja/Mawng Reciprocal: A Challenge to the Typology of Reciprocal Constructions

Nicholas Evans, Ruth Singer & B. Birch

An recent paper by König and Kokutani (in press) proposes an initial typology of reciprocal constructions, establishing the following four-way classification of strategies:

1. Verbal strategy: affixal, e.g. Swahili *V-an* (reciprocal suffix on verb)
2. Verbal strategy: deverbal, e.g. Japanese *au* 'meet', with reciprocal meaning when serialized.
3. Nominal strategy: pronominal, e.g. German *sich*, French *se*, Somali *is*
4. Nominal strategy: quantificational, e.g. English *each other*, Rus. *druga druga*

Consideration of reciprocal data from a wider range of languages, however, suggests the König and Kokutani typology is too simple, and that a comprehensive typology of reciprocal constructions will require considerable expansion from their framework. In this article we outline the special characteristics of reciprocal constructions in Iwaidja and Mawng, two closely-related Australian languages of the Iwaidjan family (non-Pama-Nyungan; Cobourg Region, Northern Territory).

In each of these languages, the reciprocal construction makes use of normal transitive coding taking one of the reciprocants as the subject and the other as object, followed by a connective plus a contrastive subject pronoun. 'They hit each other' of a man and a woman, is thus expressed as 'she-him-hits' and 'he.in.turn'. Examples are given in (1) for Iwaidja and (2) for Mawng. Note that one argument is referred to both in the pronominal prefix of the verb and in the contrastive pronoun which follows the conjunction. The conjunction followed by a contrastive pronoun resembles a sort of truncated second clause in these languages because the contrastive pronoun is usually used to emphasise who is the subject, after change of subject (although it occasionally encodes an object). The construction clearly originated from two linked clauses but evidence such as the placement of argument NPs, indicates that the construction is monoclausal. For example in (3) to (6) the second object of a ditransitive verb follows the conjunction plus contrastive pronoun. It is usual for all NPs to either precede the verb, or follow the contrast pronoun.

The fact that the construction is monoclausal raises some problems for the identification of subject in the reciprocal construction. Is the subject the participant marked as such in the pronominal prefix, or the one encoded by the contrastive pronoun? The 'pronominal strategy' in König and Kokutani's typology only covers the case where one of the two argument positions – the object in a reciprocal formed from a canonical transitive – is filled with a special reciprocal pronoun. In Mawng and Iwaidja, however, one of the participants is encoded twice - once by the bound pronominal on the verb, and once by the contrastive/reciprocating pronoun. We propose a new cell in the typology – the 'overlaid pronominal strategy' – to accommodate the Iwaidja and Mawng facts.

Reference:

König, Ekkehard & Shigehiru Kokutani. In press. Towards a typology of reciprocal constructions: focus on German and Japanese. *Linguistics*.

Systemic Typology – Systemic Optimality

Gertraud Fenk-Oczlon & August Fenk

Previous crosslinguistic studies by the authors showed

- that the mean length of simple declarative sentences (one proposition within one intonation unit) is restricted to a rather small range of 5 – 10 syllables
- a set of statistically significant crosslinguistic correlations indicating time-related constraints. For instance: the more phonemes per syllable, the fewer syllables per word and per clause
- the association of such metric properties with non-metric properties: High syllable complexity is associated with VO order, with a rather low number of cases and a tendency to fusional or isolating morphology.

Naturalness Theory and Optimality Theory determine naturalness or optimality (and “conflicts among universal well-formedness constraints”, Tesar & Smolensky 1996) on separate levels of language (naturalness or optimality in phonology, in morphology etc.) Our results, however, conform to the program of a systemic or holistic typology (von der Gabelentz 1901, Plank 1998) by linking phonological parameters of crosslinguistic variation with morphological and syntactical parameters. Language as a whole has to be “optimal” and efficient. It is a dynamical system that answers changes on level X with self-regulatory processes affecting level Y and Z. Our empirical results concern (cognitive, articulatory, time-related) constraints constraining the variability of language as a whole. In the present study, this correlational framework is being put to test from a diachronic perspective (**A**). A second touch-stone (**B**) is Hawaiian, i.e. a language that is assumed (e.g. in Colarusso 1983) to have long words and a lower speed in delivering information due to its small phonemic inventory size.

Ad A: A comparison of the Beowulf Prologue in Old English (OE) with its translation into Modern English (ME) shows an increase in syllable complexity (a mean of 2.63 phonemes in OE and of 2.88 in ME) and a concomitant decrease of syllables per word (from 1.83 to 1.22) and an increase of the total of words (from 269 to 394). This confirms some assumptions (of e.g. Lehmann 1978) about the changes from OE to ME. And it perfectly matches several synchronous results: Menzerath’s (1954) law as well as our negative crosslinguistic correlation between n of phonemes per syllable and n of syllables per word.

Ad B: The statistical evaluations of our test-sentences translated by a native speaker into Hawaiian showed the following mean values: A very low number of phonemes per syllable (1.84) together with a very high number of syllables per sentence (but with 9.55 still within our range of 5 – 10 syllables) and 2.04 syllables per word, which is rather high but much lower than claimed by several authors. (In a running text, the Hawaiian romance of Laieikawai, the mean word length turned out to be even shorter, 1.88 syllables, than in our test sentences.) Hawaiian has a lot of short function words, and also the content words are not very long. Thus, in Hawaiian (and other Polynesian languages too?) the size of clauses, i.e. of the “packages” to be handled by working memory, seems to fit the same constraints as in the languages of our original sample. But it might become necessary to specify one of our regularities as follows: Low syllable complexity tends to OV order only in agglutinative languages with rather long verbs, probably due to the general principle of putting short before long.

Results A and B support the search for “optimality”, efficiency and economy within the framework of a functional and systemic typology considering constraints of perception- and production mechanisms.

I Gave It Him

On the Rationale of the ‘Alternative Pronominal Double Object Construction’ in English and Other Germanic Languages

Volker Gast

It is a well known fact that English has two semantically near equivalent constructions to express ditransitive predications: the ‘prepositional object construction’, exemplified in (1a), and the ‘double object construction’ in (1b), where the Recipient precedes the Theme.

- (1) a. John gave a book to Mary.
b. John gave Mary a book.

Moreover, there is a third possibility to encode ditransitive predications in certain varieties of (British) English: In the ‘alternative double object construction’ (Siewierska & Hollmann to appear), the Theme precedes the (non-prepositional) Recipient. This pattern is basically restricted to predications with two pronominal objects:

- (2) So I taught it him.

The construction illustrated in (2) is much more widespread than is often assumed. In certain varieties and registers, it is even preferred vis-à-vis the ‘canonical’ double object pattern, albeit still much less frequent than the prepositional object construction. For example, THM > REC order is generally more common in fiction (cf. Biber et al. 1999: 929), and Siewierska & Hollmann (to appear) have shown that in the Lancashire dialect of England it seems to be more common in spoken language too. In the British National Corpus, THM > REC prevails over REC > THM when the Theme is pronominal *it*.

Given the observations made above, a number of questions arise: What are the factors favouring either the canonical or the alternative double object construction? How have the two constructions developed historically? How do they relate to comparable patterns in languages other than English? And finally: How could the two constructions be ‘motivated’ from a functionalist perspective?

A first point that needs to be kept in mind when considering these questions is that most varieties of Modern English obey (a specific version of) the ‘ditransitive person role constraint’ (DPRC; cf. Haspelmath to appear). One of the restrictions resulting from the DPRC is that the double object construction is impossible when the Recipient is third person, and the Theme is either first or second person (**He gave it_{REC} me_{THM}*). In combinations with third person pronouns and SAP pronouns, case ambiguities do therefore generally not arise. This is reminiscent of direct/inverse case marking systems, where the directionality of a transitive predicate is a function of the relative degree of animacy of the participants.

Second, it is important to note that the ‘alternative double object construction’ was the default option in Middle English if both objects were pronominal, whereas the canonical order of non-pronominal objects was REC > THM. This seems to point in the following direction: (Early) Middle English allowed both constructions and was thus susceptible to principles of discourse organization and information packaging. The various varieties of Modern English have developed in one of three ways: (a) they have grammaticalized the preferred word order pattern of Middle English in the pronominal double object construction (THM > REC), against the pressure of paradigm uniformity coming from the canonical pattern (REC > THM); (b) they have generalized the word order of the non-pronominal double object construction to the pronominal double object construction (REC > THM), or (c) they have given up the pronominal double object construction altogether.

The assumption that the ‘alternative pronominal double object construction’ reflects an earlier stage of English does, of course, not answer the question of why it exists in the first place. Why should THM > REC order have been preferred when both objects were pronominal? This question can be put in more general terms if we take into account that similar asymmetries can be observed in other languages too. For example, German has the preferred order REC > THM when both arguments are non-pronominal, while THM > REC is more common with pronominal objects:

- (1) a. Er gab einem Bettler einen Euro.
 he gave a beggar (DAT) a Euro (ACC)
 ‘He gave a Euro to a beggar.’
 b. ??Er gab einen Euro einem Bettler.
 he gave a Euro (ACC) a beggar (DAT)
- (2) a. Ich gab es ihm gestern.
 I gave it (ACC) him (DAT) yesterday
 ‘I gave it (to) him yesterday.’
 b. ??Ich gab ihm es gestern.
 I gave him (DAT) it (ACC) yesterday

From a cross-linguistic perspective, it is hard to generalize over the relative order of Recipients and Themes. Primus (1998) has proposed that the order of REC and THM is determined by two competing principles: on the one hand, REC/THM linearization tends to be aligned with the Thematic Hierarchy AGT > REC > THM; on the other hand, it tends to align with the Case Hierarchy NOM > ACC > DAT, so there is a conflict. However, this proposal fails to explain the asymmetry pointed out above. One of the problems is that Modern English does not have a morphological case distinction between accusative and dative, so the Case Hierarchy should be irrelevant, and we would expect to find only the REC > THM pattern. It should also be mentioned that THM > REC order appears to be unexpected under the hypothesis that Themes are typically focal, and that focal material comes last in the sentence.

I will propose an explanation that is based on observations concerning frequency distributions of pronoun combinations in historical corpora of English. Given that monotransitive predicates outnumber ditransitive ones by far, the pattern [V THM_{PRO/3}] was much more frequent than [V REC_{PRO/3}] in Middle English, irrespective of the order of the two object NPs in ditransitive predications. While [V THM REC] and [V REC THM] were both, in principle, possible, the ‘component pattern’ [V THM_{PRO/3}] was much more frequent than [V REC] because the former occurred in monotransitive predications too. As a consequence, this pattern was ‘entrenched’ and grammaticalized – in the relevant varieties of English – and was generalized to the double object construction. This explanation can also account for the fact that THM > REC order is not commonly found when the direct object is first or second person (cf. the DPRC), or when both object NPs are lexical, since SAP-pronouns and lexical NP tokens do not occur frequently enough in an object position to trigger grammaticalization, so the ‘canonical’ case pattern REC > THM was preserved.

References

- Haspelmath, M. (to appear). 'Explaining the ditransitive person-role constraint: a usage-based account'. To appear in *Constructions*.
- Primus, B. (1998). 'The relative order of recipient and patient in the languages of Europe'. In Siewierska, A. (ed.), *Constituent Order in the Languages of Europe*, 421-73. Berlin: Mouton de Gruyter.
- Siewierska, A. & W. Hollmann (to appear). 'Ditransitive clauses in English with special reference to Lancashire dialect'. To appear in Hannay, M. (ed.), *The English Clause: Structure and Usage*. Amsterdam: Benjamins.

Towards a Typology of Focus Quantifiers

Volker Gast

In using the term FOCUS QUANTIFIER I refer to a family of expressions that quantify over alternative values contrasting with the focus of a sentence, in relation to the information given in the background of that sentence. Typical instances of focus quantifiers are particles like English *only*, *also* and *even* ('focus particles'). For example, *Only [FRED]F snores* expresses that the property of snoring (BACKGROUND) can be attributed to no one other than Fred (FOCUS). In addition to focus particles, the class of focus quantifiers contains elements with different morphological properties, but with a similar semantics (e.g. phrasal and affixal elements; cf. below).

Focus quantifiers vary along four major dimensions: (i) their MEANING, (ii) their MORPHOLOGY, (iii) their SYNTAX, and (iv) COMPATIBILITY REQUIREMENTS of various types, imposed on either the context or the associated focus:

I. Meaning

The meaning of focus quantifiers can be described with reference to two major parameters. First, there are two basic types of 'quantification': While *only* is EXCLUSIVE (*Only John snores* → *No one other than Fred snores*), *also/too* and *even* are ADDITIVE (*John snores, too* → *There is someone other than John who snores*). And second, some focus quantifiers interact with unstructured sets of alternatives (NON-SCALAR focus quantifiers, e.g. *too*), while others make reference to scales (SCALAR or EVALUATIVE focus quantifiers, e.g. *even*). For example, *Peter snores, too* merely presupposes that there is someone other than Peter who snores, while *Even Peter snores* additionally suggests that Peter is particularly unlikely to snore.

II. Morphology

In addition to focus particles or LEXICAL FOCUS QUANTIFIERS like *only* or *even*, there are also AFFIXAL FOCUS QUANTIFIERS (e.g. Afaan Oromo *-s* 'too', Finnish *-kin* 'too', cf. (1)) and POLYMORPHEMIC FOCUS QUANTIFIERS (e.g. Engl. *let alone*). Polymorphemic focus quantifiers are either CONTINUOUS (*let alone*, *at least*) or DISCONTINUOUS (e.g. Modern Breton *ne ... nemet* 'only', Hebrew *lo ... éla* 'only', cf. (2)).

III. Syntax

Focus quantifiers may stand in various structural relationships to the corresponding foci. There are three major types of configurations: (i) the focus quantifiers exhibit a considerable degree of positional variability but must c-command the focus (e.g. Engl. *only*); (ii) the focus quantifiers are adjacent to the focus (Turkish *de* 'too', cf. (3)); or (iii) the focus quantifiers occupy a fixed position in the clause (e.g. the pre-verbal position in Mandarin Chinese, cf. (4)).

IV. Compatibility requirements

Some focus quantifiers impose compatibility requirements on their environment, for example insofar as they interact only with foci of specific syntactic categories (e.g. German *selbst* 'even' co-occurs only with clausal and nominal foci, cf. (5); Hebrew *aH* only focuses on verbs and adjectives). Moreover, some focus quantifiers are sensitive to the polarity of a sentence (e.g. Engl. *too* vs. *either*, Spanish *tambien* 'too' vs. *tampoco* 'either', cf. (6)).

In this paper I will present a typology of focus quantifiers based on the four major parameters of classification pointed out above, and I will aim to determine the relevant patterns and limits of variation found in the languages of the world. A number of hypotheses concerning correlations between the various parameters will be explored, for example:

- (a) Discontinuous focus quantifiers always seem to be exclusive (e.g. Hebrew *lo ... éla* ‘only’).
- (b) Affixal focus quantifiers are typically additive (e.g. Finnish *-kin*), and if they are exclusive, they are always scalar (e.g. Chalcatongo Mixtec *-ni* ‘only, merely’).
- (c) Exclusive focus quantifiers tend to precede their focus (e.g. *only*), while additive focus quantifiers often follow (*too*).
- (d) Only additive focus quantifiers may be polarity-sensitive (e.g. *too* vs. *either*).

While some of the correlations observed can probably be accounted for in terms of synchronic (functional) explanations, others should better be explained with reference to the historical development of the relevant markers (e.g. the lexical source).

Data

I. Morphology

AFFIXAL FOCUS QUANTIFIERS

(1) Finnish *-kin* ‘too’

- (a) *minä-kin olen hankkinut auto-n.*
I-too I.have got car-ACC
‘I, too, have got a car.’
- (b) *olen hankkinut auto-n-kin.*
I.have got car-ACC-too
‘I have got a CAR, too.’

DISCONTINUOUS FOCUS QUANTIFIERS

(2) Hebrew *lo ... éla* ‘only’

- hem lo hisigu éla heskem xelki.*
they not reached but agreement partial
‘They only reached a partial agreement.’

II. Syntax

FOCUS QUANTIFIER IS ADJACENT TO THE FOCUS

(3) Turkish *de* ‘too’

- (a) *Oraya [ben de] gittim.*
there I too went
‘I went there too’
- (b) *Ben [oraya da] gittim.*
I there too went
‘I went there too (as well as elsewhere)’
- (c) *Ben oraya [gittim de].*
I there went too
‘I also went there.’

FOCUS QUANTIFIER OCCUPIES A FIXED POSITION RELATIVE TO THE FOCUS

(4) Mandarin

- (a) La□o Wáng zh□ he chá.
Lao Wang only drinks tea

‘Lao Wang drinks only [tea]F / Lao Wang only [drinks tea]F / Lao Wang only [drinks]F tea.’

- (b) Wo□ ye□ xie□ shu.
I too write book

‘I also [write a book]F / I also [write]F a book / I also write a [book]F.’

IV. Compatibility requirements

RESTRICTIONS ON CATEGORY OF FOCUS

(5) German *sogar* vs. *selbst*

- Ich habe ihn sogar/*selbst gekannt.
I have him even known

‘I even knew him.’

POLARITY-SENSITIVITY

(6) Spanish *tambien* vs. *tampoco*

- (a) Yo lo conozco *tambien*.
I him know too

‘I know him, too.’

- (b) No lo conozco *tampoco*.
not him I.know either

‘I don’t know him either.’

How Exceptional Is Riau Indonesian?

David Gil

In a series of publications, David Gil has argued that Riau Indonesian exhibits a number of syntactic and semantic features that characterize it as typologically exceptional, seemingly unlike most other natural languages, at least as commonly described. The question arises whether Riau Indonesian is truly exceptional, or whether its apparently exceptional properties are a mere artefact of Gil's particular descriptive approach. In order to answer this question, it is necessary to compare Riau Indonesian to other languages through the same eyes, using the same objective and rigorous yardsticks. This paper proposes one such yardstick, in the form of a psycholinguistic experiment designed to elicit truth-value judgments in different languages. The results of the experiment provide an answer to the question posed, suggesting that although Riau Indonesian is off in one corner of typological space, there exist other languages similar to Riau Indonesian and perhaps even more exceptional with respect to the properties in question.

At the heart of Gil's semantic analysis of Riau Indonesian is the claim that when two expressions X and Y with meanings P and Q respectively are combined, the meaning of the collocation X Y is derived from that of its constituent parts by means of the association operator, A (P, Q), which says that the meaning of X Y is associated in an unspecified way with the meanings of X and Y respectively. For example, if *ayam* means 'chicken' and *makan* means 'eat', *ayam makan* means A (CHICKEN, EAT), or anything that has to do in some way with 'chicken' and with 'eat'. In particular, the semantic representation A (CHICKEN, EAT) lacks any specification of thematic roles: the chicken could assume the role of agent, patient, or whatever might make sense in the context of the utterance.

This paper presents the results of an experiment designed to measure, objectively across a variety of languages, the availability of *apparently associational interpretations*: interpretations that appear to be obtainable from the association operator without reference to thematic roles or other semantic categories. The experiment is specifically designed for isolating languages with apparent basic SVO word order: this characterization encompasses Riau Indonesian, other Malayic languages, other languages of Southeast Asia and of West Africa, and also creole languages. Two kinds of apparently associational interpretations are sought: (a) those in which what looks like a bare noun preceding a bare verb is interpreted as the patient (rather than the agent); and (b) those in which what looks like a bare noun in construction with a bare verb is interpreted as an oblique argument or even a non-argument (in the absence of prepositions or other such markings). The experiment presents subjects with a sentence in the target language and two pictures; subjects are asked which of the two pictures is best described by the sentence. The experiment is ongoing; as of December 2004, over 500 subjects in 12 languages had been tested. Some preliminary results are summarized in Table 1 below, with languages arranged in roughly decreasing availability of apparently associational interpretations. (More data will be available by July 2005.)

Table 1 contrasts Riau Indonesian with three other Malayic languages, two other isolating SVO languages, and two other non-isolating languages. In addition, within each language, Table 1 contrasts subjects of ordinary vs. high socio-economic status. The reason for doing so is clear: for all isolating SVO languages for which such data is available, a consistent and significant difference is in evidence — the availability of apparently associational interpretations is higher for subjects of ordinary than for high

socio-economic status. Given that linguists often tend to seek out speakers of higher socio-economic status, this suggests that our overall view of what language is like may suffer from a systematic sociolinguistic bias. Moreover, it is precisely this bias that makes languages such as Riau Indonesian appear more exceptional than they really are.

Cross-linguistically, the results point towards the following conclusions: Non-isolating languages have near-zero availability of apparently associational interpretations. In contrast, isolating SVO languages generally allow apparently associational interpretations to some extent, thereby setting such languages apart from most others. However, amongst themselves, isolating SVO languages exhibit substantial cross-linguistic variation with respect to the availability of apparently associational interpretations; indeed, such variation is strikingly evident within closely-related Malayic languages. In this regard, the position of Riau Indonesian amongst isolating SVO languages is not exceptional: it is roughly comparable to Vietnamese, and in the mid range of Malayic languages. In fact, the truly exceptional language, with the highest availability of apparently associational interpretations, would seem to be the language of the ALT VI venue, Minangkabau.

<i>Language</i>		<i>Availability (%) of Apparently Associational Interpretations</i>	
		<i>Ordinary Socio-Economic Status</i>	<i>High Socio-Economic Status</i>
<i>Malayic</i>	<i>Minangkabau</i>	66	39
<i>Malayic</i>	<i>Jakarta Indonesian</i>	56	22
<i>Malayic</i>	<i>Riau Indonesian</i>	43	31
Isolating SVO	Vietnamese	41	23
Isolating SVO	Bislama	34	13
<i>Malayic</i>	<i>Kuala Lumpur Malay</i>	25	9
Non-Isolating	Hebrew	3	7
Non-Isolating	English	5	3

Table 1

The Sociolinguistic Landscape of Sumatra

David Gil

This paper presents a general overview of the sociolinguistic landscape of Sumatra. Its primary goal is to provide ALT participants with a sense of the languages that they will encounter here in Padang and elsewhere across Sumatra, and of how these languages are used. However, by using Sumatra as a case study, a more general argument is also put forward, namely that in order to do serious linguistic typology, it is necessary to pay greater attention to a variety of sociolinguistic factors.

Taking as a starting point the geographical perspective of Sumatra offered by Matthew Dryer, it is argued that any two-dimensional linguistic map of Sumatra (such as, for example, those available from the *World Atlas of Language Structures*) needs to be supplemented by at least one additional dimension, representing sociolinguistic space. Typically, in any given location in Sumatra, one encounters, side-by-side, language varieties belonging to at least the following four general sociolinguistic kinds: (a) a minor and often endangered regional language and/or a local dialect of a major regional language; (b) a standardized variety of a major regional language; (c) a local variety of colloquial Indonesian; and (d) a local variety of standard Indonesian. In this talk, these four categories are illustrated with data from two locations in Sumatra: the city of Padang and a village in the province of Riau.

Cross-linguistic and areal studies which ignore the sociolinguistic dimension risk comparing apples to oranges. For example, whereas standard Indonesian has a dedicated conjunction meaning 'and', colloquial varieties of Indonesian spoken in Sumatra, as well as many of the local languages of Sumatra, have a single macrofunctional word whose meanings include, among others, both 'and' and 'with'. Examining the use of conjunctions in a multiglossic and multilingual corpus of short text messages on mobile phones in Sumatra suggests that the presence of a dedicated word meaning 'and' correlates with the written medium, thereby accounting for its presence in standard Indonesian alongside its absence in many or most of the languages of Sumatra. In other words, the reason standard Indonesian has an 'and' word is not because of its overall typological profile, nor is it because of language contact and areal factors; rather, it is because of the sociolinguistic circumstances associated with its status as a major national language.

On Weak Government: Parenthetical Verbs in English, French, Mandarin Chinese and Malagasy

Claude Hagege

Parenthetical verbs are those which seem to be complement-taking but do not, in fact, really govern their complement. In many languages, there are either purely parenthetical verbs, or verbs with parenthetical uses among others, or both. These other uses of complement-taking verbs may be considered as cases of strong government whereas parenthetical and parenthetically used verbs, which refer to an opinion or a mood, may be considered as cases of weak government. This fact may be proved by several tests, depending on the languages under study. Those used here as illustrations are English, French, Chinese and Malagasy.

Let us consider the English verb *to presume*. As shown by exx. (1a), (1b) and (1c), its first person present singular can occur in three positions: followed by *that* and an indicative clause, or inserted at the middle or end of an indicative sentence. This variation, characteristic of a parenthetical use, is less often observed when *to presume* describes a psychological condition, and is, therefore, non-parenthetical, as in ex. (2b), a reply to (2a). This contrast between two types of uses is also found in English for other verbs: *admit, assume, believe, concede, conclude, confess, deduce, estimate, expect, feel, gather, guess, infer, know, maintain, think*. All of them may be used as parenthetical, or weakly-governing, verbs.

Similarly, in French, the first person present singular of *croire* "to believe" may occur in three positions and accepts no anaphoric pronoun when it is used parenthetically, i.e. refers to a mere opinion and not to a belief, as shown by exx. (3a), (3b), (3c) and (3d). Furthermore, *je crois* may not be inserted when it is taken in its psychological meaning, and when the indicative sentence is embedded after a declarative verb, as shown by exx. (3e) and (3f) respectively. These restrictions are typical of many other complement-taking epistemic verbs in French, and serve to spot, in a negative way, their parenthetical use.

Mandarin Chinese also provides interesting clues to parenthetical verbs. It possesses a refutative negation, *bìngbù*, meaning, as in ex. (4), that things are not as has been asserted or claimed, or implicitly considered as the hearer's opinion. This negation cannot be used in subordinate clauses, such as the one in ex. (5). However, *bìngbù* may quite freely be used in utterances containing parenthetical verbs, namely those that merely express an assertion mood: *gūjì* "to presume", *huáiyí* "to suspect", *xiāngxìn* "to believe", *zhīdao* "to know", *xiǎng* "to think", *shuō* "to say". This is shown by exx. (6) and (7), where we see that the main information is given by the embedded clause.

In Malagasy, state or event verbal nouns, traditionally assigned to an "adverbial voice", and formed by suffixing *-ana* to the verb stem, with optional prefixation of *i-*, are found, among other uses, as objects of various epistemic verbs, like *hay* "know", *atao* "consider as", and verbs of volition like *asaina* "order". These can also occur in the passive, as in (8), where *nasai-ny* may be left out without any change; this shows that its use here is parenthetical. The same applies to *mety* and *azo* "be allowed". They can be used as passives, either personal, as in (9), or impersonal excluding a subject, as in (10); but they also occur in examples like (11): here, *azo* may neither be treated as a personal passive, since in this case we would have the passive verb *lazaina* instead of *ilazana*, nor as an impersonal passive, since in that case the presence of a subject *ny rivotra* and its relationship with the verb *mirohondrohona* remain unexplained. Moreover, *azo ilazana* can be left out. For all these reasons, one should consider that *azo* functions in

(11) as a parenthetical verb.

Parenthetical verbs are an interesting case of verb-complement relationship. The above data shows that verbal government is a scale, whose lower grades are occupied, among others, by parenthetical verbs, i.e. verbs which, although they seem to govern a complement, are in fact modal parentheses, and do not contain the main information.

Data

- (1) English
 - a. *I presume that this city is very old*
 - b. *This city is, I presume, very old*
 - c. *This city is very old, I presume*
- (2) English
 - a. *I presume that this city is quite new*
 - b. *- Well, I presume that it is very old.*
- (3) French
 - a. *Je crois que Pierre va venir*
'I believe that Pierre is going to come'
 - b. *Pierre, je crois, va venir*
'Pierre, I believe, is going to come'
 - c. *Pierre va venir, je crois*
'Pierre is going to come, I believe'
 - d. **Pierre va venir, je le crois*
'? Pierre is going to come, I believe it'
 - e. **Pierre va venir, je crois, car je suis optimiste*
'*Pierre is going to come, I believe, for I am optimistic'
 - f. **Jacques raconte partout que Pierre, je crois, va venir*
'*Jacques is saying everywhere that Pierre, I believe, is going to come'.
- (4) Mandarin Chinese (Hoa 1986: 255)
wǒ bìngbù è (1SG REFUT.NEG hungry)
'I am not hungry at all (in reply to a statement like 'you are hungry!')
- (5) Mandarin Chinese (Hoa 1986: 274)
**wǒ xīwàng tā bìngbù lái* (1SG hope 3SG REFUT.NEG come)
'I hope he will not come at all'
- (6) Mandarin Chinese (Hoa 1986: 275)
wǒ huáyí tā bìngbù zhù zài nàr (1SG suspect 3SG REFUT.NEG live at there)
'I suspect that he does not live there at all'
- (7) Mandarin Chinese (Hoa 1986: 275)
tā shuō nǐ bìngbù shì zhōngguó rén (3SG say 2SG REFUT.NEG be China man)
'he says that you are not Chinese at all'.
- (8) Malagasy (Fugier 1999: 70)
ny goaikana nasai-ny izaran-kena (ART crow be.ordered-3SG share-meat)
'the crows were given meat by his order'
- (9) Malagasy (Fugier 1999: 66)
tsy azo vonoina io omby io (NEG be.allowed be.killed DEM ox DEM)
'this ox may not be killed'
- (10) Malagasy (Fugier 1999: 67)
tsy azo ipetrahana (NEG be.allowed sit)
'it is not allowed to sit down'

(11) Malagasy (Fugier 1999: 68)

mirohondrohona fa tsy mifioka intsony no azo ilazana ny rivotra (howl and NEG
blow more FOC be.possible say ART wind)

"one could say that the wind was howling, and no longer blowing".

10 with 2 is 12
(in any Language?!)
Refinements of the Typology of Addition
Thomas Hanke

This paper deals with the conceptual sources of additive markers in numerals. It supplements the previous results with a sample of 300+ languages. Besides the common case of unmarked sums (*twenty-three*), there are additive "links" which can be traced to a number of source types. Greenberg's (1978: 264f.) classification is generally accepted:⁹

1. the most common "comitative" type ('and, with') <ETH> including the rare exclusively verbal coordination¹⁰,
2. superessive links ('on, over'),
3. possessive links,
4. expressions for additional objects ('extra', 'left')

My study enables the following refinements:

- ad 1. I have found no additive constructions with purely comitative constructions, i.e. without coordinative meaning (like *with ...*). Comitatives and coordinators are also used in locative and other asymmetric constructions (like *and ... over it* or *with ... extra*).
- ad 2. Besides the most common superessive, other locative sources are used, including 'under' and movement expressions like 'to place' or 'go beyond'.
- ad 3. The possessive sources include 'have' and also 'get' constructions.
- ad 4. Greenberg's hypothesis, that augends with 'left' are always unexpressed,¹¹ is falsified.

In addition, there are a range of overtly bodily sum constructions which may combine 'hands' and 'feet' with the mentioned 1.-4. or with special means like 'other (side/hand)'. The limitation of asymmetric source constructions to the asymmetric serial addition (+1, +2, ... Greenberg 1978: 265f.) can be verified for all types. Here again, clear distinctions are important, as between the bodily use of 'other' and 'other' in doublings or pair/odd-sequences, and between coordinating and pure comitative constructions. The latter are of course a common source for nominal coordination¹², but only an indirect one for addition.

⁹ Cf. Hurford (1987: 237), Heine (1997: 33f.), Greenberg (2000: 777).

¹⁰ Not mentioned by Greenberg (1978). The different coordination types as additive sources can only be analyzed with a detailed examination of part of speech assignments of numerals.

¹¹ Apparently based on the limitation of 'left' in Germanic and Baltic languages to the 1st decade, cf. *twelve*.

¹² Greenberg's use of *comitative* directly conflicts with the differentiation of asymmetric and symmetric sums, cf. the general delimitation of *comitative* against *coordination* (Haspelmath 2000: 27-29).

References

- Greenberg, Joseph H. (1978): Generalizations About Numeral Systems. p. 249-295 in: Greenberg (ed.): Universals of Human Language. Vol. 3 Word Structure.
- Haspelmath, M. (2000): Coordination. *email.eva.mpg.de/~haspelmt/coord.pdf* [published in: T. Shopen (2004, ed.): Language typology and linguistic description, 2nd ed.].
- Heine, Bernd (1997): Cognitive foundations of grammar.
- Hurford, James R. (1987): Language and Number. The Emergence of a Cognitive System.

Diversity in Phonological Word Domains: Testing the Prosodic Hierarchy Hypothesis

Kristine Hildebrandt & Balthasar Bickel

It is traditionally assumed that the domains in which phonological rules (e.g. tone spread, apocope, phonotactic constraints, vowel harmony, etc.) apply (henceforth ‘pdomains’) cluster on a universally restricted number of nodes (i.e. the Prosodic Hierarchy Hypothesis: Booij 1983, Selkirk 1984; Nespor & Vogel 1986, McCarthy & Prince 1993, Hall 1999, Peperkamp 1997, Inkelas & Zec 1990, etc.): mora, syllable, foot, p-word, p-phrase, intonational phrase, p-utterance. Apparent deviations from this are often accounted for by positing recursive nodes (e.g. a p-word in a larger p-word: Ladd 1986ff), but such an account is controversial, because unlike in true recursion, the alleged recursion here involves units of different kinds (e.g. a small domain of prefix+stem vs. a larger domain of prefix+stem+suffix) (cf. Peperkamp 1997).

In this paper we present evidence that the Prosodic Hierarchy Hypothesis has no empirical support outside the few languages which gave rise to it. We survey, in a sample of nine Sino-Tibetan languages,. For each language we determine the precise p-domains for all segmental and suprasegmental rules that our sources allow us to identify (grammars, dictionaries, fieldwork). The results of such a comprehensive inventory include:

1. Languages often differentiate more than one p-word (see Table 1). By p-word we mean a p-domain that references morphological structure (typically, a stem ± affixes or clitics), but (unlike a p-phrase) not syntactic phrase structure.
2. The biggest p-word may or may not include all affix types in a language at once (e.g. Lahu has one p-word, the reduced stress word, that references prefix+stem and another, the tone change word, that includes stem+suffix, but has no p-word including prefix, stem and suffix at once).
3. P-domains in a particular language are not necessarily contained within each other, along the lines suggested by traditional notions of recursion (Selkirk 1984). For example, in Dege Tibetan the domain of aspirated affricate consonant lenition is prefix+stem but not suffixes or clitics, and the domain of tone alternation to default high is stem+suffixes+clitics, but not the prefix.
4. There is no evidence for areal factors governing p-word diversity in Sino-Tibetan: although Sino-Tibetan languages of Southeast Asia appear to have fewer nonisomorphic p-words than their relatives in the Himalayas, this is a trivial effect of having less options to begin with (if there is only one affix, there can only be one p-word including and one excluding this affix). If we control for the number of available options in a language, the differences disappear and we find the size of the biggest p-word is similar in all (surveyed) branches of the family (Table 1); likewise, in all branches we find that the smallest p-word is the same, viz. the stem alone, or the stem plus one affix/clitic (For example, the smallest p-word in polysynthetic Belhare is the stem alone: it is the exclusive domain of pretonic vowel syncope.) Therefore, unlike the phonological (in-)coherence of individual affixes (Bickel & Nichols 2005), the size of p-words does not track the areal distribution of what is traditionally called an ‘isolating’ language.

We conclude that, contra to traditional expectations, the precise definition of p-domains in a language is not driven by a universal preference for a minimal and well structured set of p-domains, nor by areal preference patterns favoring or disfavoring the phonological coherence of specific affixes, but by individual phonological histories.

Data

	Subgroup	Area	<i>N</i> (nonisomorphic p-words)	Size of biggest p-word relative to the number of available affixes and clitics
Belhare	Kiranti	Himalayas	5	0.6
Newar (Dolakha)	Newaric	Himalayas	4	0.75
Kinnauri	West Himalayish	Himalayas	5	0.6
Kham	Central Himalayish	Himalayas	10	0.75
Manange	Tamangic	Himalayas	4	1
Tibetan (Dege)	Bodish	Himalayas	5	0.75
Lahu	Lolo-Burmese	Southeast Asia	2	0.6
Meithei	Kuki-Chin	Southeast Asia	5	0.75
Kayah Li	Karenic	Southeast Asia	2	1

Table 1

References

Bickel, Balthasar and Johanna Nichols. 2005. Fusion of selected inflectional formatives. In: Haspelmath, Martin, Matthew Dryer, David Gil and Bernard Comrie [eds]. World Atlas of Linguistic Structures, pp. 86-89, Oxford: Oxford University Press.

Booij, Geert. 1983. "Principles and parameters in Prosodic Phonology." *Linguistics* 21. 249-280.

Häsler, Katrin. 1999. A grammar of the Tibetan Sde.dge dialect. University of Bern: Doctoral dissertation.

Hall, T. Alan. 1999. The phonological word: A review. In: Hall, T. Alan and Ursula Kleinhenz [eds.] *Studies on the phonological word*, pp. 1-22. Amsterdam /Philadelphia: Benjamins.

Inkelas, Sharon and Draga Zec, eds. 1995. *The phonology-syntax connection*. University of Chicago Press.

Ladd, Robert. 1986. Intonational Phrasing: The Case for Recursive Prosodic Structure. *Phonology Yearbook* 3: 311-340.

Matisoff, James A. 1973. *The grammar of Lahu*. University of California Press.

McCarthy, John and Alan Prince. 1993. "Generalized alignment." *Yearbook of morphology*, volume 6. Ed. by Geert Booij and Jap van Marle, 79-153. Dordrecht: Kluwer.

Nespor, Marina and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Peperkamp, Sharon. 1997. *Prosodic words*. The Hague: Holland Academic Press.

Selkirk, Elizabeth. 1984. *Phonology and syntax. The relation between sound and structure*. Cambridge, Mass: MIT Press.

Asymmetries in Word Architecture: Another Look at the Suffixing Preference

Nikolaus Himmelmann

Across the world's languages there is a fairly strong asymmetry in the affixation of grammatical material in that suffixes considerably outnumber prefixes in typological databases (Cutler, Hawkins & Giligan 1985, Bybee, Pagliuca & Perkins 1990, Hall 1992, Stump 2001). This is known as the *suffixing preference*. A number of explanations have been offered for this preference, including a psycholinguistic one which maintains that prefixing is disfavored for reasons of processing (taking a hearer's/comprehension point of view). This talk explores an alternative hypothesis which relates the preference for suffixes to a widely observable difference in the prosodic phrasing of function words. In spontaneous speech, a prosodic boundary frequently occurs between a *preposed* function word and its host. Such prosodic boundaries do not occur between a host and a *postponed* function word. Apart from the suffixing preference, this difference is also observable with regard to the direction of attachment of clitics. Stray adjunction is widely attested for clitics which *precede* their (syntactic/semantic) host, but hardly ever for those following it.

The difference in prosodic phrasing between preposed and postponed function words is, of course, also in need of an explanation. Two quite different, but mutually enforcing, motivations are offered for this difference. First, it has a natural basis in the mechanics of turn taking (*turn projection*). Second, it also follows from the assumption made in many models of speech *production* that function words are ready for 'spell-out' before content words are fully specified (e.g. Levelt 1999).

References

- Bybee, Joan L., W. Pagliuca & R.D. Perkins, 1990, "On the asymmetries in the affixation of grammatical material", in: W. Croft et al. (eds), *Studies in Typology and Diachrony*, Amsterdam: Benjamins, 1-42
- Cutler, Ann, John Hawkins & Gary Gilligan, 1985, "The suffixing preference: a processing explanation", *Linguistics* 23:723-758
- Hall, Christopher J., 1992, *Morphology and mind: a unified approach to explanation in linguistics*, London: Routledge
- Levelt, Willem J.V., 1999, "Producing spoken language: a blueprint of the speaker", in Brown, Colin M. & Peter Hagoort (eds), *The neurocognition of language*, Oxford: Oxford University Press, 83-122
- Stump, Gregory T., 2001, "Affix position", in: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals*, Berlin: de Gruyter, 708- 714

From Intralinguistic Analysis to Implicational Universals: a Cognitive-Typological Perspective on the Dative Alternation

Willem Hollmann

Ditransitives have been researched extensively, both by typologists (e.g. Siewierska 2003, Haspelmath 2005) and by scholars working on a single language, usually English (e.g. Gropen et al. 1989, Pinker 1989, Groefsema 2001). The latter tend to focus on the alternation between the indirect-object (IOC) and double-object or dative construction (DOC):

1. a. John gave a gift to Mary.
b. John gave Mary a gift. (Gropen et al. 1989:204)

Not all ditransitives occur in both patterns. There is still no account that fully “predicts” which verbs in English (or any other language) occur in DOC. Groefsema (2001) observes that Gropen et al. (1989) and Pinker (1989) fail to explain why some verb classes but not others are subject to the morphophonological constraint that only monosyllabic and initially stressed polysyllabic verbs occur in DOC. She suggests that, instead, verbs occur in DOC if that results in a special perspective, as compared to IOC, on the way in which the recipient is affected. However, Groefsema’s claim that a “non-dativisable” verb such as *donate* does not portray a different effect on the recipient than *give* is hard to prove. One may argue that our frame knowledge of donations specifies that in this case recipients are typically dependent on what they receive, and may be more grateful. Additional issues are raised by differences between dialects, as well as by contrasts with other languages, especially closely related ones.

The problems associated with Gropen et al., Pinker and Groefsema may be symptomatic of the pursuit of a (set of) neat constraint(s) on occurrence in DOC by studying ditransitives without regard for crosslinguistic (and intralinguistic) variation. On the cognitive-typological perspective taken here, there are universal constraints on DOC in the guise of implicational universals. For historical reasons different languages and dialects have different cut-off points. Compare Hopper & Thompson’s (1980) transitivity parameters: whilst not predicting the exact scope of the passive in any specific language they do represent universal implicational constraints. The hypothesis is that the same can be done for DOC.

I start from the common suggestion that DOC means ‘X causes Z to have Y’, as opposed to the ‘X causes Y to go to Z’ semantics of IOC. The extent to which the recipient is causally affected depends on the nature of the event, and may be measured against Hopper & Thompson’s parameters. This claim is perhaps not new: Gropen et al. implicitly argue for the relevance of volitionality and punctuality (1989:244).

Presumably, some parameters are more important than others in determining the degree of affectedness. Now since affectedness correlates with passivisability, one expects passive ditransitives to feature highly affected recipients. In line with Croft’s (2001:107) suggestion that intralinguistic analysis can be used to discover typological universals, this study is restricted to English. Furthermore, as English *give* is compatible with ditransitive events that display many of the different values of Hopper & Thompson’s parameters (as well as of some other properties that may be relevant to the semantics of ditransitives, e.g. degree of permanence of transfer), it suffices to focus on this verb. The procedure then is to compare corpus examples of passive ditransitive *give* to tokens of this verb in the active DOC. By analysing the extent to which the different semantic properties contribute to “passivisability” we also gain insight into the properties that contribute to “dativisability”, not only of *give* but of ditransitive verbs in

general. And because the parameters are crosslinguistically valid, the conclusions may be stated as implicational universals, which should apply to (all dialects of) English and other languages (that have the double-object pattern) alike.

References

- Curnow, Timothy Jowan (1997). *A grammar of Awa Pit (Cuaiquer): An indigenous language of south-western Colombia*. Unpublished PhD Thesis, Australian National University.
- Hulstaert, Gustaaf (1938). *Praktische grammatica van het Lonkundo (Lomongo) (Belgisch Kongo)*. Antwerp: De Sikkel.
- Hulstaert, Gustaaf (1965). *Grammaire du LýmýÉngý*, Vol. 2: *Morphologie*. Tervuren: MusŽe Royal de l’Afrique Centrale. (Annales Sciences humaines, 57).
- Nau, Nicole (1999). Was schlŠgt der Kasus? Zu Paradigmen und Formengebrauch von Interrogativpronomina. *Sprachtypologie und Universalienforschung* 52.2:130-150.
- Wierzbicka, Anna (1996). *Semantics: Primes and Universals*. Oxford: Oxford University Press.

The Many Faces of SVO: Noun Phrase Structure in Chinese and Thai

Peter Jenks

Constituent order, particularly that of verbs and objects, is generally considered to be predicative of the branching patterns of a particular language, with head-initial and head-final being the clearest opposition. However, while the presence of VSO/VOS or SOV word order is a strong predictor of a language's general branching pattern, the branching properties of SVO are far from homogenous (Travis 1989). Moreover, a number of factors besides branching tendencies seem to correlate with a language having SVO word order. In particular, it seems that languages with highly isolating morphology tend to be SVO (Siewierska & Bakker 1996). This paper examines two such languages, Thai (Tai-Kadai) and Mandarin Chinese (Sino-Tibetan), which have rigid SVO word order but markedly different branching patterns—most noticeably in their noun phrase structure.

On the surface, Chinese and Thai are typologically quite similar: both are SVO; have predominantly analytic and isolating morphology; lack overt tense or aspect marking on the verb; express tense/aspect with sentence particles; form general questions using a sentential question article; have *wh*-in situ, and use numeral classifiers. Despite these apparent similarities, Thai and Chinese NPs, which use the same categories as their building blocks, are nearly mirror images of one another. This paper presents and analyzes the Thai and Chinese NPs and provides an explanation for the striking structural differences between (across) these two languages (1, 2).

Chinese NPs have obligatorily phrase final head-nouns, preceded by a Num(be)r+Cl(assifier) constituent (Tang 1990, Gao 1994), which is preceded in turn by a demonstrative. In this paper, we show that numerals and the classifiers that obligatorily accompany them form a single constituent due to their tightly bound phonetic and distributional properties. Chinese adjuncts, many of which are structurally reducible to relative clauses with a complementizer on the right, may occur either between the noun and Num+Clf constituent or before the demonstrative. True adjectives can only occur between the head noun and Num+Clf (1).

Conversely, Thai NPs exhibit obligatorily phrase initial bare nouns, followed by a Num/Q+Clf constituent, which is followed by a demonstrative or determiner. The evidence that Thai nouns are heads stems first from their obligatoriness and second from the clear subcategorization of classifiers for nouns, indicating a head-complement relationship. Like Chinese, Thai treats the combination of classifier and numeral as a single constituent, as evidenced by distributional evidence and constituency tests. Thai adjuncts, like Chinese ones, can occur either between the head noun and the Num/Q+Clf constituent or after the demonstratives, again with the exception of adjectives, which can occur between nouns and the Num/Q+Clf constituent or between the latter constituent and demonstratives.

I argue in this talk for a three-tiered NP structure, shown in (4) This structure clearly accounts for not only the rigid word order of these NPs but also the subcategorization frames exhibited by the three phrasal heads. When the structure of Thai and Chinese sentences are compared, the contrast in their branching tendencies is clearly evident (S1-2). There are other additional corollaries of these different branching patterns in Thai and Mandarin Chinese, including possessor-possessee order, Comp-RC order, and PP-V order.

As Thai and Chinese are polar extremes of NP branching, Vietnamese (Mon-Khmer), another SVO language employing numeral classifiers, appears to represent the typologically intermediate situation, with mixed branching even within its NP. While the Num+Clf constituent occurs to the left of the head noun, demonstratives and adjuncts occur the right of the noun (3) (Nguyen, 2002, Li 2004).

These dramatic differences in head-directionality in the Mandarin Chinese and Thai — as manifested in NP structure and elsewhere — exist despite the fact that both languages are very strongly SVO. This may indicate that limiting order typology to VO/OV types may not be sufficient for predicting the headedness patterns of SVO languages, contra Dryer (1991, 2001). I suggest that the SVO word order found in isolating languages is motivated not by their headedness properties but by a lack of morphology, which does not provide agreement or case-marking as means of argument encoding.

Data

(1) The Chinese NP

wo xihuan.de na san.zhi rou.de qianbi
 I like COMP that three.CLF red.COMP pencil
 “Those three red pencils that I like”

(2) The Thai NP

nangseuu jaj saam lem nii tii phom chawp
 book big three CLF this that I like
 “These three books that I like.”

(3) The Vietnamese NP

ba con chim d#en ma\$ ba`n thay d#o\$
 three CLF bird black that you see there
 ‘Three black birds that you see there.’

(S1) Chinese NP Structure

[_{DP} [_{D'} [_D] [_{ClfP} [_{Clf} [_{Num+Clf}] [_{NP} [_{N'} [_N]]]]]]]]

(S2) Thai NP Structure

[[[[[[[[_N] _{N'}] _{NP}] [_{Num+Clf}] _{Clf}] _{ClfP}] [_D] _{D'}] _{DP}]

References

- Dryer, Matthew S. 1991. SVO Languages and the OV/VO Typology. *Journal of Linguistics* 27: 443-482.
- Dryer, Matthew S. 2001. Mon-Khmer Word Order from a Crosslinguistic Perspective In *Papers from the Sixth Annual Meeting of the Southeast Asian Linguistics Society, 1996*, edited by Karen L. Adams and Thomas John Hudak, pp. 83-99. Tempe: Arizona State University Program for Southeast Asian Studies.
- Gao, Qian. 1994. Chinese NP structure. *Linguistics* 32, 475-510.
- Li, Xusheng. 2004. Vietnamese Classifiers and Nominal Phrase Structure. Unpublished Paper. University of California, San Diego.
- Nguyen, Tuong Hung. 2003. Internal Structure of the Vietnamese DP. MS. Boston University
- Siewierska, Anna and Dik Bakker. 1996. The Distribution of Subject and Object Agreement and Word Order Type. *Studies in Language*, 20, 1, 115-161.
- Tang, Chih-Chen Jane. 1990. A note on the DP analysis of the Chinese noun phrase. *Linguistics* 28, 337-354.

Travis, Lisa. 1989. Parameters of Phrase Structure, in M. R. Baltin and A. S. Kroch (eds.), *Alternative Conceptions of Phrase Structure*, University of Chicago Press, Chicago, pp. 263-279.

Grammatical Relations in Minangkabau: A Typological Analysis Jufrizal Jufrizal

In linguistic typological studies, the concept of grammatical relations has an important role, especially at the (morpho)syntactical level. In relational grammar, the grammatical relations are S(ubject), D(irect) O(bject), I(ndirect) O(bject), and Obl(ique). While S, DO, and IO are pure syntactical relations, Obl relations are semantic ones. Obl relations may be categorized as locative, benefactive, and instrumental. This paper proposes a typological analysis of syntactic grammatical relations and semantic grammatical relations. Linguistic typological analysis shows that in Minangkabau the grammatical relation S has the following characteristics: (1) preverbal; (2) agent; (3) accessibility to relativization; and (4) unmarked form. DO in Minangkabau has the following characteristics: (1) directly after verb; (2) no insertion between the DO and its verb; (3) can be promoted to S through passivization; and (4) can not be deleted in coordinative construction. The characteristics of IO in Minangkabau are: (1) directly after verb in bitransitive construction; and (2) can be promoted to S through passivization. Obl in Minangkabau is marked by prepositions.

The Indirect Object as an Indicator of Transitivity

Seppo Kittila

The purpose of the present paper is to examine how transitivity-related features (such as affectedness of the recipient and the completedness of events) are expressed by modifying the form of the indirect object (understood in a very broad sense). In addition, some features typical of three-participant events only will be examined (see (5) for an example). The paper is primarily concerned with the encoding of events, which clearly involve a recipient/goal or a beneficiary, even though some other cases are also discussed. The role the indirect object bears needs, however, to be maintained, which means, for example, that the differences in the encoding of the roles of recipient and locative will not be considered, nor will the differences which follow from the verbs used. For example, in many languages >give= differs from other trivalent verbs in formal transitivity (the Recipient bears different marking depending on the verb), but this is not considered here.

The most relevant semantic features to be discussed in the present paper are the affectedness/involvement of the recipient, the permanence of transfer, the completedness of the event, and animacy. An example of each is illustrated in (1)-(4). In Sochiapan Chinantec, the variation between ditransitive (1a) and transitive (1b) frame signals differences in the permanence of transfer. The formally more transitive (1a) profiles permanent transfer, while the meaning of (1b) is closer to ‘entrust with=’. Similar variation is attested also e.g. in Tsez and Korafe. In Macedonian, the variation between genitively and datively marked Sources follows from affectedness; in (2b) the source is considered more affected than in (2a). Further examples are found in Dutch, Zulu and Tuvaluan. The examples from Paamese illustrate a coding difference between successfully completed and non-completed events. The use of the punctual dative, as in (3a) implies that the act of giving has been completed in that the recipient has received the transferred entity, while the use of the areal dative implies that the act of giving has been only partially completed. The examples from Tigrinya illustrate the effects of animacy on the encoding of the indirect object; the animate goal occurs in the dative, while inanimate goals are coded adpositionally. Similar variation is attested in a number of languages including Finnish and Hup. The examples from Maori illustrate a case where the preposition used for encoding the Recipient changes based on whether the recipient is going to use the entity transferred to its sphere of control for a specific purpose or not. Similar cases are found in Khasi and Khmer. (5) exemplifies a feature confined to the indirect object, while those in (1)-(4) are common to transitive and ditransitive events.

The present paper examines the features briefly discussed above in a wide range of languages. As noted above, the focus lies on recipients/beneficiaries, but also other roles may be considered, if relevant to the examination. In addition to merely illustrating the different features that affect the Recipient marking, the paper also examines the relation of formal and semantic transitivity. This means that I will take a look at whether higher degrees of formal ditransitivity are related to high semantic transitivity or not. This is the case in Sochiapan Chinantec, where the double object frame profiles permanent transfer, while in Maori the difference is merely the preposition used, and the formal transitivity of the clauses is not affected.

Data

Sochiapan Chinantec (Foris 1998: 212)

- (1a) cuéh₃₂ tsú₂ pé₁ qu e₃ tsa₃háu₂
 give.FUT.3 3 Peter money tomorrow
 ‘S/he will give Peter money tomorrow’
- (1b) cué₃₂ tsú₂ qu e₃ Zi₁con₂ pé₁ tsa₃háu₂
 give.FUT.3 3 money to Peter tomorrow
 ‘S/he will give money to Peter tomorrow’

Macedonian (Janda 1998: 258)

- (2a) ani pes by od n ho k rku chleba
 nezval
 not even dog.NOM would from him.GEN crust.ACC bread not
 took
 ‘Not even a dog would take a crust of bread from him’
- (2b) ani pes by mu k rku chleba nezval
 not even dog.NOM would him.DAT crust.ACC bread not took
 ‘Not even a dog would take away his crust of bread’

Paamese (Crowley 1982: 197)

- (3a) inau na-saani ratioo mini-e
 1SG 1SG.REAL-give radio PUNCT.DAT-3SG
 ‘I gave him my radio’
- (3b) ono-mo vilme=mune na-saani-e veni rei
 kalimoo
 POSS-2SG film=ADD 1SG.REAL-give-3SG AREAL.DAT Ray Gallimore
 ‘I delivered your film to Ray Gallimore’s too (though I didn’t give it to him personally)’

Tigrinya (examples courtesy of Saliem Leghese Tesfazghi)

- (4a) ti m mh r n ti m t?s af n - akim s d dd- -l-u
 the teacher ACC.the book DAT-doctor sent-he-DAT-him
 ‘The teacher sent the book to the doctor’
- (4b) ti m mh r n ti m t?s af nab dm nt n s d dd-o
 the teacher ACC.the book to Edmonton sent-itM
 ‘The teacher sent the book to Edmonton’

Maori (Bauer et al 1993: 272)

- (5a) i hoatu a mere i te keke ki tana tama
 T/A give PERS Mary DO the cake PREP SG.GEN.3SG son
 ‘Mary gave the cake to her son’ (without any necessary expectation what to do with it)
- (5b) i hoatu a mere i te keke maa tana tama
 T/A give PERS Mary DO the cake INT.GEN SG.GEN.3SG son
 ‘Mary gave the cake to her son’ (for himself to keep, with the expectation that he will eat it)

References

- Bauer, Winfried (with William Parker and Te Kareongawai Evans). 1993. *Maori*. London/New York: Routledge.
- Crowley, Terry. 1982. *The Paamese language of Vanuatu*. Canberra: Pacific Linguistics.
- Foris, David 1998. Sochiapan Chinantec GIVE: A window into clause structure. In Newman (ed.), 209–248.
- Janda, Laura 1998. GIVE, HAVE and TAKE in Slavic. In Newman (ed.), 249–265.
- Newman, John 1998. *The linguistics of giving*. Amsterdam/Philadelphia: John Benjamins.

Differential Goal Marking: Concerning the Animacy Effects on the Encoding of Goals

Seppo Kittila

It is received wisdom in linguistics that animate objects are marked in a more elaborate way than inanimate arguments in the same function. This phenomenon is known as Differential Object Marking (henceforth DOM). There are dozens of papers dealing with the topic at the level of direct objects. The present paper, on the other hand, adopts another perspective in examining the effects of animacy on the encoding of goals. The focus lies on the formal types of Differential Goal Marking (DGM for short). In addition, the rationale behind the attested types of DGM will also be touched upon. The goal of the paper is to show that animacy effects are not confined to direct objects, but they also make an important contribution to the object marking in more general terms.

Formally, DGM has three distinct manifestations. The first type of DGM is illustrated by languages in which animate Goals are marked in the same way as the direct object, while inanimate Goals receive an oblique-like treatment. The encoding of animate Goals can employ the mechanisms used exclusively for animate Goals, or the marking can coincide with direct objects in general. The former is illustrated in (1) from Hup, while Athpare in (2) exemplifies the latter. Moreover, in some languages, such as English and Icelandic, animate Goals are only optionally accorded core-like status via dative shift, while this is not possible for inanimate Goals. The second type of DGM is represented by languages in which Goals invariably bear marking distinct from direct objects, but in which the marking is nevertheless sensitive to animacy. Examples are provided in (3) and (4). The last type of DGM is labelled as Extended DOM based on the features it has in common with canonical DOM. The type is exemplified in (5) from Gujarati. In Gujarati, animate Goals are overtly marked, while inanimate Goals bear zero marking, which coincides with canonical instances of DOM. Note, however, that I have not come across a single language, which invariably marks the Goal following DGM of the last type, but it usually affects only specific types of Goals (like names of places). In addition to the three different types of DGM there are also languages in which animacy makes no contribution to the encoding of Goals. These languages are not considered in the present paper.

In addition to examining the formal types of DGM, the paper at hand also discusses the rationale behind the attested types. Features of DOM are highly relevant to the discussion here. These comprise the generally exceptional treatment of animate arguments, disambiguation, highlighting of the atypical nature of the object argument, and affectedness. The first of the proposed functions serves as a kind of default explanation. It definitely makes a contribution to DOM and also DGM, but another explanation should be favoured, if available. The contribution of disambiguation to DOM is clear, since the more elaborated marking of animate objects is the only way to disambiguate the semantic roles of two potential agents. On the other hand, the contribution of disambiguation is less clear in the case of DGM. For example, disambiguation does not aid us in explaining the occurrence of type 2 of DGM in any way. Disambiguation is very closely related to the highlighting of the atypical nature of animate objects, whence it is unsurprising that it does not make a major contribution to explaining DGM. Rather, the more elaborated marking of animate Goals highlights the typical nature of its referent in most cases. Affectedness associated with animate arguments in general, on the other hand, plays an important role in DGM, if we consider

the oblique-like marking of less affected transitive objects. Moreover, the close relation of affectedness and individuation has been long recognized (see Hopper and Thompson 1980: 252), which makes the occurrence of DGM understandable.

Data

Hup (examples courtesy of Pattie Epps)

- (1a) §ameriku §awponsu-an pij deh-an d'o§-ham-yæh-æh...
 Americo Alfonso-OBJ cabari village-OBJ take-go-order-DECL
 'Americo sent Alfonso to cabari village...'
 (1b) tiyi§ tvh=tæh-an wan n]§-jh
 man 3SG=son-OBJ knife give-DECL
 'The man gave the knife to the boy'

Athpare (Ebert 1997: 107, 111)

- (2a) khan na topi sa-ða a-choös-u-e-na?
 you this cap who-OBL 2-sell-3U-PAST-NML
 'Who sold this cap to you'
 (2b) sak ka-paö-i tay-e-na
 who 2POSS-house-LOC come-PAST-NML
 'Who came to your house'

Finnish

- (3a) musiikkipedagogi lähett-i kirje-en performanssitaiteilija-lle
 music.pedagogue send-3SG.PAST letter-ACC performance.artistr-ALL
 'A/the music pedagogue sent a/the letter to a/the performance artist'
 (3b) musiikkipedagogi lähett-i kirje-en turku-un
 music.pedagogue send-3SG.PAST letter-ACC Turku-ILL
 'The music pedagogue sent the letter to Turku'

Balinese (examples courtesy of I Wayan Arka)

- (4a) Guru-ne nto ngirim buku sig/*ke anak-e nto
 teacher-DEF that AV.send book to person-DEF that
 'The teacher sent a book to the person'
 (4b) Guru-ne nto ngirim buku ke/*sig Indonesia
 teacher-DEF that AV.send book to Indonesia
 'The teacher sent a book to Indonesia'

Gujarati (examples courtesy of Babu Suthar)

- (5a) sikshak-e vidaarathi-ne pustak mokl-y-un
 teacher-ERG student-DAT book.NEUTR.SG send-PASTPERF-NEUTR.SG
 'The teacher sent a/the book to the student'
 (5b) sikshak-e delhi pustak mokl-y-un
 teacher-ERG Delhi book.NEUTR.SG send-PASTPERF-NEUTR.SG
 'The teacher sent a/the book to Delhi'

References

- Ebert, Karen. 1997. *A grammar of Athpare*. München/Newcastle: Lincom Europa.
 Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language* 51, 251-299.

Renewing Connection with the Data: Reciprocal Constructions in Melanesian and Polynesian Languages

Ekkehard König & Claire Moyse-Faurie

During the last ten years or so a wide variety of descriptive, theoretical as well as cross-linguistic studies have been published on reciprocity. Among the latter we find Kemmer (1993), Frajzyngier (2000), Siloni (2002), Maslova & Nedjalkov (2004), Heine & Miyashita (2004) as well as a large survey initiated and coordinated by V. Nedjalkov (to appear). In spite of providing a wealth of interesting data, observations and analyses for individual languages, however, these cross-linguistic studies do not give us an integrated and reasonably adequate typology of reciprocal constructions comparable - in quality and simplicity - to that developed by Faltz (1985) for reflexives. At least three different attempts are currently being made to capture the relevant patterns and limits of variation in this domain (Nedjalkov, Evans, König & Kokutani). In the first part of our paper we will discuss the current state of these projects, their basic assumption and preliminary results, but also the differences between them. The simultaneity of these three projects provides the unique opportunity to observe the process of formulating a typology and to find out how much of its results are determined by basic theoretical assumptions, the size and quality of the language sample as well as the generalizations made about basic types and connections between variant properties.

In the second part of our paper we will do what J. Firth called “renewing connection with the data”: Some of the typological generalizations discussed before will be confronted with the languages of one particular family (Oceanic) and one particular area. Melanesian and Polynesian languages exhibit interesting properties in their encoding of reciprocity which are rarely found elsewhere:

- the co-occurrence of strategies that are alternative choices elsewhere
- the use of bound or co-referential personal pronouns together with affixes
- the almost complete lack of quantificational expressions like *each other* in English
- the use of intensifiers for both reflexivity and reciprocity
- the frequent use of the same expression for reciprocity and the middle voice in opposition to markers of reflexivity
- the incipient grammaticalization of lexical items from unusual source domains

The Writing Systems of Sumatra

Uli Kozok

The writing systems of Sumatra are one of the least researched writing systems in Indonesia. Although it is known that they ultimately derive from the South Indian Pallava script, very little is known about the internal relationship between the northern (Surat Batak) and the southern (Surat Ulu) group, and even within the two groups little is known about the genetic relationship of the various writing systems within each group.

My research on the internal relationship between the five closely related scripts of the northern group shows that there is a clear pattern of a development from south to north, suggesting that the Batak scripts probably originated south of the present location of their distribution, and there is good reason to believe that the predecessor of the Batak scripts was used by the Minangkabau people of West Sumatra.

A similar pattern of development still needs to be established for the southern group, which, however, seems difficult since it is very likely that at a certain stage of the development from an Old Sumatran post-Pallava script, the southern script was radically redesigned to create a homogeneously looking script. A similar development is likely to have taken place in Sulawesi, where this has also cause difficulties for scholars to establish a genetic relationships with any post-Pallava script.

Some stone inscriptions from Lampung, and the recently re-discovered 14th century Tanjung Tanah manuscript, have provided new insights into the development of the scripts of the southern group, and its relation to the northern group. The presentation will evaluate the newly discovered evidence and their relevance for the obscure history of the Surat Ulu scripts.

"Detached Head-Marking" in NPs

Yuri Lander

Despite the fact that Nichols' (1986, 1992) influential head-marking/dependent-marking taxonomy focuses mainly on morphological patterns, a number of constructions with syntactically autonomous markers (e.g., prepositions and auxiliaries) also have been included in this typology. Indeed, many patterns of this kind show association of relation markers with heads or dependents, and as is argued by some scholars, syntactically autonomous relators are often at the intermediate stage in their development into morphological head- or dependent-marking. In this paper, however, we wish to discuss two kinds of structures which pose certain problems for Nichols' taxonomy. In general, these structures (found among adnominal possessives) are usually identified as head-marking patterns, since they employ markers whose position is determined at least partly by the head and/or display morphology that (for a given language) is recognized as head-marking. Further, they are found in languages that actively use the head-marking strategy. Yet relators in the patterns under discussion are associated with both participants of a relation - in contrast to markers that tend to be associated with either head or dependent or seemingly are associated with any or neither of them (as is the case in "neutral" marking).

First, in many languages the position of a marker is determined not by the head but by a projection of its. Sometimes this results from incorporation of some modifiers, as it does in Adyghe (1). But occasionally such a relator signals the boundary of a constituent to which a modifier is added; an example of this phenomenon is presented by the *izafet* construction in a number of Iranian languages (2). A further development of this strategy marks not only the presence of a modifier but also its function; cf. Zaza Kurdish (3). Finally, we find that a relator can indicate not only the function of a dependent constituent but also its inherent properties. This is observed in some languages of Western Indonesia and Oceania, where pronominal agreement markers - while immediately preceding possessors, their controllers - also clearly refer to their referentiality (4), as is occasionally reflected by their complex interaction with optional determiners. That this function of a head-marking formant indeed can be determined by its position with respect to dependent is further supported by the fact that where determiners became obligatory, the position of agreement often turns out to be independent of the place of the possessor, and the system develops into pure head-marking (5).

Second, sometimes "head-marking" morphology appears on an "intermediate" element that already identifies a construction. This is a common situation with possessive classifiers (regardless of whether they mark the type of the relation or the type of the possessor); cf. (6)-(7). It can be argued, however, that these elements are themselves heads or NPs, which are either dominated by another phrase headed by the possessum (6; analogical structures are observed for dependent-marking patterns as well) or include the possessum as a non-referential modifier (7). Either case, we find strong association of the agreeing classifier with the possessor together with a head-marking structure.

These two kinds of constructions are likely to be quite stable in spite of the fact that they do not fit in pure head-marking or dependent-marking pictures. We can expect, then, that they may come as stoppages during the drift of certain components of a construction from a dependent to the head. This indeed may be the case for some possessives with "auxiliary pronouns" in European languages (see Koptjevskaja-Tamm 2003: 665-670 among others), which occasionally seem head their own NPs (as is seen

in (8) from the plural agreement in the verb). The alternative scenario, with such pronouns marking the boundary of a projection of the head along with the role of the possessor is not, of course, excluded as well. Additionally, we may speculate that the structures of this kind can hamper (formal) grammaticalization of certain components of possessives including possessive classifiers and possessive pronouns.

Finally, the structures discussed here show that the conceptions of head-marking and dependent-marking actually involve tendencies to cluster properties that associate a marker with various components of a construction, these properties being by no means mutually exclusive.

Examples

- (1) ADYGHE (Gorbunova 2004)
 s-jE-pHe-LaqW
 1SG.PR-POSS-wood-leg ‘my wooden leg’
- (2) PERSIAN (Ghomeshi 1997: 739)
 [[kif-e charm]-e bozorg]-e dust-am
 bag-IZF leather-IZF big-IZF friend-PR.1SG ‘my friend’s big leather bag’
- (3) ZAZA KURDISH (Smirnova & Ejubi 1998: 35)
 láz-ə m̄n-o pil
 son-IZF.POS.M 1SG:OBL-IZF.ATR.M elder ‘my elder brother’
- (4) LAMPUNG (Walker 1976: 10, 20)
- a. Hadat pekon asal-ni bapaq-ni kajong-ku pagun kuat.
 custom village origin-PR.3 father-PR.3 spouse-PR.1SG still
 strong
 ‘The traditions of the native village of my wife’s father are still strong.’
- b. Tanggay-ni lemaong-hina tajom-tajom.
 claw-PR.3 tiger-DEM sharp:RDP
 ‘The claws of the tiger are very sharp.’
- c. Lemaong-sa nganiq daging binatang lain.
 tiger-DET ACT:eat flesh animal other
 ‘Tigers eat the flesh of other animals.’
- (5) TONDANO (Sneddon 1975: 120)
 m-pasar-ea s̄ra/
 DET-market-PR.3PL fish ‘their fish market’
- (6) MONO-ALU (Fagan 1986: 34)
 [mafa e-gu] iana
 1SG POSCLR-PR.1SG fish ‘my fish’
- (7) KOKOTA (Palmer 1999: 131)
 [ge-na viri] havi
 POSCLR-PR.3SG tobacco Havi ‘Havi’s tobacco’
- (8) LOW SAXON (Strunk 2004)
 Peter sien un Hinnerk sien
 Peter.M.SG.ACC his.N.SG and Hinnerk.M.SG.ACC his.N.SG
 Hus sünd schön
 house.n.SG.NOM be-PL beautiful
 ‘Peter’s and Hinnerk’s house(s) are beautiful.’

References

- Fagan, J.L. 1986. A Grammatical Analysis of Mono-Alu (Bougainville Straits, Solomon Islands). Canberra: ANU.
- Ghomeshi, J. 1997. Non-projecting nouns and the ezafe construction in Persian // *Natural Language & Linguistic Theory* 15 (4): 729-788.
- Gorbunova, I. 2004. Kategorija ottořžimosti/neottořžimosti v adygejskom jazyke. Ms.
- Koptjevskaja-Tamm, M. 2003. Possessive noun phrases in the languages of Europe. In *Noun Phrase Structure in the Languages of Europe*, ed. F. Plank, 621-722. Berlin/New York.
- Nichols, J. 1986. Head-marking and dependent-marking grammar // *Language* 62 (1): 56-119.
- Nichols, J. 1992. *Linguistic Diversity in Space and Time*. Chicago: The University of Chicago Press.
- Palmer, B. 1999. A Grammar of the Kokota Language, Santa Isabel, Solomon Islands. Ph.D. diss., Univ. of Sydney.
- Smirnova, I.A. & K.R. Ejubi. 1998. *Kurdskej dialekt zaza*. Moscow.
- Sneddon, J. 1975. *Tondano Phonology and Grammar*. Canberra: ANU.
- Strunk, J. 2004. The missing link? An LFG analysis of the prenominal possessive construction in Low Saxon. Ms.
- Walker, Dale F. 1976. *A Grammar of the Lampung Language: the Pesisir Dialect of Way Lima*. (NUSA 2.) Jakarta.

Lexical Typology of AQUA-Motion

Yuri Lander & Timur Maisak

The study of lexical and grammatical means that languages use to describe motion in space attracted much attention in the recent decades (see Talmy 1975, 1985, 1991; Fillmore 1983, 1997; Slobin 1996, 2004, among others). At the same time, most studies devoted to this issue focus on the Path component of the motion situation, i.e. on the description of direction, whereas cross-linguistic semantics of manners of motion – e.g., running, jumping, climbing or crawling – remains mysterious to a great extent (among rare exceptions, see Taylor 1996; Fillmore & Atkins 2000 for some discussion).

As a part of a larger project concerning typological description of manner of motion verbs, this paper is confined to a narrow lexical field of AQUA-motion¹³, i.e. verbs describing motion in a liquid medium (cf. Russian *plyt'*, *plavat'* or English *swim*, *sail*, *float*, *navigate*, etc.). Here we present our preliminary generalizations based on the data from about 30 languages (including Slavic, Germanic, Romance, Uralic, Turkic, Semitic, Caucasian, etc.).

Within the broad AQUA-motion field, there are three main domains that the languages distinguish most consistently on the lexical level: these are “active motion in water”, “passive motion in water” and “motion of vessels or aboard a vessel”. Using English words as (rough) metalanguage labels, we can call the respective groups ‘verbs of SWIMMING’, ‘verbs of FLOATING’, and ‘verbs of SAILING’. “Average” AQUA-motion systems distinguish between these three domains. Apart from them, there are “rich” systems with additional lexical oppositions within one of the domains, as well as “poor” systems where some of the oppositions are neutralized (and one and the same means is used in more than one domain). Note that in order to distinguish between several types of AQUA-motion on the lexical level, it is not necessary for a language to have several special words, whose main meaning is a particular manner of motion in water. Lexical “gaps” may be filled with verbs with a general meaning: e.g. in many languages there are no special SAILING verbs, and this type of motion is described normally by means of ‘go’ and ‘come’ (cf. Turkic languages).

As a rule, verbs of SWIMMING are applicable to all animate beings capable of “active” swimming; it happens very rarely that some type of animate subjects stands in opposition to other “swimmers” (cf. Arabic and Hebrew, where swimming of water birds is normally described by means of verbs of FLOATING, not SWIMMING). In the SAILING domain additional oppositions are also not very typical, but in some cases motion in a certain environment (usually in the sea), or on a certain type of a vessel (e.g. sailing or rowing) may get separate lexical expression: cf. Korean *hanghay hata* ‘to sail the seas’ or Portuguese *velejar* ‘to sail’ (English *sail* has acquired a more general meaning and now can be used with other types of vessels as well, not necessarily having sails: e.g. *US aircraft carrier sails for the Gulf* [BNC]).

On the contrary, the FLOATING domain is not so homogeneous, and several subtypes of “passive” AQUA-motion are often distinguished, such as “non-controlled drifting” (like a log in a river) vs. “directionless motion within the water” (like a leaf in a puddle), etc. Thus, in Japanese *nagareru* is a typical “directed floating” verb, whereas *tadayou* is a verb of directionless motion, and *uku* / *ukabu* are verbs describing “buoying” on a surface without significant motion.

Sometimes a language may use one and the same verbal root in all three domains (cf. Russian *plyt'*/*plavat'*, which have a common root diachronically, or Georgian *curwa*

¹³ We are grateful to Philippe Bourdin who suggested this term for us.

and its prefixed derivatives). However, such situation seems to be rare, and even in such cases there exist some specialization of derivatives. Thus, in Russian only the iterative derivative occupies a BUOYING subdomain (cf. *v luzhe plavaet/*plyvet list* ‘a leaf is floating in the puddle’), and in Georgian only prefixed verbs can be used when a directed motion is being described.

References

- Fillmore, Charles J. How to know whether you're *coming* or *going* // Rauh, G. (ed.). Essays on deixis. Tübingen: Narr, 1983.
- Fillmore, Charles J. Lectures on Deixis. Stanford: CSLI Publications, 1997.
- Fillmore, Charles J. and B. T. S. Atkins. Describing polysemy: the case of ‘crawl’ // Yael Ravin and Claudia Leacock (eds). Polysemy: Linguistic and Computational Approaches. Oxford: Oxford University Press, 2000.
- Slobin, Dan I. Two ways to travel: Verbs of motion in English and Spanish // Thompson, Sandra A. & Masayoshi Shibatani (eds.). Grammatical constructions: Their form and meaning. Oxford: Oxford University Press, 1996.
- Slobin, Dan I. The many ways to search for a frog: Linguistic typology and the expression of motion events // Strömquist, Sven & Ludo Verhoeven (eds.). Relating events in narrative. Vol. 2: Typological and contextual perspectives. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.
- Talmy, Leonard. Semantics and syntax of motion // Kimball, J. (ed.). Syntax and semantics. Vol. 4. New York: Academic Press, 1975.
- Talmy, Leonard. Lexicalization patterns: semantic structure in lexical forms // Shopen, Timothy (ed.). Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon. Cambridge: Cambridge University Press, 1985.
- Talmy, Leonard. Path to realization: A typology of event conflation // Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society. Berkeley: Berkeley Linguistic Society, 1991.
- Taylor, John R. On running and jogging // Cognitive Linguistics 7: 21-34, 1996.

Reciprocity in a Polysynthetic Language: Adyghe Reciprocals

Aleksandr Letuchiy

Adyghe language is a polysynthetic language (Abkhaz-Adyghe family). Both ergative and absolutive arguments of the verb must be cross-referenced in the verb form. In my report I would like to analyze means of expressing the reciprocal meaning in this language and to address a problem, whether these features are connected with specific properties of polysynthetic languages.

The means of expressing reciprocity are the prefixes *ze-/zE-* and *zere-* and the pronouns *zEm zEr* ‘one another’ (literally ‘one one’) and *zEm adrEr* ‘one another’. In most cases use of *ze-* or *zere-* depends on whether transitivity of the initial verb is changed in the reciprocal construction or not, but this rule has exceptions: cf. *zewE{aZ’ERex* ‘they wounded each other’ (initially transitive verb). The prefix *ze-* is initially the reflexive prefix, so the Adyghe examples are interesting because they show that in the modern language this prefix occupies new domains: after it became the reciprocal marker, it begins to be used in new reciprocal contexts. Some verbs can form the reciprocal construction both by means of *ze-* or by means of *zere-*, with different meanings: cf. *LeRWE* ‘see’, *zeLeRWE* ‘see each other’ and *zereLeRWE* ‘meet’.

The most problematic case is “possessive reciprocals” [Nedjalkov 2002]: in this case either *zere* or *zere-Re* (the combination of the reciprocal and the causative marker) can be used: cf. *a{axer zere-/zereRe-we{aZ’ERex* ‘they wounded each other’s hands’. The transitivity of the initial verb does not change, because the reciprocal construction is also transitive, but reciprocity is stated between the initial agent and the possessor of the initial patient (‘hand’). Therefore, semantically and syntactically these constructions are very close to reciprocals like ‘they wounded each other’, except for the fact that the body-part is defined.

Another difficult case is the reciprocal construction, formed from so-called “inversive verbs” (*zexexE* ‘hear’, *gwEpSE* ‘forget’), which have an ergative argument (experiencer) and an absolutive one (stimulus) and have intermediate properties between transitive and intransitive verbs. Reciprocals are formed from these verbs by means of the prefix *ze-*. Reflexives are formed by means of the same affix, but the reflexive marker, in contrast to the reciprocal one, substitutes the absolutive affix. This difference reflects deep differences between reflexive and reciprocal: the reflexive derivation is “subject-sensitive” (the ergative argument of the inversive verbs has more subject properties than the absolutive one), whereas the reciprocal derivation is “absolutive-sensitive” (the reciprocal affix is always controlled by the absolutive argument).

The most intriguing variants of reciprocal construction are variants with pronominal markers *zEm zEr* and *zEm adrEr*. There are many combinations of pronouns, reciprocal prefixes and agreement markers: cf. the following variants:

- 1) The construction with the pronoun and the reciprocal form of the verb: *te zE-r zE-m tE-ze-fe-KWe-Z’E* ‘we visit each other’.
- 2) The pronoun is the only marker of reciprocity, the third person singular agreement markers for both arguments are used: *te zE-m zE-r zex-je-xE* ‘we hear each other’.
- 3) The pronoun is the only marker of reciprocity, the third person singular agreement marker of the ergative argument is used, the absolutive argument is cross-referenced in a usual way: *a-xe-r zE-r zE-m fe-KWe-Z’E-x* ‘they visit each other’ (-x is the third person plural absolutive agreement marker).

The variants of reciprocal construction reflect different stages of grammaticalization of the reciprocal pronouns. These variants are distributed according to simultaneity/non-simultaneity of the actions of the participants, transitivity of the initial verb, the number of the participants etc. Rich variance of the Adyghe reciprocal constructions lets us make conclusions about properties of reciprocal constructions in polysynthetic languages:

- 1) Reflexive-reciprocal polysemy can be avoided if reflexive and reciprocal markers occupy different slots in a verb form.
- 2) Though many specialists (cf. [Baker 1996]) suppose that pronouns and nominal phrases have properties of adjuncts in polysynthetic languages and so-called “agreement markers” are arguments, we can see that in Adyghe both agreement markers (cf. *ze-*, *zere-*) and free pronouns (cf. *zEm zEr*) have some properties of arguments.

References

- Baker M. 1996. *The Polysynthesis Parameter*. Oxford UP, New York. Nedjalkov V.P. Karachay-Balkar reciprocals. // *Turkic Languages*, 2002, 6: 19-80.
- Rogava G.V., Kerasheva Z.I. 1966. *Grammatika adygejskogo jazyka (A Grammar of Adyghe)*. Krasnodarskoje knižnoje izdateljstvo. Krasnodar – Maikop.
- Testelets Ya.G. 2003. Subject in Adyghe. Working materials for the report in the Adyghe expedition.

Typology and Universals of Comparative Correlatives

Tsz-Cheung Leung

Language always employs a particular construction which expresses a dependency relation between two conjoining clauses, as in English *the more you eat, the fatter you get*, i.e. the degree that one gets fatter depends on the degree that one eats more. Call it *comparative correlatives* (henceforth CC; Fillmore et al 1988, McClawley 1988, Culicover and Jackendoff 1999, etc). A typological study of CC reveals that it usually (i) involves the semantics of comparatives in both clauses (see the list in e.g.1), whereas this could be morphologically realized (e.g. Romance) or not (e.g. Chinese, Thai, Indonesian in e.g.2); (ii) contains parallel occurrence of *correlative markers* (henceforth CM) in both clauses (e.g.2); (iii) concatenates two clauses *paratactically* without any intervening connector (e.g.3).

Property (ii) signifies the notion of *correlatives* in this construction. First by analysing the CM as a *clausal operator* meaning *to the degree/extent that*, it explains why the CM must be sentence-initial in both clauses since it has to take scope over its c-commanded IP constituent (e.g.4). The overt movement of comparatives which attaches to the CM in some languages (e.g. English, German) suggests that the operator is modified by the comparatives, which is syntactically represented by a single constituent. Such a combination is highly rare elsewhere, e.g. *the+Adj-er* in English (e.g.5).

Second, languages differ in that the pair of CMs in both clauses could be morphologically identical to each other or not (see the list in (6)). For languages in which the two CMs are identical to each other (call it Type I languages), the second CM is a fully parallel occurrence of the first one (i.e. it is not an anaphoric expression). That the second CM is coindexed with but not syntactically dependent on the first CM leads to the hypothesis that the two clauses are connected via simple concatenation at the underlying level (property (iii)). This is verified by Condition C violation via the intervention of a pronoun between the two clauses (e.g.7).

On the other hand, for languages in which the two CMs are not identical to each other (call them Type II languages), the relation between them is always similar to that between *wh- \bar{e}* and *th- \bar{i}* in English, i.e. the first CM signals an indefinite and the second one a deitic expression (c.f. Polish *t-*, Russian *t-*, Hindi *ut-* as the second CM are morphologically deitic). This is highly similar to correlatives construction in general (Srivastav 1991, Bhatt 2003, den Dikken (to appear)). For instance Hindi correlatives are highly productive in generating all types of free relatives, relative clauses and comparative correlatives (e.g.8). Bhatt (2003) suggests that in Hindi correlatives the first correlatives CP adjoins to the second CM (i.e. the DEM-XP) and form a constituent at the underlying structure (e.g.9), verified by the coordination test (e.g.10). Coincidentally, the hypothesis that an antecedent (c.f. the first relative CP) and an anaphor (i.e. the CM in the second clause) form an underlying constituent is also observed in the analysis of antecedent-pronoun sentences in which the antecedent forms an underlying constituent with its anaphor at the underlying level (Kayne 2002, Zwart 2002; e.g.11). Analogous to the analysis in which the antecedent undergoes overt movement to some sentence-initial position at the surface level, in correlatives the relative CP undergoes a similar overt movement. Therefore the primary distinction between Type I and Type II languages is more than the morphological realization of the CM, rather it concerns the level in which the two CMs form an underlying constituent. For Type I languages, the underlying constituent is formed at the clausal (i.e. CP) level, whereas in Type II languages it is formed at the phrasal (i.e. DEM-XP) level. Such a distinction is expected: anaphoric relations (i.e. Type II languages) are more

syntactically constrained since the anaphor in the second clause has to be syntactically licensed (in this case, by forming an underlying constituent). On the other hand, in Type I languages, the semantics of the second CM is easily identified since it is morphologically identical to the first one (i.e. a fully parallel occurrence). As a result, language does not resort to a syntactic strategy (e.g. c-command, forming an underlying constituent) to resolve its semantic identity.

Examples

- (1) a. Im bardziej jesteś zmęczony, tym gorzej pracujesz (Polish)
 IM more you-are tired TYM worse you-work
 'the more you are tired, the worse you work'
- b. Chem bol'she vin, tem veseleye (Russian)
 What-INST more wine-GEN that-INST merrier
 'the more wine, the merrier'
- c. Oso pio poli diavazo, toso pio poli katalaveno (Greek)
 As-much more much I-read that-much more much I-understand
 'the more I read, the more I understand'
- (2) a. *The* more you eat, *the* fatter you get (English)
- b. ta *yue* piaoliang, wo *yue* kaixin (Mandarin)
 she CM pretty I CM happy
 'The prettier she is, the happier I am'
- c. Ni *jyut* gong, keoi *jyut* ganzoeng (Cantonese)
 You CM talk he CM nervous
 'The more you talk, the more nervous he is'
- d. Khun *ying* phuut, khaw *ying* dandeng (Thai)
 You CM talk he CM nervous
 'The more you talk, the more nervous he is'
- e. Ali *makin* besar *makin* sombong (Indonesian)
 Ali CM old CM arrogant
 'The older Ali gets the more arrogant he becomes'
- (3) *The more you eat, *I think that* the fatter you get.
- (4) a. The more he thinks that you are hardworking, the happier he is.
 (more>think)
- b. The more hardworking he thinks that you are, the happier he is.
 (more>hardworking)
- (5) *John is the taller.

(6)

Type I		Type II	
English:	the... the...	German:	je... desto...
French:	Ø... Ø...	Polish:	im... tym...
Mandarin:	yue... yue...	Russian:	chem... tem...
Cantonese:	yuht... yuht...	Hungarian:	Minél... annál...
Thai:	ying... ying...	Greek:	Oso... toso...
Indonesian:	makin... makin...	Hindi:	jitnaa... utnii...

- (7) *The more John_i eats, he_i thinks that the fatter he_i gets (c.f. *He_i thinks that the more John_i eats, the fatter he_i gets)

- (8) a. *Jitnaa* suuraj chamk-aa *utnii*-(hii) ThanD baRhii (comparative correlatives)
How-much sun shine that-much-(only) cold increase
'The more the sun shines, the colder it gets'
- b. *jo* larkii kharii hai vo lambii hai (relative clause)
REL girl standing is DEM tall is
'which girl is standing, that girl is tall'
- c. *Jo* aapne banaayaa, *vah* meine khaayaa. (free relatives)
Which you-erg make-pst that I eat-pst
'I ate what you cooked'
- (9) [which CD is on sale], [Ram bought [_{t_i} that CD_i]]
(c.f. Ram bought the CD which is on sale)
- (10) Rahul a:jkal [_{DP} [_{DP} [_{jo} kita:b Saira-ne likh-i:] vo1] aur [_{DP} [_{jo} cartoon Shyam-ne bana:-ya]2
R nowadays REL book S-erg write-perf-F DEM and REL cartoon S-erg make perf
vo2]] parh raha: hai
DEM read prog be-pres
'Nowadays, Rahul is reading the book that Saira wrote and the cartoon that Shyam
made' (Lit. 'Nowdays, Rahul is reading [[which book that Saira wrote] that (book)]
and [[which cartoon that Shyam made] that (cartoon)]
- (11) thinks [John he] is smart --> Johni thinks [_{t_i} he] is smart

The Typology of Reported Speech: Semi-Indirect Speech in Papuan Languages

Robyn Loughnane

The traditional typology of reported speech¹⁴ is still being expanded and fully understood. Reported speech has traditionally been divided into several subcategories which represent the different styles used to report the speech and thought of others. These are: direct reported speech, indirect reported speech and free indirect reported speech.

In this presentation, I will attempt to expand on the typology of reported speech by describing a style of speech found in a number of Papuan languages. The presentation will focus on data from the Chimbu language Golin. The particular combination of spatial deixis, mood marking and person deixis in this style of reported speech make it fall outside the traditional types of reported speech listed above. The mood marking is like that traditionally associated with indirect speech, the spatial deixis is like that traditionally associated with direct speech, and the person deixis appears to be primarily direct speech style but with a few interesting alterations. I will also refer to previous work done by Resink (1987, 1993) and Haiman (1980) on the topic.

References

- Coulmas, F. E. (1986). Direct and indirect speech. Berlin, Mouton de Gruyter.
- Güldemann, T. and M. von Roncador (2002). Preface. Reported discourse : a meeting ground for different linguistic domains. T. Güldemann and M. von Roncador. Amsterdam ; Philadelphia, PA, J. Benjamins Pub. Typological studies in language, v. 52.
- Haiman, J. (1980). Hua : a Papuan language of the eastern highlands of New Guinea. Amsterdam, John Benjamins.
- Longacre, R. E. (1994). The dynamics of reported dialogue in narrative. *Word* 45(2): 125-143.
- Mathis, T. and G. Yule (1992). The role of staging and constructed dialogue in establishing speaker's topic. *Linguistics* 30: 199-215.
- Reesink, G. P. (1987). Structures and their Functions in Usan. Amsterdam, John Benjamins.
- Reesink, G. P. (1993). "Inner Speech" in Papuan Languages. *Language and Linguistics in Melanesia* 24: 217-25.

¹⁴ This term is used by e.g. Coulmas (1986). The phenomenon is also known as constructed dialogue (Mathis and Yule 1992), reported dialogue (Longacre 1994), reported discourse (Güldemann and von Roncador 2002)

Example data from Golin (Chimbu):

- (1) *abal i [awi na si-m-u-a] di-n-g-w-e*
 woman TOP dog 1sg hit-3-REP-DIST say-3-AS-3-PROX
 She_i said that the dog bit me_j.
 or
 She_i said that the dog bit her_i. / She_i said “the dog bit me_i”.
- (2) *[na i yal-su si-ø-w-a] di-n-g-e*
 1sg 2 man-two hit-1sg-REP-DIST say-2-AS-PROX
 You_i said you_i hit us_j (2)./You_i said “I_i hit you_j (2)”
- (3) *yal i [na na si-m-u-a] di-n-g-w-e*
 man TOP 1sg 1sg hit-3-REP-DIST say-3-AS-3-PROX
 He_i said I_j hit him_i.
- (4) *yal kane [inin er-o-m-u-a] di-n-g-w-e*
 man many 1pl do-go-3-REP-DIST say-3-AS-3-PROX
 They said we are going.
- (5) *yal i [wan wai n-a-m-a] de] pru-n-g-w-e*
 man TOP son good go-IRR-3-DIST say think-3-AS-3-PROX
 He_i wants his son_j to go well.
- (6) *na [yal i bia yo-m-a] di] pri-ø-g-e*
 1sg man TOP old be-3-DIST COMP think-1sg-AS-PROX
 I assume he is old.
- (7) *yal i [i si-ø-w-a] di-n-g-w-e*
 man TOP 2 hit-1sg-REP-DIST say-3-AS-3-PROX
 He_i said he_i hit you_j. / He_i said “I_i hit you_j”
- (8) *ebal benamble [solwara kar-a-l-a] pri-m=ba*
 person many saltwater(TP) see-IRR-1sg-DIST think-3=CONC
 Although many people want to see the sea... /Many people think “I will see the sea”, but...

Abbreviations:

1	First person
2	Second person
3	Third person
AS	Assertive
COMP	Complementizer
CONC	Concessive
DIST	Distal
IRR	Irrealis
PROX	Proximal
REP	Reported speech
sg	Singular
TOP	Topic
TP	Tok Pisin

Verb Type Hierarchies and a Typology of Transitivity Splits

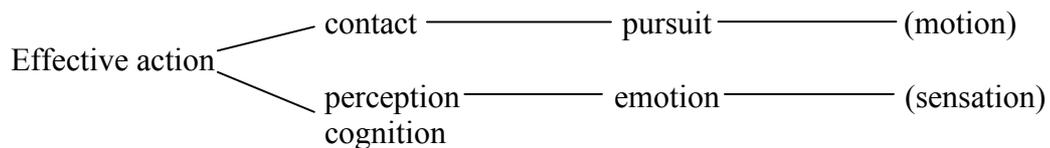
Andrei Malchukov

The present paper deals with a typology of ‘transitivity splits’, that is, variation in case marking conditioned by the verb class of the predicate. As a starting point I take the verb type hierarchy in (1) proposed in publications by Tsunoda (Tsunoda 1981; Tsunoda 1985), which is intended to predict distribution of intransitive and transitive patterns in individual languages:

- (1) Effective action >> Perception >> Pursuit >> Knowledge >> Feeling >> Relation

While it is generally recognized in the literature (cf., e.g., Drossard 1991; Lazard 1998) that this hierarchy correctly captures the general tendency (transitive encoding becoming less likely as we move down the hierarchy), the strict ordering of verb types in the central part of the hierarchy (e.g., pursuit verbs vis-à-vis mental verbs) has been questioned (Lazard 1998). In my paper I argue that this hierarchy will gain in semantic coherence and predictive power if it will be recast in the form of a two-dimensional semantic map, as in (2).

- (2)



In the first part of the paper I present cross-linguistic data from languages of different alignment type in support of the particular “routes” (sub-hierarchies) on this map.

In the second part I show how the proposed two-dimensional hierarchy can be used to predict which case frames would be preferred by different verb types. Building on the previous work in the functional-typological tradition (in particular, on the ‘competing motivation model’ of DuBois 1985) and optimality-theoretic tradition, I propose a set of universal, functionally based constraints that can account for cross-linguistic variation in case-marking patterns of individual verb types. The constraints proposed include ‘Faithfulness’ to semantic role (as manifestation of the indexing function of case marking), Distinctness (as manifestation of the distinguishing function of case-marking), Unmarked Case Constraint (as in Tsunoda 1981), that requires that every clause contains one unmarked ‘primary’ argument, and Transitive Default (as in Primus 1999), on the assumption that the transitive pattern is an unmarked construction for two-argument verbs (cf. Lazard 1998). Applying this analysis to a wide range of languages with different alignment, it is shown that it can correctly predict the available (resp. non-available) case marking patterns for individual verb types. In particular, it accounts for an asymmetry between accusative and ergative languages in syntactic encoding (i.e. the choice of case frames) of the pursuit type, their similarities in encoding of the emotion type, as well as resolves a controversial issue of the relative predisposition of emotion and sensation predicates for the non-canonical subject marking (see Haspelmath 2001 and Onishi 2001 for opposing views on this issue). Some apparent counterexamples to the proposed analysis are attributed to other interfering factors related to the word class of the predicate and pattern inheritance attested for polysemous items. In conclusion I integrate the proposed two-dimensional

hierarchy into a more comprehensive semantic map showing further connections between the transitive and intransitive domains.

References

- Aikhenvald, A., R.W.M. Dixon & M. Onishi (eds.). (2001). *Non-canonical marking of subjects and objects*. (Typological studies in languages, 46). Amsterdam: Benjamins.
- Drossard, W. (1991). Verbklassen. In: Seiler, H. & W. Premper (eds.), 150-182.
- DuBois, J.A. (1985). Competing motivations. In: Haiman, J. (ed.) *Iconicity in syntax*. Amsterdam: Benjamins, 343-66.
- Haspelmath, M. (2001). Non-canonical marking of core arguments in European languages. In: Aikhenvald et. al. (eds.), 53-85.
- Lazard, G. (1998). *Actancy*. Berlin: Mouton.
- Onishi, M. (2001). Non-canonically marked subjects and objects: Parameters and properties. In: Aikhenvald et. al. (eds.), 1-53.
- Primus, B. (1999). *Cases and thematic roles*. Tübingen: Niemeyer
- Tsunoda, T. (1981). Split case-marking in verb types and tense/aspect/mood. *Linguistics*, 19, 389-438.
- Tsunoda, T. (1985). Remarks on transitivity. *Journal of Linguistics* 21, 385-396.

Complexity in Functional Domains

Matti Miestamo

Language complexity has recently become a major issue in linguistic research (e.g. McWhorter 2001, Kusters 2003, Dahl 2004). The metric of overall grammatical complexity introduced by McWhorter (*ibid.*) pays attention to overt signalling of phonetic, morphological, syntactic and semantic distinctions beyond communicative necessity: a grammar is more complex than another to the extent that 1. its phonemic inventory has more marked members, 2. its syntax requires the processing of more rules, 3. it gives overt and grammaticalized expression to more fine-grained semantic and/or pragmatic distinctions, and 4. to the extent it uses inflectional morphology. The metric may be suitable for the specific purpose it was designed for, i.e. to reveal very clear complexity differences, but in general, measuring overall grammatical complexity is an extremely difficult task. All such metrics are affected by two general problems: the problem of representativity, i.e. it is impossible for any metric to exhaustively take into account all relevant aspects of grammar, and the problem of comparability, i.e. the contributions of the complexities of different areas and aspects of grammar to overall complexity are incommensurable. The problem of comparability affects all levels of linguistic structure and thus renders any attempt to compare languages in terms of overall complexity impossible (except perhaps when complexity differences are extremely clear). Instead of trying to address the the question of overall complexity directly, cross-linguistic studies of grammatical complexity must be primarily concerned with the complexity of specific aspects of grammar.

Functional domains provide a *tertium comparationis* for cross-linguistic studies of grammatical complexity. It can be argued that of two languages that grammaticalize a given functional domain, one that makes a larger number of distinctions within the domain is more complex in this respect than the other, e.g. the functional domain of aspect is more complex in a language that makes five aspectual distinctions than in another language that makes only two. In other words, a language that splits a given area of conceptual space into a larger number of parts is more complex in this respect than a language that makes fewer divisions in the same area of conceptual space. The number of grammaticalized distinctions figures as one of the criteria of complexity in McWhorter's metric as well (criterion 3 above), and a related principle is found in Kusters (2003).

In this paper I will present results of a typological study of grammatical complexity that pays attention to the grammaticalized distinctions languages make in different functional domains. The study is based on a sample of 50 languages stratified genealogically and geographically as well as in terms of a few relevant typological and sociocultural parameters. Some theoretical and methodological problems involved in the identification of the functional domains and the distinctions grammaticalized within these domains will be addressed. A number of functional domains grammaticalized by each sample language will be examined (e.g. tense, aspect, mood, deictic reference, number, gender, argument coding, etc – depending on which domains each language grammaticalizes), and cross-linguistic observations are then made about the complexity of specific functional domains. Statistical correlations between the complexity or implicitness (or non-grammaticalization) of different functional domains will be examined; this will bring some light to the question whether complexity in one area is compensated by simplicity in another. Correlations will also be sought between the complexity of functional domains and other aspects of grammar (morphological complexity and morphological type, word order, etc).

In addition to the investigation of the distinctions made within functional domains on the paradigmatic axis, the study also involves an investigation of how these distinctions are realized on the syntagmatic axis. Short texts from the sample languages are examined and attention is paid to the overt grammatical expression of different functional domains, e.g. how densely different functional domains find grammatical expression in running text. Again, correlations are examined both within and across languages.

References

- Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Studies in Language Companion Series 71. Amsterdam: Benjamins.
- Kusters, Wouter. 2003. *Linguistic Complexity, the Influence of Social Change on Verbal Inflection*. Utrecht: LOT.
- McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5:125–66.

The Functions of Case: Unifying the Discriminatory and Indexing Approaches

Åshild Næss

Current typological theory is divided in its views on the primary function of case-marking on core arguments. The perhaps most widely accepted “discriminatory” view argues that case-marking is essentially a device used to distinguish clearly between the arguments of a clause – that is, to clarify which of the arguments is the subject and which is the object. The “indexing” view, on the other hand, claims that case-marking reflects certain semantic properties of the case-marked argument, such as affectedness (for accusative case) or control (for ergative case).

In this talk I will argue that an adequate theory of core case-marking must recognise that case-marking has **both** discriminatory and indexing functions. I will present data from a number of languages where case-marking appears to discriminate not between subjects and objects, but between agents and patients. In other words, case has a discriminatory function, but it discriminates between semantically rather than structurally defined entities.

The data in question concerns languages where case-marking of one argument depends on semantic properties not of the case-marked argument but of the **other** argument of the clause – or in one case (passives in Marwari) on an argument not even overtly present in the clause. The general tendency in such languages is for marking of the patient/object to depend on the volitionality of the agent – the object gets marked only when the agent is construed as acting volitionally – or for marking of the agent/subject to depend on the referentiality or affectedness of the object, so that ergative case only applies in clauses with an affected object.

Such marking patterns are difficult to explain either from a purely discriminatory or a purely indexing point of view. It is not at all clear why a volitional agent would be more in need of overt distinction from the object than a nonvolitional one. Similarly, these case-markers do not appear to index any properties of the case-marked arguments, and it is unclear why case-marking on one argument should index properties of another. I will argue that the solution lies in recognising that case-marking in these languages depends on semantic transitivity, that is, on the presence in the clause of both a semantic agent and a semantic patient. Accusative case marks the patient when it is opposed to a semantic agent in the clause, but not necessarily when no agent is present, and vice versa for ergative case. That is, case-marking in these languages discriminates between agents and patients rather than between subjects and objects.

This view of case-marking as comprising both a semantic and a discriminatory dimension not only accounts for the problematic data discussed, but also provides a framework within which to account for crosslinguistic variation in case-marking systems. From a basic function of discriminating between semantically defined entities, case-marking can potentially be extended in two directions: along the semantic dimension or along the discriminatory dimension. In the first case, case-marking would be extended to apply to certain semantically specified entities independently of the presence of other types of entities in the clause – for instance, ergative case might apply to semantic agents regardless of whether the clause also contains a patient. The extreme case of such semantic extension is split-intransitive languages, where case-marking applies to semantically agentive (and/or patientive) arguments regardless of whether there is a second argument in the clause at all.

Extension along the discriminatory dimension would mean that case-marking would apply in all or most cases where two arguments need to be distinguished, regardless of whether or not they are semantic agents and patients – or in the extreme case, **only** where there is a need for the overt discrimination between two arguments. Some languages come quite close to such a purely discriminatory system (e.g. Malayalam) – but most languages with case probably fall somewhere in between a purely discriminatory and a purely semantically based case-marking system.

Selected data

- (1) Marwari (Indo-Aryan; Magier 1985)
- | | |
|---|---|
| a. Rām mārīj-iyo
Ram kill-PASS
'Ram was killed
(intentionally or otherwise)' | b. Rām- ne mārīj-iyo
Ram-NE kill-PASS
'Ram was murdered' |
|---|---|
- (2) Chepang (Tibeto-Burman; Caughley 1982)
- | | |
|---|--|
| a. pū-nis-i h□w sat-aka-c-u
OBro-DU-AG YoBro kill-PAST-DU-AG
'The two older brothers killed the younger brother (accidentally)' | b. h□w- kay pū-nis-i sat-a-th□y
YoBro-GL OBro-DU-AG kill-PAST-GL
'The two older brothers killed the younger brother (on purpose)'
(‘GL’ = ‘goal marker’) |
|---|--|
- (3) Shipibo-Conibo (Panoan; Kittilä 2002)
- | | |
|--|---|
| a. ea-Ø-ra yapa-Ø pi-kas-ai
I-ABS-AS fish-ABS eat-DES-INCOMPL
'I want to eat fish (referential or nonreferential)' | b. e- n -ra yapa-Ø pi-kas-ai
I-ERG-AS fish-ABS eat-DES-INCOMPL
'I want to eat fish (referential only)' |
|--|---|
- (4) Niuean (Polynesian, Seiter 1980)
- | | |
|---|--|
| a. Kua mohe a ia he fale
PERF sleep ABS he in house
'He has slept in the house' | b. Kua mohe e ia e timeni
PERF sleep ERG he ABS floor
'He has slept on the floor (some effect on the floor implied)' |
|---|--|
- (5) Bunuba (Bunuban, Rumsey 2000)
- Nyirraji-**ingga** бага'ranganggi
this one-ERG sleep 3sg-ra-PRES-2sg.OBL
'This one is sleeping with your wife (lit. sleeping, affecting you)'

Rebuilding Semantic Maps

Heiko Narrog, S. Ito & E. Tokita

The basic idea of semantic maps is to represent cross-linguistically documented similarity of meanings and the diachronic relationship between these meanings. This idea has been around since Anderson and the early 80s. However, particularly in the past five years or so, there has been a noticeable surge in interest in semantic maps which has manifested itself in an increasing number of publications. Currently, two of the most prominent approaches are those by Croft (2001 and others.), who highlights the idea of conceptual “space” represented in the semantic map and “regions” within that space, and the more traditional approach, which highlights connections between meanings. The latter has been given explicit theoretical footing in Haspelmath (2003).

The appeal of semantic maps lies in the fact that they reveal universal relationships between meanings. Therefore, if the research enterprise proves to be successful, they may open a unique window on human conceptualization. On the other hand, the validity of semantic maps is limited at present by some serious methodological problems. As it stands now, the structure of the semantic map drawn relies on the knowledge and experience of the individual researcher. This inevitably leads to occasional mistakes in the maps, such as meanings being overlooked or misplaced. Furthermore, the way a particular map is spatially arranged depends almost entirely on the personal preferences of the researcher. Therefore, a map of the same semantic area can look totally different depending on who has drawn it.

The most desirable solution to these problems is the development of a computer program that, based on explicit principles, would transform data input from a sample of languages automatically into an optimal semantic map. Unfortunately, this is technically not feasible since the drawing of an optimal semantic map is at best a NP-complete problem, that is, theoretically possible but computationally intractable. Very recently, Croft (unpublished research) has developed a powerful alternative to semantic maps, namely the application of multidimensional scaling. This method allows to reliably represent the similarity of meanings while the connection component of the semantic maps is being lost. In our presentation, we present an attempt at another possible solution. Taking a specific meaning as the point of departure (in this case “instrumental”), and working with a DV sample of 200 languages, we developed a method by which the implicational universals connected with this meaning are automatically extracted. The implicational universals that are identified can then serve as the building blocks to construct a semantic map. This method does not solve the problem of how to draw the optimal semantic map, but it provides a tool to verify the implicational universals whose existence is claimed in the map.

Nias in Typology Perspective

Waozisokhi Nazara

Nias is an Austronesian language spoken by around 700,000 people (Brown 2001), most of whom live on Nias island, which lies about 100 kilometres west of Sumatera island. This vocalic language (Nazara 2001) has three main varieties: (1) northern variety, (2) central variety, and (3) southern variety. Amongst the three, northern variety has most speakers. (Most) speakers of the other two varieties can even speak or at least understand the northern variety, from which the data of this analysis have been gathered.

Typologically Nias performs some interesting characteristics. In terms of word order, a clause in this language positions the predicate before the subject. The predicate, which precedes the grammatical subject, subcategorizes for the grammatical object after the verb. It holds then that Nias is a VOS language.

Morphologically Nias is characterized by at least two interesting facts. First, a pronominal prefix is attached to the verb requiring two NPs (subject and object). This pronominal prefix reveals the person (first, second, or third) and the number (singular or plural) of the NP functioning as the subject. Second, generally ‘who does X to whom’ can identified be referring to the morphological form of the NP occupying argument position in a clause. The S argument in an intransitive clause is marked the same as the P argument in a transitive clause, while A argument is marked differently. A language behaving this way is commonly labeled as morphologically ergative.

In some circumstances, however, where a prefix *ma-* or *mo-* rather than a pronominal prefix is attached to the verb even though the verb subcategorizes for two arguments. Prefix *ma-* and *mo-* do not reveal the person or number of the NP being positioned as the A argument. In such occurrences the A argument of a transitive clause is morphologically marked the same as the S argument of an intransitive clause, the P argument is marked differently. Such a language as this is generally said to be morphologically accusative.

Syntactically Nias seems to behave like a S/A Pivot language, in the framework for pivot investigation proposed by R.M.W. Dixon (1994). This language tends to group S and A together rather than S and P in a coordinate clause. If the P argument of the second clause is the same as the S argument of the first clause, the coordinate clause resulted from the two clauses tends to be resistant to the deletion of the P argument. The reverse works if it is the A argument of the second clause is the same as the S argument of the first clause. If the P1 = S2, the derived clause can be accepted despite the S2 is overtly deleted.

Interestingly, however, both A and P arguments can be relativised. This leads me to believe that Nias tends to be a fertile soil for discussions in typological perspective. Are there any colleagues who would like to share ideas about this?

Data

1. Mörö Döngöni. ‘Töngöni sleeps’
2. Tohare Tolamanö. ‘Tolamanö comes’
3. Tohare Dolamanö ba mörö Döngöni.
4. Tohare Dolamanö ba ibözi Döngöni. ‘Tolamanö comes and hits Töngöni’
5. Tohare Gayaso ba ibözi ia/*Ø Gadödö. ‘Gayaso comes and Gadödö hits him’
6. Itifa Dehe’aro Tahadödö ba me’e (Dehe’ao) ‘Tahadödö kicks Tehe’aro and Tehe’aro cries’
7. Itifa Dehe’aro Tahadödö ba mofanö ia ‘Tahadödö kicks Tehe’aro and leaves’
8. I’usu Dehemanö Tolamanö ba itifa Tahadödö. ‘Tolamanö bites Tehemanö and Tahadödö kicks Tehemanö’
9. Itezu ba itifa Dehemanö Töngöni. ‘Töngöni punches and kicks Tehemanö’
10. Manezu ba manifa Tehemanö Döngöni. ‘Töngöni punches and kicks Tehemanö’
11. Ibözi Dehemanö Tolamanö ba itifa Döngöni Tehemanö.’Tolaman ö hits Tehemanö and Tehemanö kicks Töngöni’
12. Ibözi Dehemanö Tolamanö ba itifa Dolamanö/*Ø Tahadödö. ‘Tolamanö hits Tehemanö and Tahadödö kicks Tolamanö’
13. Ibözi Dehemanö Tolamanö ba itifa Dolamanö Tehemanö. ‘Tolamanö hits Tehemanö and Tehemanö kicks Tolamanö’
14. Ibözi nono alawe si no mogete ono matua amagu. ‘My father hits the girl who has pinched a boy’
15. Ibözi nono matua si no ikete ono alawe amagu. ‘My father hits the boy who has been pinched by a girl’

Against a Parameter-Setting Model of Typological Variation

Frederick Newmeyer

In standard approaches to generative grammar, the theory provides principles of Universal Grammar (UG), which have (ideally, a small number of) different parameter settings for different languages. The Head Parameter, for example, posits that complements are consistently to the left or to the right of the head. English and Swahili, say, would choose the parameter setting HEAD-LEFT, while Japanese and Lakhota would choose HEAD-RIGHT. As far as ‘typologically inconsistent’ languages and typologically rare phenomena are concerned, the usual assumption has been that both rarity and inconsistency are reflexes of ‘less-than-fully-optimal’ grammars. That is, typologically rare grammars are more complex than grammars of typologically consistent languages, perhaps by requiring special marked parameter settings or by violating a hierarchical relation among parameter settings. Furthermore, parameters and their settings are considered to be abstract enough to capture seemingly unexpected clusterings of typological properties (the voluminous literature on the Null-Subject Parameter provides examples). In this talk I develop a number of arguments against a parameter-setting model of typology, among which are the following:

- To capture the full complexity of any language, the number of parameter settings would have to be as great as the number of rules in a non-parametric approach, undercutting an economy-based argument for parameters.
- There is little or no ‘abstract clustering’. The much-vaunted null-subject correlations do not exist. The clusterings of typological properties that do exist (e.g. the many correlations with VO or OV order) are quite close to the surface.
- Typologically rare grammars are not necessarily more complex grammars. Preposition-stranding (e.g. *Who did you talk to?*) is one of the most typologically rare phenomena. Yet it is easy to demonstrate that stranding grammars are less complex than non-stranding ones.
- Typological generalizations are probabilistic. VSO languages are more likely to have *wh*-movement than SVO languages, which are more likely to have it than SOV languages. That sort of statistically-phrased generalization is incompatible with the ‘algebraic’ nature of UG.

I argue (briefly, given time limitations) that the most robust typological generalizations have an explanation based in language processing. In other words, UG provides the set of possible grammars of human languages, while performance considerations determine why some of those grammars are crosslinguistically more common than others.

Grammatical and Semantic Variation in Serial Verb Constructions

Melanie Owens

Serial Verb Constructions (SVCs) are monoclausal utterances containing sequences of verbs with identical tense, aspect, mood, and polarity categories, where there is no marking of coordination or subordination between the verbs. I propose an additional definitional property: letting *X* be the list of thematic roles selected in a construction *A*, if *A* is a SVC, then there exists in some language a construction *B* which contains a single verb, and which selects *X*. Consequently, ‘covert coordinations’ such as (1) are excluded from the class of SVCs, since the object of each verb has the thematic role of *patient*, and in no language does there exist a clause containing a single verb whose thematic role list duplicates the patient role. In the recent formal literature, SVCs are identified on the basis of *internal argument sharing* (Baker 1989, Collins 1997, Stewart 2001, etc.), according to which (1) is excluded from the class because each of the two verbs selects a distinct internal argument, rather than sharing a single internal argument. Internal argument sharing, its deficiencies as a defining property and how these impact upon currently accessible crosslinguistic SVC data will be revisited in conclusion; the main purpose of this paper is to lay out types of cross-linguistic grammatical variation in the class of SVCs, and to evaluate whether these correlate with semantic properties of SVCs.

SVCs exhibit two main types of grammatical variation, in terms of (a) contiguity status, and (b) argument sharing. Contiguity status refers simply to the possibility for non-verbal material to intervene between the verbs (as in non-contiguous (2)), or whether the verbs must be linearly adjacent (as in contiguous (3)). Crowley (2002) provides the most comprehensive enumeration of argument sharing schemes in SVCs, distinguishing between same subject sharing, switch subject sharing, inclusory serialization, multiple object serialization and ambient serialization. Does any of this grammatical variation correlate with semantic properties of SVCs? I approach this question from the angle of event structure, and in particular Aikhenvald’s description of a SVC event structure continuum which runs from single indissoluble events to looser packages of subevents, and her accompanying claim that “the place of a serial verb on [this] continuum correlates with grammatical parameters – such as contiguity ... of components, and argument sharing” (2003, 17). Do non-contiguous SVCs with freer argument sharing schemes allow looser event structures than contiguous SVCs with tighter argument sharing restrictions? ‘Looseness’ is here identified in terms of independence of the actions described by each verb, time lapse between subevents, and productivity of verb combinations in SVCs.

The primary data for this paper is from Bimanese, a Central Malayo-Polynesian language with contiguous SVCs whose loosest possible combinations, exemplified in (4)-(5), consist of an initial verb describing a preliminary action of motion or positioning, followed by a verb expressing the main action. Compare this with the looser event structures permitted in the non-contiguous SVCs of Kwa (and Kwa substrate creole) languages in (6)-(8); for these combinations of transitive verbs ‘anything pragmatically apt goes’ (Veenstra 1996, 103), and there are no Bimanese SVC equivalents. Does the freer argument sharing and noncontiguity of Kwa SVCs account for their looser event structures? I suggest that this is not a promising solution, since Haitian and Papiamentu, for example, have non-contiguous SVCs with freer argument sharing, yet very tightly constrained event structures. Nonetheless, the contrast in permissible event structures for Bimanese and Kwa SVCs raises some

pertinent questions. Can the looser event structures of Kwa SVCs be related to, or predicted from, other grammatical properties of these languages (such as their phrase structures or inventories of coordinating elements)? Do Bimanese and Kwa SVCs really instantiate the same grammatical phenomenon? Or, does there exist some property which could exclude these looser Kwa structures from the class of SVCs? I know of no such property, but an important point is that examples such as (6)-(8) appear in force in the formal literature by virtue of fulfilling the internal argument sharing property. Alarming, (4)-(5) do not fulfill this property, and so would be excluded from the class of SVCs. Current typological SVC research is subject to the lack of both accessible data and unanimous definitions for this multi-faceted phenomenon. When extracting data from the current formal literature, with its preoccupation with internal argument sharing, it must be borne in mind that the latter property arguably obscures the SVC phenomenon and skews its true distribution.

Examples

- (1) Òzó gb□□ ìvín bolo kà. (dó; Stewart 2001, 51)
 Ozo plant coconut peel corn
 ‘Ozo planted coconut and peeled corn.’
- (2) A tei goni suti di pingo. (Saramaccan; Byrne 1985, 281)
 3SG take gun shoot DET pig
 ‘He shot the pig with a gun.’
- (3) Amu hanta wa’a.na meja ese uma. (Bimanese)
 Amu pick-up carry.3 table up-to house
 ‘Amu (picked up and) carried the table up to the house.’
- (4) Rao lu’u ka.maru weki.na. (Bimanese)
 Rao enter CAUS.sleep REFL.3
 ‘Rao went in and slept.’ (lit. ‘made herself sleep.’)
- (5) Fero tu’u reb’o.na. (Bimanese)
 Fero get-up dance.3
 ‘Fero got up and danced.’
- (6) Ò d LGB tié. (dó; Stewart 2001, 49)
 Ozo buy LGB read
 ‘Ozo bought LGB and read it.’
- (7) Ajé jí aš gbé w□. (Yorùbá; Awóyalé 1988, 3)
 Aje steal dress take wear
 ‘Aje stole some clothes and put them on.’
- (8) A kisi dí fou náki kii limbó bóí njan. (Saramaccan; Veenstra 1996, 103)
 3SG catch DET bird hit kill clean cook eat
 ‘He caught the bird, struck it dead, cleaned, cooked and ate it.’

References

- Aikhenvald, A. 2003. *Serial Verb Constructions*. Ms. at: <http://www.latrobe.edu.au/rclt/Workshops/2003/position%20paper.pdf>
- Awóyalé, Y. 1988. *Complex Predicates and Verb Serialization*. Cambridge, MA: Lexicon Project Working Papers 28.
- Baker, M. 1989. 'Object Sharing and Projection in Serial Verb Constructions'. *Linguistic Inquiry* 20: 513-553.
- Byrne, F. 1985. *Verb Serialization and Predicate Complementation in Saramaccan*. Ph.D. dissertation, University of Arizona.
- Collins, C. 1997. 'Argument Sharing in Serial Verb Constructions'. *Linguistic Inquiry* 28: 461-497.
- Crowley, T. 2002. *Serial Verbs in Oceanic*. Oxford: Oxford University Press.
- Stewart, O. 2001. *The Serial Verb Construction Parameter*. New York: Garland.
- Veenstra, T. 1996. *Serial Verbs in Saramaccan*. Dordrecht: HIL Dissertations.

This paper presents the results of a typological survey of possessive subject-indexing in 19 NWS languages, characterising the functional and structural characteristics of the relevant constructions in each language. The paper also presents evidence on a distinction between dominant and subordinate possession and on the possession of adverbial nominalised clauses in synchronic NWS, revealing a complex formal relationship between subjects and possessors in this group of languages.

Evidence for a Major Typological Shift at the Dawn of the History of Dravidian Languages Christiane Pilot-Raichoor

Dravidian languages enter in history a little before the Christian era through records in the Tamil language (two main sources: a poetic corpus, known as the *Sangam* (a compilation of anthologies) and epigraphic records from the 2nd century BC, cf. Mahadevan 2003). For more than half a millennium, Tamil remained the only attested language (next is Kannada, with its first inscription in 450AD) and literary works did not appear in other languages before the turn of the first millennium. Most of the other – unwritten- languages of the family were only recorded in the 19th – 20th centuries.

The development of Dravidian comparative linguistics since the middle of the 19th century (Caldwell 1856) has brought out the picture of a fairly unified family of languages with a well established common vocabulary (Burrow & Emeneau 1984) and a bulk of common grammatical features: strong SOV order, highly developed ‘agglutinative’ morphology with a clear noun/verb distinction (based on case system for the noun and finite/non finite forms for the verb), use of converbs in complex syntax, etc. The results of the comparative works lead to the reconstruction of a Proto-Dravidian grammar which, broadly speaking, typologically resembles the features of the modern languages (Krishnamurti 2003).

The point I wish to raise here is that the presumed typology reconstructed for Proto-Dravidian is at odd with the Early Old Tamil data actually recorded centuries prior to any other Dravidian languages. The main differential features are: (i) use of bare lexical stem (for would be ‘nouns’ as well as would be ‘verbs’) in all functions (reference, modification and predication), (ii) prevalence of semi-free words used as post-positions instead of case system, (iii) use of polyfunctional ‘specializing particles’ (Caldwell) which would become the base of the morphological constructs, (iv) no clear grammatical distinction between noun and verb forms, etc. The specific features of early old Tamil prompted the early dravidianists and specialists of Old Tamil to suggest, as Zvelebil explicitly states it, that Tamil has gone through an important typological shift :

“It can be supposed that the period preceding the EOT [Early Old Tamil] stage was in a state similar to ‘isolation’ while the EOT period [the first historically attested stage of development] had a transient character with disappearing traces of isolation, typical features of agglutination and nascent features of inflection” (Zvelebil 1967:37)

The idea of a typological shift has been shadowed by the development of comparative grammar and the fact that the earliest recorded data are already in a ‘transient’ state, with traces of the previous typology as well as occurrence of most of the formatives or even fully established forms of later morphological constructions. These formal resemblances gave rise to a fallacious tendency to assign modern values (for instance in terms of tense or case) to the formatives of the older constructions – which, as many authors (Caldwell, Ramaswami Aiyar 1938:769, Rajam 1992:933...) alert us, they probably did not have originally.

The aim of my communication is to illustrate the specific features of Early Old Tamil and to put into correspondence the well established findings of comparative Dravidian linguistics and the internal evolution of Tamil in order to show that what is presumed to be ‘Proto-Dravidian’ grammatical typology fits actually with a middle stage of development in Tamil. This frame of typological evolution allows reinterpretation of the earlier formatives and offer consistent explanations for the

development of the Dravidian morphology. The question raised is: Should we assign to Proto-Dravidian the typology evidenced by early old Tamil data (i.e. an isolating language with unclear noun-verb distinction) and what are the consequences for long range comparison of the language families of the region?

References

- Burrow, T; & Emeneau M, 1984, *A Dravidian Etymological Dictionary*. Clarendon Press, Oxford.
- Caldwell, R. [1956/ rep.1976] (1856 1st ed., 1875 2nd ed., 1913 3rd ed. revised by J. L.Wyatt and T. Ramakrishna Pillai ; reprinted) *A comparative grammar of the Dravidian or South-Indian family of languages*. Madras: University of Madras.
- Krishnamurti, Bh.. 2003. *The Dravidian languages*. Cambridge: Cambridge University Press.
- Mah Devan I., 2003, *Early Tamil Epigraphy. From the Earliest Time to the Sixth Century A.D.* Chennai, Cre-A & Harvard University, USA.
- Rajam, V.S. 1992. *A Reference Grammar of Classical Tamil Poetry (150 B.C-pre-fifth/sixth century A.D.)*. Philadelphia: American Philosophical Society.
- Ramaswami Aiyar, L.V. 1938. The Morphology of the Old Tamil Verb. *Anthropos* 33:747-781.
- Zvelebil, K., Y.Y. Glasov et M. Andronov. 1967. *Introduction to the Historical Grammar of the Tamil Language. Pt. 1: Preliminary Textual Analysis*. Moscou: Nauka.

Stories behind Implications:
On the Relationship between Typology and Diachrony
Frans Plank

Typology is about variation across languages, diachrony is about variation in time; neither typological nor diachronic variation appears to be unlimited. The question is how constraints on variation across languages and in time bear upon each other. Two alternative answers have been given, attributing (possibly exclusive) responsibility to either typology (a) or diachrony (b):

- (a) Universals, seen as timeless laws, impose (absolute or conditional) limits on variation across languages at any and all times; they thereby constrain change insofar as a language must not change so as to violate such a universal, or at any rate not without subsequent changes swiftly redressing the balance. (There are no laws of change itself.)
- (b) Particular targets (forms, categories, constructions, rules, constraints) can only result by particular mechanisms of change operating on particular sources; such laws of change thereby impose limits on how languages can differ: they can only be what they could become. (There are no timeless universals; co-variation is due to co-evolution.)

The view that has diachrony in charge (b) is the more traditional one; for a while superseded by position (a), it has lately been regaining popularity, inspired by Greenberg's programme to "dynamicise" typology. In the case of implicational constraints, the most straightforward way of dynamicization has been to read "implies" as "derives from" (e.g., Prep NP \supset / $<$ N Genitive, NP Postp \supset / $<$ Genitive N; that is, when head nouns in attributive constructions are grammaticalized as adpositions, if nothing else happens, they will be remain in the same position relative to the NPs they are in construction with).

My aim here is to caution against too facile reductions of typology to diachrony. Implicational statements of constraints on co-occurrences are shorthand, and their diachronic translations are in danger of missing the connecting stories they abbreviate. These stories are to do with mental representations of grammars and lexicons: it is these that are variable across languages and in time, subject to constraints which typology and/or diachrony have to account for. Just as there are constraints on REanalyses, forcing different representations of the same primary linguistic data upon successive generations, there arguably are constraints proscribing any differences in representations across generations.

For illustration I propose to look at a constraint on affix positioning relative to stems that has alternatively been given a diachronic reading (b), supposedly rendering the typological one (a) redundant.

- (a) Infixes imply adfixes.
- (b) Infixes can only derive from adfixes through metathesis or entrapment.

In the implicational statement (a), the adfixation preference is easily seen as a special case of a more general preference of continuous over discontinuous constituents at whatever constructional level; and there is a plausible functional explanation, to do with the relative ease of storing, producing, and processing continuous and discontinuous constituents, that one would regret to see sacrificed. In the derivational statement (b), however, at least where the mechanism of reanalysis is metathesis, there is an equally plausible explanatory consideration, to do with the optimisation of syllable

structure. Looking at some infixation histories in different families, with particular attention to morphological and phonological representations which determine the fate of infixes (re-externalisation from, or fixation in, edge-bound positions), I will seek to disentangle the typological and diachronic strands of the respective plots.

Double Object Constructions: The Division of Labor between Semantics and Information Structure

Maria Polinsky

In the double object construction in English, some grammatical behaviors give superiority to Goal, while others favor Theme (object asymmetries). Some of these asymmetries are also observed in other constructions involving double complements. Same recurrent asymmetries between the objects of the double object construction and applicative construction are observed cross-linguistically. This paper reviews a series of asymmetries that recur cross-linguistically and addresses the following questions:

- (1) Do any of the morphosyntactic effects reflect the semantic or information-structural status of multiple objects?
- (2) If some effects are reducible to semantics, what do they show with respect to the hierarchy of semantic roles, in particular, the hierarchical relationship between Theme (Patient) and Goal (Recipient)?
- (3) If some effects are reducible to information structure, what do they show with respect to the general organization of MULTIPLE topics in a clause?
- (4) Which effects if any remain definitional for the grammatical function object?

The main conclusion drawn in the paper is that a number of object asymmetries are reducible to information-structural effects and subsequently, to processing (Erteschik-Shir 2004). Experimental evidence from English, which has a true double object construction, and from Russian and Japanese, both of which do not have such a construction but show similar effects, is reviewed. However, prosodic evidence from English is used to show that structural asymmetries between the two objects have to be maintained. Information structural asymmetries, which give preference to Goal, allow us to isolate purely semantic relationships as well. These relationships, which are rooted in the compositional structure of the events associated with ditransitives, have a direct bearing on the morphosyntax of ditransitives. One of the major typological challenges lies in separating semantics and information structure, and the paper will conclude with a brief discussion of possible diagnostics relevant for such a separation.

Causatives with Passive Meanings in Languages without Passives

Amara Prasithrathsint

Studies show that passives and causatives share certain common characteristics. Keenan (1990: 261-262) notes that passives and causatives exhibit more similarities across languages than one might reasonably expect. He cites examples from Japanese and German, which show identical markings on agent phrases in passives and causatives, and also adds that in some languages, such as Korean, the same verbal morphology may sometimes be interpreted as causative and sometimes as passive. Other studies, such as those by Whitman and Hahn (1988), Washio (1993), Shibatani (1994), and Kim and Pires (2004), focus on the same construction that can be interpreted as either passive or causative, especially in Korean. They propose several syntactic and semantic criteria for determining when such a construction should be interpreted as passive and when as causative.

Looking at the causatives in certain languages without passives, I observed that some of them could be interpreted as passive. This study aims to analyze causative constructions in three ancient languages in order to find out the elements that generate their passive meanings.

The data on which the analysis is based was taken from corpora of ancient Sukhothai Thai (the 13th Century), early Bangkok Thai (The 18th Century), and Ahom Tai (the 18th Century). These languages have been documented as languages without passives (See Prasithrathsint 2004).

The findings show that a causative sentence that has a passive meaning in the three languages has three syntactic characteristics: 1) the verb is transitive, 2) the subject NP and the Agent NP are omitted, and 3) the Patient NP is topicalized, as can be seen in the following example (from Ahom Thai):

Khon-raw	saam	tua	nii	<u>hay</u>	song	thii	caw-faa
people	three	body	this	Caus.	take	to	prince

'These three people, (someone) caused (someone) to take them to the prince.'

In a normal causative sentence, there are two noun phrases in front of and after /hay/, the causative marker; the one in front is the subject and the other is the agent phrase. However, in the above example there are none of those, and the Patient ('these three people') of the transitive verb /song/ 'take' is topicalized. Therefore, reading the sentence superficially, one may interpret it as a passive construction meaning 'These three people were taken to the prince.' Such a causative is rare in current Thai, in which a passive would rather be used to convey the same meaning.

In addition to their common syntactic structures, causatives and passives also share some semantic similarities. It can be inferred from the findings of this study and previous studies of passives in Thai that there is some connection between the /hay/ causative marker and the /thuuk/ passive marker, in terms of their sense of *obligation*. Indeed, it is found that some causative constructions have the same meaning as a construction containing an auxiliary meaning 'must/ have to' (e.g./tong/). As for the /thuuk/ passive marker in modern Thai, it was found by Prasithrathsint (2004) that at one stage of its development, it was used as an auxiliary meaning 'must/ have to'.

This study suggests that causatives and passives are related. Universally, causatives seem to emerge in a language before passives, and they may be used to convey passive meanings in languages without passives.

A Survey of Studies of the Minangkabau Language

Rina Marnita

The Minangkabau language is the native language of the Minangkabau people who mainly live in area which is now known as West Sumatera, a province in the island of Sumatera. There have been various studies on Minangkabau language. The earlier book about Minangkabau language was “Minangkabausche – Maleich - Nederland wooden-book” written by J.L ven der Toorn in 1891 who also compiled the first Minangkabau dictionary and “Minangkabau Spraakkuns” in 1899. The second dictionary, “Kamoes Bahasa Minangkabau-Bahasa Melayoe-Riau was compiled by Thaib St Pamoentjak in 1936.

Minangkabau has undergone some changes in its ortography system. First, the language was written in an Arab-Melayu script consisting of 34 letters utilizing both Minangkabau and Arabic sounds. Toorn’s spelling system has 26 phonemes, 19 cononants and 7 vowels. MC Emeis (1932) simplified Torn’s system by reducing the vowels to 5 and changing fonem /k/ to the glottal stop /?/, adding four allophones. Among those who studied the Minangkabau sound system are G.E. William (1961) who published “Colloquial Minangkabau: a description of phonological and morphological structure”. In 1990, the Indonesian language Agency published “A Guidance to Minangkabau Spelling ystem” which consists of explanation of spelling system and regulations used in written Minangkabau. However, there is almost no study conducted by the native on the Minangkabau script yet.

Studies on Minangkabau language as well as culture and literature greatly increased after a grant from the Indonesian government for the study of Indonesian regional languages became available. Among them are study on preposition (Syamsir Arifin 1980) ; a tagmemic analysis of the nominal phrase (Husin 1980); structure of Minangkabau language spoken in Lima puluh kota, Agam, Tanahdatar and Pesisir (Shawin Nikelas 1980). Studies on dialects were done by Tamsin Medan (1982), Marjusman Maksan (1981). A description of Minangkabau grammar, written in French , was done by a missionary, Gerard Moussay, in 1981. This book has become the most popular and important reference in writing about the Minangkabau language.

During the last decade, studies on various aspets of Minangkabau language have increased significantly with the increasing numbers of native people studied for either Master or PhD program. Besides semantic, sociolinguistic and pragmatic studies, there is a typological study on Minangkabau classifiers (Marnita, 1996); the Minangkabau reduplication Gaffar, 1998); a lexicostatistics methods for a reconstruction of the proto-phonemes of the Minangkabau and Kerinci (Sastra, 1994) and an intensive and in depth dialect study to find the origin of he Minangkabau people (Nadra, 1994;1998, now).

Person Marking Systems in Dani Languages

Yusuf Sawaki

The Great Dani language family is a Papuan language family of Trans-New Guinea Phylum spoken by people in the central highlands of New Guinea, around the valley of the Balim River. In this paper, I present person marking systems of Dani languages. In common, Dani languages constitute a great distribution of pronominal markers in morphological inflection on verbs, nouns, adjectives, and postpositions. This distribution shows a clear characteristic of active and non-active person marking systems. There are two common sets of pronominal markers; Set I is active person markers that suffix to transitive verbs and non-stative intransitive verbs. Set II is non-active person markers that prefix to transitive verbs, functioning as object markers, stative verbs and adjectives, functioning as subject markers, subject markers in postposition and possessors in nouns. This phenomenon is common across Dani languages.

Key terms: The Great Dani language family, person marking systems, active and non-active, transitive and intransitive verbs.

The Life Cycle of Antipassive Constructions

Sergey Say

The life cycles of antipassive constructions (henceforth, ANTI, i.e. intransitive clauses derived from transitives with the O demoted or deleted and A converted into the S function) is analyzed within the framework of the theory of grammaticalization. Grammaticalization is understood not only as a hyper-model of diachronic change but also as an insightful dynamic approach to the study of synchronic distribution of more and less obligatory uses of a construction.

Most students of ANTI constructions tend to emphasize the alleged distinction between pragmatic-semantic (usually signaling indefinite, habitual or not fully affected Os) and structural (A-promotion for participation in syntactic processes tuned to the S/O pivot) varieties thereof. However, it can be shown that even if ANTI serves syntax it still has pragmatic or semantic motivation.

Despite significant cross-linguistic diversity, typological generalizations have been put forward with respect to pragmatic/semantic functions of ANTI. They are most commonly used in absolutive constructions, i.e. signaling generalized or existential patients. The development of the so-called “full” uses of ANTI (i.e. clauses with demoted rather than deleted original Os) is a step towards grammaticalization that makes antipassivization a potentially productive syntactic rule (this is attested in e.g. some Uto-Aztecan and Daghestanian languages). Typical pragmatic properties of “full” ANTI clauses include downgrading potential Os with low referential persistence and backgrounding the meaning of the verb-patient complex, so that its meaning is either in presupposition or is asserted as a more unitary concept than in the corresponding transitive construction.

The paper examines a number of syntactic contexts in which ANTI is obligatorily required / frequently encountered in various languages. The contexts examined include answers to questions of the ‘what is X doing?’ type (Pipil, Dargwa); A-based wh-questions; focused A-negation (‘it was not X who Verbed’, e.g. in Mayan languages); coordinate reduction of As (obligatory in e.g. Dyirbal and optionally used as a disambiguation mechanism in Chukchee); A-based relative and other types of subordinate clauses; agentive participles (e.g. Eskimo); A-resultatives of the ‘to have eaten one’s fill of sth.’ type (e.g. Dargwa); actor nominals (e.g. Dyirbal, Chinook and optionally in Eskimo). In each case an explanation is offered of how these syntactic functions emerged from arguably more basic pragmatic functions in those languages. For instance, antipassivization for A-relativization is shown to stem from the discourse tendency to use relative clauses for coding backgrounded information. Potential intermediate stages in the conventionalization of ANTI in A-based relative clauses are discussed (e.g. Mam has a contrast between present and past tense relative clauses; obligatory use of ANTI in the latter is explained as a generalization of a typical pattern of discourse organization).

The second part of the paper examines the interaction of antipassive derivation with the lexical semantics of individual verbs and in particular the processes of lexicalization of antipassive derivatives. It appears that in most languages that have morphological antipassives, their derivation is not fully productive and there are numerous semantic idiosyncrasies in such derivatives. It is shown that antipassive derivatives are normally diachronically unstable which is accounted for by the fact that pragmatic / contextual suppression of object-oriented meaning components and concomitant stress on agent-oriented ones may lead to significant changes in the semantic structure of individual lexemes (e.g. aspectual). That is why antipassive often

constitutes a fossilized degrammaticalized class of derivatives rather than a full-fledged member of the voice system of a language.

The final conclusion is that even if there are languages that systematically use antipassives for purely syntactic and not pragmatic reasons, such a state of affairs is unlikely to be diachronically stable and it most likely reflects a prior pragmatically-determined stage that have been rearranged due to syntactic analogy and conventionalization; formation of antipassives usually starts to dissipate by force of lexicalization sooner that it gets fully grammaticalised. Such a relatively short life cycle of antipassives accounts for the wide-spread co-existence of several independent antipassive-like processes that are attested in one and the same language (Dyirbal, Eskimo, Halkomelem, several Mayan languages, a few Athabaskan languages, probably some Philippine languages). Typically, the two processes are on discrepant stages in their life cycle. These complex systems pose some problems for purely syntactic analyses in which they appear to be unnecessary superfluous doublets. However, they can be nicely handled in the perspective of the theory of grammaticalization.

Universalities and Diversities in Synaesthetic Metaphors: A Contrastive Study of Hebrew and Indonesian Yeshayahu Shen & David Gil

All languages make use of a variety of kinds of figurative expressions, such as metaphors, similes, oxymorons, analogies, proverbs and idioms. The cross-linguistic study of figurative expressions reveals the existence of important universal properties alongside a significant amount of diversity. To cite just one example: many languages make use of a universal conceptual metaphor in accordance with which emotions are stored in a physical container which is associated with a particular body organ. However, languages differ with respect to the choice of organ: while in English it is the heart, as evidenced in expressions such as *broken heart*, in Minangkabau it is the liver, as in *patah hati*, literally "break liver".

This paper presents some preliminary results from a large-scale ongoing crosslinguistic study of figurative expressions. Its concern is with metaphors of one particular type, namely, *synaesthetic metaphors*: metaphors that involve a mapping between two sensory domains. For example, the English expression *sweet silence* is a synaesthetic metaphor, since it involves a mapping from the source domain of taste to the target domain of sound. The paper presents some results of a contrastive study of synaesthetic metaphors in two languages: Hebrew and Indonesian. The contrastive study draws upon two complementary sources of data: (a) production, as manifest in naturalistic texts of different kinds, written and oral, artistic and everyday; and (b) production and comprehension, as evidenced by experimental tasks of various kinds, such as recall, judgements of appropriateness, context generation and interpretation generation.

The main result to emerge from the study is a pattern common to both Hebrew and Indonesian, in which synaesthetic metaphors are governed by the following *hierarchy of senses*:

- (1) touch < taste < smell < sound < sight

The above hierarchy of senses determines the structure of synaesthetic metaphors in accordance with the following principle:

- (2) Metaphors in which the source domain is lower on the hierarchy than the target domain are preferred over metaphors in which the source domain is higher on the hierarchy than the target domain.

For example, in a metaphor such as *sweet silence* the source domain taste is lower on the hierarchy than the target domain sound and hence this metaphor is preferred over its counterpart *silent sweetness*, in which the source domain sound is higher on the hierarchy than the target domain taste. Indeed, our findings suggest that in both Hebrew and Indonesian corpora, synaesthetic metaphors in compliance with principle (2) tend to occur more frequently than their counterparts in violation of the principle. And correspondingly, subjects in both Hebrew and Indonesian perform better on a variety of experimental tasks when presented with synaesthetic metaphors that satisfy principle (2) than they do when faced with other metaphors that violate the same principle. On the basis of the above results, it would seem, then, that the hierarchy of senses and principle (2) are promising candidates for universality in the domain of synaesthetic metaphors.

Alongside the above fundamental similarity between Hebrew and Indonesian, our study also reveals the existence of certain differences between Hebrew and Indonesian in the realm of synaesthetic metaphors. At present, we have no systematic account of these differences, other than to suggest that they might be due to any or all of a handful of heterogeneous factors, such as, perhaps, arbitrary conventionalized properties of particular lexical items and phrases, the distinct grammatical properties of the two languages, and cognitive or cultural differences between speakers of the two languages.

The present study is part of a large-scale research programme currently under way, combining corpus studies with experimental techniques in order to investigate a wider range of figurative types in Hebrew, Indonesian, and other languages. Our hope is to lay the foundations for a cross-linguistically informed typology of figurative language, and, in so doing, achieve a more empirically adequate understanding of figurative language, and thus also of language in general.

Bound Person Forms in Ditransitive Clauses Revisited

Anna Siewierska & Dik Bakker

It has been frequently noted that as far as bound person forms are concerned, in ditransitive clauses the R is clearly favoured over the T. While there are languages which exhibit bound person marking for the T to the exclusion of the R (e.g. Guarani, Jacaltec, Paamese, Yapese) the converse is far more common (e.g. Alamlak, Anejom, Araki, Bimoba, Cahuilla, Cora, Halkomelem). Moreover, bound person forms for the R may exist even in the absence of any bound person forms for the P in monotransitive clauses. Such is the case, for example in the Chadic language Gude or languages such as Malayalam, Lepcha, Tsez and Waskia. By contrast there are no languages which have bound person forms for the T but not for the P. Unlike in the case of monotransitive clauses in which marking of both the A and P by bound person forms is more common than marking of just either the A or the P, in ditransitive clauses bound person marking of both the T and R is relatively infrequent. In the languages which do allow for person marking of both the T and R, the T and R forms may be phonologically identical as in (1) or phonologically distinct either completely as in (2) or at least for some of the forms as in (3).

(1) Diola Fogany (Sapir 1965:91)

	T	R
1sg	- m/-a:m	- m/-a:m
2sg	-i	-i
3sg	- -	- l
1 excl	-uli/-oli	-uli/-oli
1incl	- la/ - lal	- la/ - lal
2pl	-u/-ul	-u/-ul
3pl	-i:/-il	-i:/-il

(2) Tamazight of the Ayt Ndhir (Penchoen 1973: 26, 27)

	T	R
1sg	i-/yi	i-
2sg m	(i)š-	aš-
2sg f	(i)š m-	am-
3sg m	(i)t-	as-
3sg f	(i)tt-	as-
1pl	aw-	ax-
2pl m	(i)k n	awn
2pl f	(i) k _w nt	ak _w nt
3pl m	(i)t n	as n
3pl f	(i)t nt	as t

(3) Ekari (Doble :82, 84)

	T	R
1sg	na-	na-
2sg	ka-	ka-
1pl	ne-/ni-	niya-
2pl	ke-/ki-	kiya-
3	e-	ya-

Symposium: Ditransitive Constructions

Anna Siewierska & Willem Hollman

One of the consequences of the increasing influence of construction-based approaches to grammar is the recent resurgence of interest in ditransitive constructions (e.g. Newman 1998; Primus 1998; Comrie 2001; Haspelmath 2001, 2005; Siewierska 2002, 2003, Gensler 2003). While much of the earlier work on ditransitive clauses focused on the similarities between ditransitive and transitive clauses and in particular on which of the non-subject arguments qualifies as the direct object, current research is more directed at ditransitive constructions, as constructions in their own right. This research has revealed much more internal variation in the structure, use and diachronic development of ditransitive constructions, both within and across languages, than previously recognised.

The aim of this workshop is to provide further insights into the nature of ditransitive constructions by bringing together research conducted from three different perspectives; of a specific language, of an areally and/or genetically determined group of languages and of languages in general. The major issues to be discussed are:

- the factors determining the emergence, distribution (with different predicates, different participants, in different dialects) and use of ditransitive constructions in individual languages and genetically and areally related groups of languages,
- the diachronic changes manifested in the productivity and grammatical encoding of ditransitive constructions
- the morpho-syntactic reflexes of changes in degrees of semantic transitivity and of the interplay between semantics and information structure in ditransitive constructions.

Given the enormous amount of attention that has been devoted to the analysis of ditransitive constructions in English, the discussion of the development and choice between alternative ditransitive constructions will begin with English and related Germanic languages. It will then branch out to cover other genetic groupings and geographical areas, in particular Oceania, and Eurasia as well as languages in general.

Comparing Constructions across Languages: A Case Study of the Relationship between the Inclusory Construction and Some Related Nominal Constructions.

Ruth Singer

Despite claims that syntactic constructions are not comparable between languages (e.g. Croft 2001), typologists have described striking similarities across languages in a range of functional, semantic and formal domains. Some such as Moravcsik (2003) define the domain of investigation using semantic properties alone. Others use a combination of semantic, formal and functional properties to define the domain of investigation (Haspelmath 2004). Stassen (2000) combines these three types of properties into hierarchical lists he calls 'strategies'. These strategies define a domain that has a radial structure.

I look at how these approaches shape our perspective on language-internal and cross-linguistic relationships between constructions through a case study of the relationships between the inclusory construction and two closely related nominal coordination constructions: conjoined nominals and comitative constructions. These three constructions are described for a range of languages in the recent volume *Coordinating constructions* (Haspelmath 2004). The relationship between inclusory constructions and associatives will also be discussed.

The types of relationships found between these constructions in particular languages is vast. In some languages one formal construction covers the functional range of all three constructions (Stassen 2000). In Russian and French the inclusory construction shares the same form as the comitative construction while in other languages it has a unique form. Inclusory constructions and associatives take the same form in some languages, whereas in others they differ.

The large range of patterns of relations between the three constructions in individual languages suggests that a deeper investigation of language internal relations is required before cross-linguistic comparison can be carried out. It is easy to assume that a single formal construction has a single meaning within a language, or that constructions which share the same grammatical marker are variations of a single construction. Although meaning is often used to define the domain of investigation in typology, description of the domain tends to focus on the formal. This obscures a number of relationships that can exist between constructions - where they share one form that is polysemous between two constructions, and where two constructions in a language share a single marker, despite being otherwise distinct. In addition, constructions without a unique grammatical marker, such as the inclusory construction, tend to be overlooked altogether. Through investigation of these related constructions I aim to defend the activity of comparing constructions across languages while highlighting some of the many pitfalls of such an endeavour.

References

- Croft, W. (2001). *Radical construction grammar*. Oxford, Oxford University Press.
- Haspelmath, M. (Ed.) (2004). *Coordinating constructions*. Amsterdam, John Benjamins.
- Moravcsik, E. (2003). A semantic analysis of associative plurals. *Studies in Language* 3: 469-503.
- Stassen, L. (2000). AND-languages and WITH-languages. *Linguistic Typology* 4(1): 1-54.

Trade-offs between Linguistic Subsystems: Evidence for Equal Complexity of Languages?

Kaius Sinnemäki

It has often been claimed that there is a trade-off between the complexities of different subsystems in languages. By subsystems I mean the formal domains of phonology, morphology, syntax, and lexicon, all of which may be divided in further subsystems (i.e. phoneme inventory, syllable structure, morphological elements, syntactic rules, etc.). The purported trade-off occurs e.g. when languages with little or no (inflectional) morphology are said to compensate this by elaborate syntax (e.g. Hockett 1958: 180,181). Thus, complexity in one domain of grammar is “balanced out” with simpler phenomena in another. However, this claim, which I refer to as the hypothesis of equi-complexity, recurs in the literature often without being systematically tested against cross-linguistic evidence. In addition, the few prior typological studies have limitations in sample size and diversity, or reliability. Quite recently, though, the testing of the hypothesis has received more attention in cross-linguistic studies (Fenk-Oczlon & Fenk 1999, McWhorter 2001, Dahl 2004).

In this presentation, I have two broad aims. Firstly, I present an overview of the evidence in the few relevant earlier studies that either support or refute the equi-complexity hypothesis. Secondly, I present results of a typological pilot study that directly tests the hypothesis by assessing the proportions of different subsystems in coding the primary roles in simple declarative clauses.

Apparently, there is evidence in earlier literature both for and against the hypothesis; some pressure towards a balanced system economy seems credible (e.g. Fenk-Oczlon & Fenk 1999) whereas clear deviations, i.e. both scantiness and abundance of formal coding, occur as well (e.g. McWhorter 2001). However, earlier studies have usually addressed the subsystems that are most readily measurable, that is, size of phoneme inventory, syllable structure, word length, degree of synthesis and fusion, etc. Domains that are more difficult to measure, such as syntax, have received much less attention. It is understandable that very few measures of syntactic complexity have been developed, not least due to the great diversity of structures present in languages and due to the difficulty of delimiting syntax and morphology from one another. However, an assessment of the proportions of morphological and syntactic subsystems may be possible by approaching grammar as in the Systems Interaction model (Frajzyngier & Shay 2003). In this model grammar is perceived as a network of independent yet interrelated coding means (which are roughly equivalent to linguistic subsystems). This enables formal systems to be kept apart in determining whether they compensate one another in coding a particular functional domain (in the sense of Givón 1981)

Thus, my second aim for this presentation is to present results of a typological pilot study of how languages code the primary roles in simple declarative clauses (i.e. subject, and object, agent and patient, topic and focus). I develop a method based on the Systems Interaction model that estimates the proportions of different formal coding means which have usually been associated with either morphology or syntax, including e.g. agreement, case marking, word order, and adpositions. A small sample of less than 25 languages will be used at the outset to arrive at a workable method and at some preliminary results. The languages are analysed according to type, internal structure, and number of coding means. Koromfe (see example 1), for instance, distinguishes its subject and object from one another on the basis of (SVO) word order alone. Wari' (2) distinguishes NP functions by subject and object agreement on a verbal inflectional

clitic alongside with word order (where the former indicates the functions more reliably). Wardaman (3), on the other hand, marks its nouns by morphological cases (absolutive and ergative) besides using subject and object agreement on the verb. To attain more extensive and reliable generalizations, a larger set of languages will be consulted after the initial stage. The languages are stratified according to genetic affiliation, macro-areas, and morphological typology.

Finally, I will also address some theoretical and methodological problems involved in identifying the primary roles and the formal coding means. There are also some pitfalls to avoid in adapting the Systems Interaction model because it assumes one-to-one mapping of meaning and form. In order to avoid any possible inbuilt equi-complexity in the method, I assume a possible one-to-many mapping; several coding means may interact and overlap in coding a particular functional domain.

Data

- (1) Koromfe (Niger-Kongo, Gur; Burkina Faso; Rennison 1997: 69)

d□ *pa a jana a k*□□ *ho*□
 PRON.3SG.HUM give ART millet.PL ART woman.SG DET.HUM.SG

‘He gives the millet to the woman’

- NP functions are distinguished by word order: subject (*d*□) precedes the verb (*pa*) and object (*a jana*) follows it.

- (2) Wardaman (Australian, Gunwinyguan; Australia; Merlan 1994: 64)

dang-mulu-yi yibyan-mulu-yi wunggunburr-bu-ndi
 that-PL-ERG man-PL-ERG 3NSG/3NSG-hit-PST

‘Those men hit them’

- Verbs are prefixed with subject and object pronominals (*wunggunburr-*) and nouns suffixed with morphological cases (*-yi* marking ergative case in the example).

- (3) Wari' (Chacapura-Wanham, Madeira; Brazil; Everett & Kern 1997: 123, 132-134)

to' non [na-on] 'Orowao'
 hit 3SG.RP/P-3SG.M name

‘He hit 'Orowao’

- Both subject and object nouns may trigger person, number, and gender agreement on a verbal inflection clitic (*non* in the example). In addition, word order interacts with agreement in indicating the NP functions.

Abbreviations

PRON	pronoun
3	3rd person
SG	singular
PL	plural
HUM	human
ART	article
DET	determiner
ERG	ergative
PST	past
RP/P	realis past/present
M	masculine

References:

- Dahl, Östen 2004. *The Growth and Maintenance of Linguistic Complexity*. Studies in Language Companion Series 71. Amsterdam: Benjamins.
- Everett, Daniel L. & Barbara Kern 1997. *Wari': The Pacaas-Novos language of Western Brazil*. London: Routledge Press.
- Fenk-Oczlon, Gertraud & August Fenk 1999. Cognition, Quantitative Linguistics, and Systemic Typology. *Linguistic Typology* 3, 151-177
- Frajzyngier, Zygmunt & Erin Shay 2003. *Explaining Language Structure through Systems Interaction*. Typological Studies in Language 55. Amsterdam: Benjamins.
- Givón, Talmy 1981. Typology and Functional Domains. *Studies in Language* 5 (2): 163-193.
- Hockett, Charles F. 1958. *A Course in Modern Linguistics*. New York: Macmillan.
- McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5 (2-3): 125-156.
- Merlan, Francesca C. 1994. *A Grammar of Wardaman*. Berlin: Mouton de Gruyter.
- Rennison, John R. 1997. *Koromfe*. Descriptive Grammars. London: Routledge.

Towards a Cross-Linguistic Typology of Extra-Clausal Constituents

Niels Smit

In the domain of information management, many scholars (Prince 1978; 1984; 1998; Ward 1988; Jacobs 2001; Gregory & Michaelis 2003, as well as numerous others) distinguish two types of constructions that involve extraclausal constituents: Dislocation (DIS) and Topicalisation (TOP). They differ in that the former, but not the latter, requires the presence of a resumptive item clause-internally (cf. ex. [1]). DIS has been shown to have two discourse functions, *simplification* and *set establishment*, which are unified by the superordinate function of *Topic promotion*. TOP, on the other hand, functions as a *presupposition creating* device. As such, it is not a Topic promoting construction, even though it shares the function of set establishment with DIS. Departing from these functional specifications, formal constraints on the use of both constructions have been shown to hold.

However, if we extend our view beyond English, we find an unsettling amount of formally distinct constructions which suggest that the class of DIS is much more complex than suggested above. It seems that, formally, at least three additional types of DIS constructions must be distinguished: DIS constructions where the lexical element occurs both clause-internally and -externally, DIS constructions with a clause-internal lexeme that is ‘thematically’ related to the extraclausal constituent, and constructions which combine formal properties of DIS and TOP (ex. [2]-[4]). Moreover, we find unexpected functional behaviour in that some of these constructions appear to have a Focus marking rather than a Topic promoting function.

This research (part of an ongoing project that aims at a comprehensive typology of information structure) explores the various configurations that hold between clauses and extraclausal constituents cross-linguistically. Capitalising on insights from Functional Discourse Grammar (Hengeveld & Mackenzie *fc.*) and Role and Reference Grammar (Van Valin 1993), it aims to provide a unificational typology of DIS and TOP constructions. To this end, use is made of a broad, typologically adequate sample of languages. Furthermore, Michaelis & Gregory (2003) observe that remarkable functional parallels exist between TOP and cleft sentences on the one hand, and DIS and presentational constructions on the other. It will be shown that these constructions can be incorporated in the unificational typology as well, with formal variation attributed to various language-specific interactions of the interpersonal and morphosyntactic levels of grammar.

Examples

1. DIS and TOP
 - a. *John I'd never betray him* (DIS)
 - b. *Mary she's doing better in life than I am* (DIS)
 - c. *Francis I lent him my copy of "Pride and Prejudice"* (DIS)
 - d. *He's a nice guy, Peter* (DIS)
 - e. *Cats I don't like Ø* (TOP)
 - f. *Nancy I gave Ø a book* (TOP)

2. DIS, lexical copy
- a. *Gbá wε S'éná gbá xwé l∅*
 Build FOC Sena build.PF house Spf_[+Def]
 'Sena BUILT the house' (Gungbe; Aboh 2004:250)
- b. *Saiatu saiatuko gara*
 Try try.FUT AUX
 'As for trying, we will try' (Basque; Hualde & Ortiz de Urbina 2003:458)
- c. *sɔtyɛ̃ a ɲwa sɔtyɛ̃ a ɲu*
 hide-DEC I hide-DEC FACT
 'I really DID hide (it)' (Lisu; Hope 1974:171)
3. DIS, 'thematically linked'
- a. *h∅ ∅∅ na-qh∅∅ yɛ ve yò*
 elephant TOP nose long PRT DECL
 'Elephants, noses are long' (Lahu; Li & Thompson 1977:462)
- b. *O le aiga o Melefítipo o Meredith,*
 PRES ART family POSS Melefiti Q PRES Meredith
e toa=fitu na maliliu
 GENR HUM=seven PAST die.pl
 'The family of the Melefitis, seven people died' (Samoan; Mosel & Hovdhaugen 1992:465)
- c. *Nei-chang huo∅ xìngkui xi∅ofang-duì lái de kuài*
 That-CLF fire fortunate fire-brigade come PRT (ADV) quick
 'That fire, fortunately the fire-brigade came quickly' (Mandarin Chinese; Li & Thompson 1977:462)
4. DIS – TOP combination
- a. *Vrienden_i die_i kan je nooit genoeg Ø_i hebben*
 Friends DEM can.3sg IMPR never enough have
 'One can never have enough friends' (Dutch)

References

- Gregory, Michelle L. & Laura A. Michaelis. Topicalization and Left-Dislocation: A Functional Opposition Revisited. *Journal of Pragmatics* 33, 1665-1770.
- Hengeveld, Kees & J. Lachlan Mackenzie (in prep.) *Functional Discourse Grammar*.
- Prince, Ellen F. (1981). Towards a Taxonomy of Given-New Information. In Cole, Peter (ed.), *Radical Pragmatics*, 223-255. New York: Academic Press.
- Prince, Ellen F. (1998). On the Limits of Syntax, with Reference to Left- Dislocation and Topicalization. In Culicover & MacNally (eds.).
- Van Valin, Robert D. jr. (1993). A Synopsis of Role and Reference Grammar. In Robert D. Van Valin jr. (ed.), *Advances in Role and Reference Grammar*, 1-164. Amsterdam: John Benjamins.
- Ward Gregory & Betty Birner (2004). Information Structure and Non-canonical Syntax. In Horn (ed.), *The Handbook of Pragmatics*. Blackwell Handbooks in Linguistics 16. Malden MA: Blackwell Publishing.

Three out of Two Ain't Bad: The Development of a Three-Participant Construction in Oceanic Languages¹⁵

Jae Jung Song

In Mokilese (Micronesian subgroup, Oceanic; Mokil Atoll and Ponape), the presence of a recipient in three-participant events is indicated by the verbal suffix *-oang*, as in (i):

- (i) joam-oai-o ki-oang ngoahi ekij mwani
father-1SG:POSS-DET give-to me some money
'My father gave me some money.'

The marking of a recipient can also be carried out by so-called possessive classifiers with or without the verbal suffix *-oang*, as in (ii) or (iii), respectively:

- (ii) ngoah rapah-ki ih pwa ngoa-n ki-oang
I look.for-COMP him COMP I-MODAL give-to
nah-Æ mwani-he
CL-3SG:POSS money-DET
'I looked for him (in order) to give him the money.'
- (iii) ngoah insingeh-di kijinlikkoauoaw nih-mw
I write-PFV letter CL-2SG:POSS
'I wrote a letter to/for you.'

The primary function of possessive classifiers (e.g. *a-*) is to encode alienable possession (e.g. *mwinge*), as in (iv):

- (iv) *in-oai* *li-ho* *onopda* *a-sai* *mwinge-hu*
mother-1SG:POSS woman-DET prepare CL-1PL food-DET
'My mother prepared our food.'

The possessive classifiers are exploited to encode the semantic role of recipient (e.g. (ii) and (iii)) or of beneficiary (e.g. (iii)) in Mokilese. This secondary use of possessive classifiers is found to be widespread in other Micronesian and Oceanic languages, although it is typically attested in the context of the encoding of a beneficiary (e.g. Song 1997).

In this paper, I will describe and explain the development of a three-participant construction (e.g. (iii)) from a two-participant construction (e.g. (iv)) in Mokilese and other Oceanic languages. Evidence will be brought to bear in demonstrating that possessive classifiers have grammaticalized as benefactive markers (*pace* Croft 1985, who argues for the beneficiary realized as the possessor via a synchronic "indirect object lowering" rule). I will also demonstrate that this development has a pragmatic motivation. When the possessor occurs in the context of an activity that brings the possessor into possession of the direct object or theme, the possessor can be construed

¹⁵ Three-participant constructions typically involve an agent, a theme, and either a recipient or a beneficiary. I have chosen to use the term "three-participant constructions" in preference to "ditransitive constructions", because identical constructions may be used to encode three-participant events, regardless of whether they involve a recipient or a beneficiary. Moreover, it is not always possible to discriminate between a recipient and a beneficiary in terms of complement/adjunct status. Newman (1996) discusses four conceptual domains by which three-participant constructions can be recognized (i.e. spatial-temporal, control, force-dynamics and human interest domains). By his definition, thus, both (i) and (iii)—with a recipient and a beneficiary, respectively—will be identically analysed as three-participant constructions.

as the beneficiary of that activity (*The man built the boy's canoe* *The man built a canoe for the boy*). Thus, the use of possessive classifiers as benefactive markers tends to be attested in conjunction with verbs of creation (e.g. *build, cook, sew, write*) or verbs of obtaining (e.g. *hunt, kill, fish*) (as originally observed by Croft 1985). Finally, I will discuss the implications of the development for the widely held view that structural scope decrease is a manifestation of grammaticalization. However, the reanalysis of possessive classifiers as benefactive markers involves structural scope increase, not decrease (also see Tabor and Traugott 1998).

Abbreviations

CL	Possessive Classifier
COMP	Complementizer
DET	Determiner
PFV	Perfective
PL	Plural
POSS	Possessive
SG	Singular
1	First Person
2	Second Person
3	Third Person

References

- Croft, William (1985) 'Indirect object "lowering"', *Berkeley Linguistics Society* 11,39-51.
- Newman, John (1996) *Give: a cognitive linguistic study*. Berlin: Mouton de Gruyter.
- Song, Jae Jung (1997) 'The history of Micronesian possessive classifiers and benefactive marking in Oceanic languages', *Oceanic Linguistics* 36, 29-64.
- Tabor, Whitney and Elizabeth Traugott (1998) 'Structural scope expansion and grammaticalization', in *The limits of grammaticalization*, A. Giacalone Ramat and P. J. Hopper (eds), Amsterdam: John Benjamins.

Deictics in Oma' Lung, a Kenyah Language of Kalimantan

Antonia Soriente

Oma' Lung is one of the Kenyah languages spoken in Kalimantan (Indonesia) and Sarawak (Malaysia). Oma' Lung, an endangered language counting not more than 3,000 speakers nowadays spread in about 10 villages in East Kalimantan has never been described before.

This paper presents a preliminary study of the deictics in Oma' Lung, in particular all the cues provided by the language to localise the speech event and its participants, in space and time.

As outlined by Anderson and Keenan (1985), there are three major categories of deixis: person deixis, spatial deixis and temporal deixis to which should be added a fourth category, the psychological deixis. As Svorou (1993) observed social and psychological conditions are also relevant factors in the deictic anchorage of language. This fully applies to the Oma' Lung language. Person deixis refers to the personal pronouns. Other than designate the speaker, the hearer and the third person, Oma' Lung encodes additional information about the referent, for example the number of individuals referred to (singular, dualis, trialis, plural), (impersonals, deferential pronouns). The subsystem of pronouns and possessives displays two sets, the first one as free morphemes and the second as post-posed particles that function as pronouns or possessives. Spatial deixis involves the demonstratives. In Oma' Lung there are three locative dimensions: designating locations in space with reference to the position of the Speaker (*jĴ* near the Speaker, versus *ji* near the Hearer and/or *ki'i* not-near any of the Speech participants). Temporal deixis localises the speech event in time by means of adverbs ('now', 'then'), therefore it is expressed by means of adverbs or demonstratives.

One of the scopes of this research is to answer basic questions about deictics: to find out whether, as pointed out by Himmelmann (1996), there is a way to distinguish between demonstratives and 3rd person personal pronouns, whether demonstratives are also used anaphorically and whether there is any correlation between the spatial and temporal deictic differentiation in three locative dimensions.

Psychological deixis encodes information about the referent that is related to the psychological framework of the Speech participants rather than to the localisation of the Speech event in space and time. For this reason a certain number of evidentially terms are used in narrative, where every statement must contain a specification of the type of evidence on which it is based: for example, whether the speaker saw it, *dito*, or heard it *dae*, or inferred it from indirect evidence, or learnt it from someone else *re/de*. These terms that according to the local speakers might not seem relevant because untranslatable in the contact language, Bahasa Indonesia, appear indeed in every single sentence and seem to be obligatory in narrative texts.

Data for this study are taken from a number of texts recorded by the writer in her recent field work in Setulang, Malinau, in the Indonesian province of East Kalimantan and analysed following the guidelines given by Engelenhoven (2000).

If from one side Oma' Lung is considered 'endangered' because of the pressure of the national language, Bahasa Indonesia on this marginal language as demonstrated for example by the use of many Indonesian demonstratives like 'ini' or 'itu' in the recorded speech of Oma' Lung speakers, on the other hand some typical Oma' Lung deictic functions like the use of dualis and trialis in the personal pronouns is calqued in Indonesian.

Areal Factors in Two Sets of Mobile Clitics
in the Verbal Systems
of the Languages of the South Caucasus and Northern Iran
Don Stilo

The languages in the S. Caucasus/N. Iran have 2 different sets of mobile clitics in their verbal systems: (1) Set₁ (enclitic forms of copula) cross-reference the S/A in (non-copula) verb paradigms (see p. 2), and (2) Set₂ (of possessive pronominal origin) in some languages crossreference core arguments (S/A and/or P) of the verb. The table below details restrictions on both Set₁ and Set₂. I will examine languages from 5 different language families of the area, but languages outside this area (e.g., Persian, C. Kurdish, Arabic) will also be included.

Languages and Affiliations			Features (Isoglosses) (see list below)													
			1	2	3	4	5	6	7	8	9	10	11	12	13	14
Iranian	NWI: Tatic	N. Talyshi	+	+	(-)	+	+	+	+	+	-	+	-	-	+	+
	NWI: Tatic	S. Talyshi	+	+	-	+	+	+	-	+	-	+	-	-	-	+
	NWI: Tatic	N. Tati	+	+	-	+	+	+	-	+	-	+	-	-	+?	+
	NWI: Tatic	Vafsi	+	+	+	+	+	+	-	+	+	+	+	+	-	+
	NWI: Caspian	Gilaki	+	-	N/A	-	+	+	-	+	-	N/A	N/A	N/A	N/A	N/A
	NWI: Kurdish	Kurmanji	+	-	N/A	-	+	+	-	+	+	N/A	N/A	N/A	N/A	N/A
	SWI:	Caucasian Tat	+	+	+	-	+	+	+	+	-	N/A	N/A	N/A	N/A	N/A
Indo-European	Indo-European	Armenian	+	+	+	-	+	+	+	+	-	N/A	N/A	N/A	N/A	N/A
	Altaic	Azeri	+	+	+	-	+	+	+	+	-	N/A	N/A	N/A	N/A	N/A
	Semitic	Neo-Aramaic	+	+	+	+	+	+	+	+	+	+	+	+	+	-
	Caucasic/Lezgian	Udi	+	(+)	-	(+)	+	+	+	+	-	-	-	-	+	-
	Kartvelian	Georgian	+	+	-	+	+	+	((-))	+	(+)	+	+	+	+	-

1. Language has Set₁ clitics (1, 2, 3 person singular, plural)
2. Language has Set₂ clitics¹ (1, 2, 3 person singular, plural; only singular for Armenian)
3. Set₂ function enclitically as pronominal possessives on noun/NP hosts
4. Set₂ has one or more functions in the verbal system
5. Set₁ function predicatively as enclitic forms of the copula (with nouns, adjectives, PPs, etc.)
6. Set₁ are *suffixal* agreement markers cross-referencing S/A argument of verb (see data, Pt.1)
7. Set₁ (copula) function as enclitic agreement markers in present tense (progressive aspect)
8. Set₁ (copula) function as enclitic agreement markers in the present perfect tense
9. Set₁ function to cross-reference P argument in the verb in the past tenses²
10. Set₂ function enclitically to cross-reference S/A argument in the verb in the past tenses^{1,2}
11. Set₂ function enclitically to cross-reference P argument in the verb (mostly in non-past tenses)
12. Set₁ & Set₂ *forms swap functions* from present to past tenses: Set₁A/Set₂P > Set₁P/Set₂A²
13. Set₁ of the verbal system front to various hosts in clause, depending on triggering factors
14. Set₂ of the verbal system front to various hosts in clause, depending on triggering factors

¹ in Georgian Set₂ is affixal only; Georgian also has a Set₃ that cross-references S/A in various tenses;
² where above rules refer to past tenses, the equivalent rules apply to Georgian present and past perfect

While I will give a brief run-down of both sets of clitics, the focus of my paper lies in two areas:

- *Fronting* (clause-level leftward mobility) of: 1) Set₁ (Feature 13) in the present and perfect tenses (\pm transitive); (in Iranian, this feature only occurs in the narrow geographic area in focus here); and of 2) Set₂ (Feat. 14) (typical only of Iranian, also outside this area, *transitive only*). Since N. Talyshi / N. Tati are Iranian and participate within this Sprachbund, they are the only languages that have both sets in the verbal system **and** allow them both leftward mobility.
- *the areal nature* of Feature 13 — including four very different languages participating in this Sprachbund (Talyshi, Armenian, Neo-Aramaic, Udi) — will be plotted onto isogloss maps. I will draw parallels to the mobility of verbal clitics in the perfects of certain Balkan languages and contrast Feature 13 with the non-mobility of Persian Set₁ clitics in the perfect tenses.

Finally, I will show that Set₁ and Set₂ represent two totally independent, interacting clitic types within the verbal system and that the factors triggering the mobility of each set to a variety of hosts within the clause are totally different. The mobility of Set₁ in the four different languages is clearly triggered by the information structure of the clause (Data, Point 3). The focused word of the clause (generally marked by clause-level stress) serves as host for Set₁. The triggering factors for Set₂ mobility in the Iranian languages of this area (N. Talyshi, N. Tati), however, involve neither stress, nor Wackernagel position, nor object marking, but so far remain elusive.

Data (Set₁ enclitics only)

(*singular paradigms only; feminine forms deleted for Aramaic*)

	<u>Neo-Aramaic</u>		<u>Armenian</u>		<u>N. Talyshi</u>	
Set₁	enclitic	verbal	enclitic	verbal	enclitic	verbal
Forms	copula	suffix	copula	suffix	copula	suffix
	(Feat. 5, 7)	(Feat. 6)	(Feat. 5, 7)	(Feat. 6)	(Feat. 5, 7)	(Feat. 6)
1sg.	=(i)vin	-in	=em	-em	=im	-im
2sg.	=(i)vit	-it	=es	-es	=iš	-iš
3sg.	=(i)li	-Ø	=e	-i	=e	-i

1. Set₁ function predicatively as enclitic forms of the copula (Feature 5)

<u>Neo-Aramaic</u>	<u>Armenian</u>	<u>N. Talyshi</u>
šidæenæ=vin	xént-em	jondó-m
šidæenæ=vit	xént-es	jondó-š
šidæenæ=li	xént-e	jondó-e [jondóy]
crazy=Set ₁	crazy=Set ₁	crazy=Set ₁
"I am crazy, you are crazy, he is crazy."		

2. Feature 6: Set₁ as **suffixes** (no mobility) vs. Feature 7: Set₁ as **enclitics** (highly mobile)

Subjunctive			Periphrastic Present		
Neo-Aramaic	Armenian	N. Talyshi	Neo-Aramaic	Armenian	N. Talyshi
Ø-xæzy-in	tesn-Ø-ém	bí-vind-im	bə-xzáya=vin	tesn-úm=em	vín-dæ=m
Ø-xæzy-it	tesn-Ø-és	bí-vind-iš	bə-xzáya=vit	tesn-úm=es	vín-dæ=š
Ø-xæzi-Ø	tesn-Ø-i	bí-vind-i	bə-xzáya=li	tesn-úm=e	vín-dæ=Ø
SUBJ-see-Set ₁	see-SUBJ-Set ₁	SUBJ-see-Set ₁	PRES-see=Set ₁	see-PRES=Set ₁	see-PRES=Set ₁
“that I see, that you see, that he see” (Subjunctive)			“I see, you see, he sees” (where PRES is also a locative adposition grammaticalized as a TAM)		

3. Feature 13: Set₁ may float leftwards in the sentence (clause-stressed syllable is bolded)

Non-fronted Version			Fronted Version		
Aramaic:	é̄næ ktǣvæ bə-qr̄áya=vin.		é̄næ ktǣvæ =vin bə-xzáya.		
Armenian:	yes girk ḵart-úm =em̄.		yes girk =em̄ tesn-úm.		
Talyshi:	az kitób handé-dæ =im̄.		az kitób =im̄ vín-dæ.		
	I book read-PRES=SET ₁		I book=SET ₁ see-PRES		
	“I am <i>reading</i> a book.” (With stress on the verb for reasons of contrastive focus: I am <i>reading</i> a book, as opposed to <i>writing</i> one)		“I see a book.” (with neutral focus and unmarked place- ment of stress on the element before the verb, often the direct object)		
<u>Distance Fronting</u> (leftward mobility to any word not contiguous to the verb)					
Aramaic:	mæn =il̄i ktǣvæ bə-qr̄áya?		é̄næ =vin̄ ktǣvæ bə-qr̄áya.		
Armenian:	ov =ē girk ḵart-úm?		yes =em̄ girk ḵart-úm.		
Talyshi:	ki =ē kitób handé-dæ?		az =im̄ kitób handé-dæ.		
	who=SET ₁ book read-PRES		I=SET ₁ book read-PRES		
	“ Who is reading a book?”		“ I am reading a book.” (not someone else)		

(Udi has isomorphic correspondences in Points 1-3 but equivalent data was not available for this abstract.)

Where was the Homeland of Malay?

Uri Tadmor

In 2000, a colloquium entitled ‘Borneo as the Homeland of Malay’ was held in Malaysia. The participants, who included the leading experts in the field, seemed to agree that the title of the colloquium was appropriate, and that the homeland of Malay was indeed Borneo. This paper points out a few problematic aspects of this hypothesis.

One problem is the inability, or reluctance, to distinguish between Malay (a language) and Malayic (a group of related languages). The distinction may not be very consequential for the reconstruction of Proto-Malayic, but it is of crucial importance for situating the homeland. After all, there is no logical reason why the place where Malay first arose as a language should be identical to the place where its ancestor Proto-Malayic originated.

A more serious problem, however, lies with the main argument used to support the Borneo homeland hypothesis, namely diversity. It has long been accepted by historical linguists that the area where a language (or a group of related languages) exhibits the greatest diversity is the most likely site for the original homeland of the language (or proto-language) in question. Nonetheless, this criterion should not be used in an unqualified manner. Specifically, the diversity argument is only valid when the diversity in question is the result of internal language change. Contact-induced change can be much more rapid than internal change; in fact, it may cause the emergence of a new isolect within as little as one generation. Therefore, diversity brought about by language contact cannot be used in support of a hypothesis regarding the homeland of a language.

This paper suggests that much of the linguistic diversity among the so-called Malayic isolects of Borneo is due to language contact, and therefore does not constitute strong evidence for the ‘Borneo as the Homeland of Malay’ hypothesis. It is, at best, a weak argument; in the case of substrate influence, it may even constitute counter-evidence. Finally, it is suggested that based on linguistic and other evidence, Sumatra is a more convincing locus for the homeland of Malay.

Interpretation of the Notional Subject in Serial Verb Construction in Seediq

Naomi Tsukida

Seediq is an Austronesian language spoken in Taiwan. This language has serial verb construction, in which the subject of the non-main verb always cannot appear. The notional subject must be coreferent with the actor (A or S) of the main verb. This pattern of coreference is interesting in that two different entities are involved: the actor and the subject. The one which is coreferred to is the actor of the main verb and the one which corefers is the (notional) subject of the non-main verb.

Seediq is a predicate initial language and the verb that comes first in the clause is the main verb. Non-main verbs follow the main verb (example (1)). Non-main verbs are usually the neutral form, regardless of the tense/aspect of the main verb (example (2)).

The notional subject of a non-main verb is interpreted to be coreferent with the actor of the main verb. In the example (1), the one who works is the one who goes, and in the example (2), the one who bought is the one who went.

One might think that such cases can be explained in either of the following two ways: (a) the two verbs share the same subject or (b) the two verbs share the same actor. Such explanations are both rejected, as shown below.

This language has so-called Philippine type focus system, in which one can change the voice of the clause and make nominals of various semantic roles into the subject, when pragmatically necessary. Sentences that involve serial verb construction can be Actor Voice, Goal Voice or Conveyance Voice, and examples of the latter two voices make us reject the explanation (a). Example (3) is GV and has the Patient as the subject of the main verb. The notional subject of the non-main verb, however, is coreferent with the actor of the main verb (= *nami* 'we (exclusive)'), not with the subject. Example (4) is CV and has the Beneficiary as the subject of the main verb. The notional subject of the non-main verb is coreferent with the actor of the main verb (= *na* 'she/he').

The non-main verbs also can be GV or CV (example (5) and (6)). Such examples make us reject the explanation (b). In such examples, the referent of the notional subject of the non-main verb, which is Goal or Conveyed theme, is interpreted as coreferent with the actor of the main verb.

Because of this constraint, only 'same subject serialization (Crowley 2002: 40)' is allowed, and 'switch subject serialization (Crowley 2002: 40)' is not. If one wants to say 'A forgives B to go' one must serialize 'to forgive' and 'to make go', not 'to forgive' and 'to go' (example (7)).

Seediq belongs to Formosan languages, which has close relationship with the Philippine languages. Philippine languages are often said to be information-salient languages (e.g. Klaiman 1991: 245), and many 'subject' properties are attributed to the 'topic' and few to the actor (e.g. Shibatani 1988). In Seediq also, many syntactic properties are attributed to the 'subject', which correspond to this 'topic' in Philippine literature. To be referred to by the null subject of non-main verbs in serial verb construction is one of the few syntactic properties of the actor in Seediq.

- (1) m-usa q-em-pahka se'diq niyi.
 AV-go AV-work NOM person this
 This person goes to work.
- (2) m-en-sa m-arig bawa ka laqi.
 AV-PRF-go AV-buy steamed:bread NOM child
 The child went to buy a steamed bread.
- (3) diy-un=nami t-em-atak rena'aw ka sawki ni'i.
 hold-GV1=1pe.GEN AV-cut bush NOM sickle this
 We will hold (=use) this sickle to cut down a bush.
- (4) se-'usa=na m-aNal sapuh ka payi=na.
 CV-go=3s.GEN AV-take medicine NOM old:woman=3s.GEN
 S/He goes to take medicine for her/his grandmother.
- (5) me-'isug kese'N-an ka laqi ni'i.
 AV-fear scold-GV2 NOM child this
 This child fears to be scolded.
- (6) s-em-ekuxul se-'apa ka laqi ni'i.
 AV-like CV-carry:on:back NOM child this
 This child likes to be carried on one's back.
- (7) s-em-eruwa pawsa leqi-'an gesiluN ka tama.
 AV-forgive AV.CAUS.go child-OBL sea NOM father
 The father forgives and let children go to the beach.

References

- Crowley, Terry. 2002. *Serial verbs in oceanic: a descriptive typology*. Oxford: Oxford University Press.
- Klaiman, M.H. 1991. *Grammatical Voice*. Cambridge University Press.
- Shibatani, Masayoshi. 1988. Voice in Philippine languages. In: Shibatani, Masayoshi. (ed.) 1988. *Passive and voice*. John Benjamins publishing company. 85-142.

Sensory and Cognitive Perceptions in a Crosslinguistic Perspective: A Semantic and Lexical Analysis

Martine Vanhove & Bruno Gaume

The interest for semantic associations between the lexicon of sensory and cognitive perceptions has been growing since the last two decades because of the regularity of semantic changes and the highly polysemous terminology noted in many languages and language phyla and the consequences thereof for cognitive linguistics. In her work *From Etymology to Pragmatics*, Eve Sweetser (1990:21) was underlying the fact that in Indo-European languages, and probably also crosslinguistically, “Deep and pervasive metaphorical connections link our vocabulary of physical perception and our vocabulary of intellect and knowledge” and that “a metaphorical analysis motivates the otherwise strange fact that certain semantic sub-domains within perception are naturally and regularly historical sources for certain sub-domains of cognition, rather than others.” Since then, the universal claim that the semantic structure of sensory verbs extends into the domain of cognition was tested for other language phyla, e.g. in Australian languages (Evans and Wilkins 1998) for the notion of ‘hearing’. The reverse was also claimed, i.e. that the connection between the two domains was particular, to some extent, to a particular typological area, e.g. Africa for the lexicon of sensory perception and understanding (Heine and Zelealem (to appear) who use the methodology of isopleth mapping). Because of these contradictory claims, we thought it worth investigating the matter crosslinguistically in more details.

In a joint project about the typology of semantic associations in the lexicon from a synchronic (polysemy) and diachronic (semantic change) viewpoint, we investigated the lexical domains of perception and cognition in a first sample of 23 languages (see appendix 1) for the notions of ‘hear’ / ‘listen’, ‘see’, ‘feel’ (and also ‘take’) and their extension into the domains of ‘heeding’, ‘obeying’, ‘learning’, ‘understanding’, ‘knowledge’, ‘desire’, ‘intention’. The work is mainly based on first hand data collected by the field linguists participating in the project and is, so far, restricted mainly to the verbal category and its derivatives. If the results achieved show a pervasive connection between the two semantic domains crosslinguistically, they also show that it cannot be considered as a universal property of languages at the lexical level. Still interesting typological observations can be made, some of which will be discussed in more detail, i.e. the links between ‘hear’ / ‘listen’, ‘see’ and the domain of cognition.

The most pervasive association is between ‘hearing’ and intellectual perception and it is remarkable that the extension of ‘hear’ concerns all the cognition sub-domains mentioned above. It occurs in all the languages of the sample but one. Further semantic investigation shows various limits to the extensions into the domain of intellectual perception according to the languages, some of them being obviously due to social reasons (e.g. ‘obeying’ is simply unknown in non hierarchical societies). A tentative classification is proposed according to the intellectual sub-domain concerned.

Less pervasive is the association between ‘see’ and intellectual perception which is limited to 13 of the 23 languages of the sample. Furthermore the semantic analysis shows a limitation of the semantic association in the lexicon to the sole cognitive sub-domains of understanding and knowledge (not withstanding other attested semantic associations outside the domain of cognition). This finding at the lexical level goes contrary to the alledged belief that vision is cognitively and universally connected with knowledge.

The investigation of the semantic associations between physical manipulation and cognition, which will not be presented into detail, leads to the following typological correlation hypothesis, which is valid for all the languages of the sample:

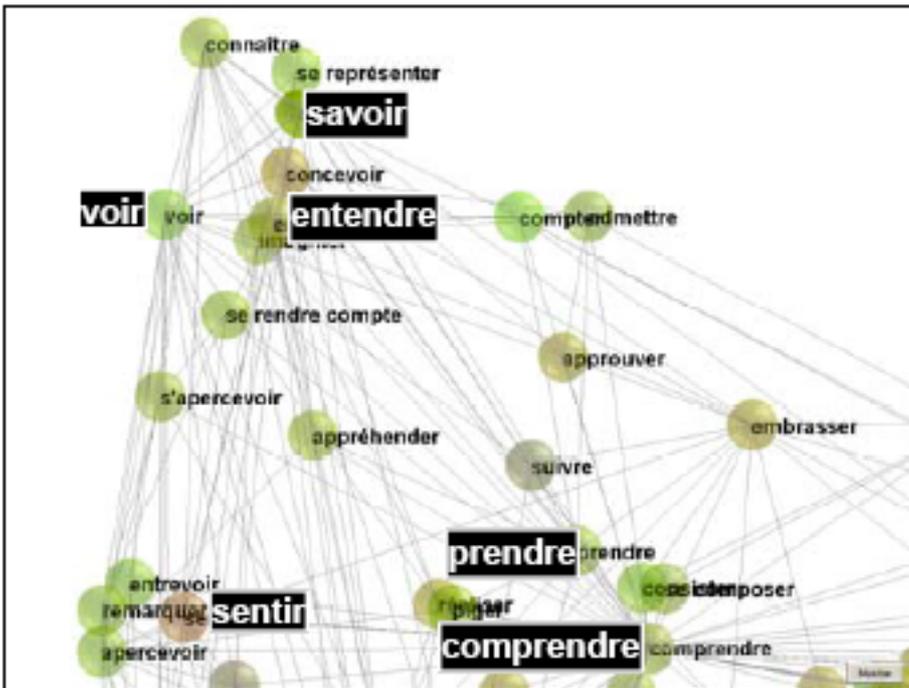
If a language has a semantic association between physical manipulation ('take') and cognition it also has it at least with one physical perception ('hear', 'see', or 'feel'), but the reverse is not true.

To conclude, the typological results of the empirical research will be compared with the results obtained for French with Prox, a newly developed mathematical and computerized approach of the lexical structure of a language through a 3D spatialization of dictionary entries. This device shows that the higher the number of paths between a word and another, the closer their proximity in space, and thus, semantically, which is just the case for the semantic fields under survey (see appendix 2 which shows the semantic structure calculated with Prox for 'understand' and 'know' in French in connection with sensory and manipulation verbs, using dictionaries of synonyms).

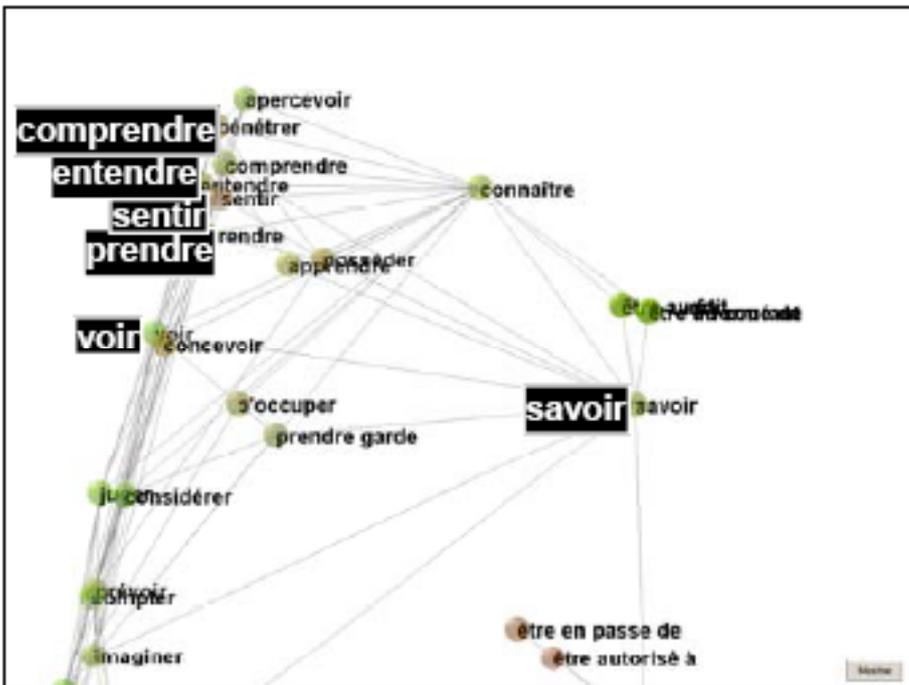
Appendix 1 : Languages under survey

ARABIC (CLASSICAL) (Semitic, Afroasiatic)
ARAKI (center-north Vanuatu, Oceanian, Austronesian)
BEJA (north Cushitic, Afroasiatic)
ENGLISH (Germanic, Indo-European)
FRENCH (Romance, Indo-European)
GBAYA 'BODOE (Gbaya-Manza-Ngbaka, Niger-Congo)
GERMAN (Germanic, Indo-European)
INUIT (eastern Eskimo)
ITALIAN (Romance, Indo-European)
KASIM (western Gurunsi, Gur, Niger-Congo)
LAKON (center-north Vanuatu, Oceanian, Austronesian)
MAKONDE (Bantu, Niger-Congo)
MWOTLAP (center-north Vanuatu, Oceanian, Austronesian)
OLRAT (center-north Vanuatu, Oceanian, Austronesian)
PALENQUERO (Spanish based Atlantic Creole of Colombia)
RUSSIAN (Slavic, Indo-European)
SAR (Sara-Bongo-Bagirmian, Nilo-Saharan)
SWAHILI (Bantu, Niger-Congo)
TAMANG (Tibeto-Burman, Sino-Tibetan)
VILI (Kikongo dialect, Bantu, Niger-Congo)
WOLOF (west-Atlantic, Niger-Congo)
YULU (Sara-Bongo-Bagirmian, Nilo-Saharan)
YUPIK (western Eskimo)

Appendix 2: Semantic structure of verbs: a. 'understand', b. 'know'



(a)



(b)

References

- Duvignau, Karine and Gaume, Bruno. 2004. Linguistic, Psycholinguistic and Computational Approaches to the Lexicon: For Early Verb-Learning. *Cognitive Systems*, Janvier, Vol 6 (1) (to appear)
- Gaume, Bruno, Hathout, Nabil and Muller, Philippe. 2004. Word sense disambiguation using a dictionary for sens similarity measure in *Acte COLING 2004, The 20th International Conference on Computational Linguistics, COLING 2004*, Geneva.
- Evans, Nicholas and Wilkins, David. 1998. The knowing ear: An Australian test of universal claim about the semantic structure of sensory verbs and their extension into the domain of cognition. *Arbeitspapier. Institut für Sprachwissenschaft, Universität zu Köln*, 1998, vol. 32, Neue Folge.
- Heine, Bernd and Zelealem, Leyew. to appear. Is Africa a linguistic area? in Heine, B. and Nurse, D. (eds.), *Africa as a linguistic area*.
- Sweetser, Eve. 1991. *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

On the Distinction Between Core and Oblique Arguments in the Austronesian Languages of Indonesia

I Wayan Arka

The classification of arguments into core and oblique is perhaps one of the key distinctions in grammar. Core arguments share certain properties that obliques or adjuncts lack. Broadly, core arguments are A, P (or O), and S (as defined in Comrie 1978 and Dixon 1979) together with any other argument which shares the core properties of the language.

In recent years there have been a growing number of studies focussed on the individual Austronesian languages of Indonesia. General information on core and oblique relations is more or less now available. However, the exact nature and significance of the core/oblique distinction for a comparative study of the Austronesian languages of Indonesia have generally not been well explored. The main purpose of this paper is to develop a typology of the Austronesian languages of Indonesia on the basis of whether or not core/oblique distinctions are made, and if they are, in what ways. Another is to utilise advances in current research on Indonesian-type languages (mainly my own on Balinese and Indonesian) to probe into other Indonesian languages of different types, particularly languages of eastern Indonesia.

Indonesian and Balinese can be demonstrated to show clear core/oblique distinctions, mainly thanks to the existence of voice alternations and applicatives in these languages. The distinctions are also manifested in a number of subtle grammatical properties (quantifier float, topicalisation of possessor phrase, resumptive pronoun, depictive predicate and reflexive binding) and coding properties (categorical expressions, order and arrangement).

Problematic cases among the languages of eastern Indonesia will be discussed where voice and/or applicativisation are unusual or absent. It remains a matter of debate whether core/oblique relations could be posited at some level of grammatical representations for these languages (cf. Croft 2001, Dryer 1997). We conclude that while overt coding often provides a good start to identifying core/oblique distinctions in some languages, in other languages these distinctions are often manifested in language-specific properties which need an in-depth understanding of the semantics, grammar and discourse of the language under investigation.

References

- Comrie, Bernard. 1978. Ergativity. In *Syntactic typology: Studies in the phenomenology of language*, edited by W. P. Lehman, 329-394. Austin: University of Austin Press.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
- Dixon, R.M.W. 1979. Ergativity. *Language* 55:59-138.
- Dryer, Matthew S. 1997. Are grammatical relations universal? In *Essays on language functions and language type, dedicated to T. Givon.*, edited by J. L. Bybee, J. Haiman and S. A. Thompson, 115-143. Amsterdam: John Benjamins.

Prolegomena towards a Typology of Relexification Acts

Paul Wexler

Relexification is the process of replacing one's native lexicon by foreign "phonetic strings" which are calibrated according to the morphosyntactic and lexico-semantic parameters of the native lexicon. The grammar and phonology of the relexified language undergo no major changes while the lexicon is now by and large formally non-native. Native speakers and non-native observers alike typically believe that the relexified language is related genetically to the language which provided the phonetic strings.

Over the last half century, relexification has been identified in the birth of a number of creole and non-creole languages, e.g. Southeast Asian Portuguese Creole (Kristang), colloquial Arabic (other than Yemeni), Medieval and Modern Hebrew, Yiddish, Old Church Slavonic, Mbugu (or Ma'a, Northeast Tanzania), Michif (a mixed Cree-French language in the Canadian Prairie provinces, North Dakota and Montana), Haitian, Papiamentu (Curaçao), and other Caribbean Creoles, Palenquero (Colombia), Media Lengua (Ecuador), Callahuaya (or Machaj Juyai, Bolivia), Damin (Australia), Shelta and Anglo-Romani (British Isles), some forms of Irish English, Esperanto, Medieval Latin, some forms of contemporary Belarusian, Ukrainian and Sorbian (eastern Germany), and Canaanite Akkadian, to name just a few.

There is still considerable disagreement about whether some of the above languages were indeed created by relexification; cases *sub judice* include Russian, Hungarian, Rumanian and non-British Romani. Nevertheless, there are now about a dozen diagnostic tests that can be applied to determine if relexification might have taken place: prime among these tests is our ability (i) to predict (i.e. motivate retroactively) which lexical elements (phonetic strings) from the superstratum can be accepted by the relexifying language, and (ii) to posit different sources for the lexicon of a language and its phonology and grammar. *The present paper will examine some of the above case studies of relexification in an attempt to identify the major variables that are likely to influence the decision to relexify a language and the resulting form the relexification is likely to take.*

This paper discusses the importance of the following nine variables:

- (i) Accessibility to more than one lexifier language in the relexification process (e.g. Yiddish relexifies simultaneously to medieval German, Classical Hebrew, Balkan Romance).
- (ii) Accessibility to only a modest unique lexicon (e.g. the Romá [Gypsies] relexify their native non-Romani languages to approximately 400 unique, primarily Indic, Iranian and Byzantine Greek roots; Bahasa Indonesian recalibrates Arabic *alamat* 'sign' and *darurat* 'necessity' to 'address' and 'emergency', respectively).
- (iii) Accessibility to an unspoken language of liturgy or literature as the lexifier language (e.g. Yiddish and Israeli Hebrew relexify to unspoken Classical Hebrew).
- (iv) The relative chronology of relexification of a spoken language and of an unspoken language of liturgy (e.g. spoken and liturgical Judeo-Spanish, Yiddish).
- (v) The likelihood of *inventing* a lexifier source (e.g. Yiddish invents Hebrew lexicon).
- (vi) Since the motivation for relexifying is exclusively identificatory and not communicatory, we need to explore (a) whether the speech community exhibits non-linguistic analogues to relexification (e.g. a new religion) and (b) what the relative chronology of the non-linguistic analogues and the relexification process

is (e.g. a newly syncretistic crypto-Judaism was practiced by speakers of relexified Iberian Judeo-Arabic; the myth of a Jewish “people” is maintained by relexifiers of Yiddish to Classical Hebrew in order to create Modern Hebrew).

- (vii) New speakers of a language relexify (e.g. Finno-Ugric and Baltic speakers relexify their adopted Russian to unspoken Old Church Slavonic; non-Jewish converts relexify their native Slavic languages to German to create Yiddish, which is adopted by Slavic-speaking Jews).
- (viii) The existence of a pattern of cyclical relexification in the history of the speech community (e.g. the languages spoken by North European Jews over the last millennium).
- (ix) The attempt, after the completion of the relexification process, to “reverse” the process by imitating the lexifier language itself (e.g. the creation of a so-called “creole continuum” in Indonesian and Malaysian Portuguese Creole [Kristang], by replacing unique Portuguese lexicon with, and by becoming bilingual in, standard Portuguese; the Classical Arabization (Semiticization) and standardization of all African and Asian dialects of colloquial Arabic).

An Inflecting Question Particle in Oroqen

Lindsey Whaley

Oroqen, a Northern Tungusic language spoken in China, possesses a typologically unusual feature of a question particle that can be inflected for voice, aspect, tense, and person/number. This paper describes the conditions under which the inflected particle can be used and the forms that it can take. The paper concludes with speculations about the origin of the inflected particle.

Oroqen, like all Tungusic languages, has a set of sentence final evidential particles that can be used to express a speaker's degree of certainty or surprise about the information contained in a sentence (1b+c). The set of particles includes *yE*, which is used optionally in the formation of yes-no questions (1d).

Among these particles, only *ye* can be inflected. When it is, it gives rise to alternative interrogatives (2a+b) in which the affirmation of negation of the same clause is being questioned. Although *ye* can carry the identical set of inflectional suffixes as the verb (as in 2a+b), it is only required to be marked for tense and person/number (cf. 1c+d), and it cannot be marked for any category that is not also marked on the main verb (1e).

It is important to note that Oroqen also creates alternative questions by using a negative auxiliary (3), which is apparently identical in meaning. This construction is the usual means in the Tungusic language family for creating alternative questions. Indeed, it is the only means employed by all the other Northern Tungusic languages, which raises the question as to why the inflecting question marker exists in Oroqen. We speculate that the construction is derived from a pro-verb **ya*, which still occurs in one southern Tungusic language, Udihe, where--tellingly--it is used for alternative questions. While this pro-verb has been lost in most Tungusic languages, it has remained in Oroqen, where the root has been re-analyzed as one of the sentence final evidential particles.

Data

- (1a) tari murin ukti-tse
 that horse run-PST
 'That horse ran.'
- (1b) tari murin ukti-tse o:
 that horse run-PST EVID
 '(Maybe) that horse ran'
- (1c) tari murin ukti-tse dze
 that horse run-PST EVID
 '(Of course) that horse ran.'
- (1d) tari murin ukti-tse (ye)
 that horse run-PST QM
 'Did that horse run.'
- (2a) bi: tari-duk yarala-p-tsa-w ye-p-tsa-w
 1S that-ABL injure-PASS-PST-1 SQM-PASS-PST-1S
 'Was I injured by him or not?'
- (2b) bi: beyu-mun-e-m ye-mun-e-m
 1S hunt-DESID-epen-1S QM-DESID-epen-1S
 'Do I want to go hunting or not?'

- (2c) bi: tari-duk yarala-p-tsa-w ye-tsa-w/*ye-p/*ye-w/*ye-tsa
 1S that-ABL injure-PASS-PST-1 SQM-PASS-PST-1S
 ‘Was I injured by him or not?’
- (2d) bi: beyu-mun-e-m ye-m/*ye-mun/
 1S hunt-DESID-epen-1S QM -1S
 ‘Do I want to go hunting or not?’
- (2e) *bi: beyu- m ye-mun-e-m
 1S hunt-1S QM-DESID-epen-1S
 ‘Do I want to go hunting or not?’
- (3) si gune-ndi e-si-ndi
 2S go-2S NEG-AOR-2S
 Are you going, or not?

The Portrait of Minangkabau Women as Depicted in Various Proverbs and Sayings

Yusrita Yanti

This article is to investigate the social and cultural norms of the Minangkabau society especially in terms of women's role as described in the various proverbs and sayings. In the manners and customs of Minangkabau, a woman has two functions namely (1) as a *parampuan* (woman), and (2) as *Bundo Kanduang* (a married woman). As a *parampuan*, she has to work hard and keep trying to do something (*namuah bajariah dan bausaho*) she is able to draw and make holes in the materials (*pandai malukih manarawang*), she is able to engrave and weave (*pandai maukie jo batanun*), and knows how to behave and wise in making decision (*tahu di suri mato karok, tahu di batang rabah tagak, arif bijaksano digunokan*). If Minangkabau women do not possess those characters, she is called as *simarayuan* (who does not have self concept) and *mambang tali awan* (an arrogant/conceited woman). The two names belong to the other types of women in Minangkabau society. As a married woman-*Bundo Kanduang*- she plays an important role in the family and society. She is not only has a function of educating children and possessing properties but also as a symbol of honor in the society (Hakymi, Dt. Rajo Penghulu. 1991; Zulkarnaini. 1994). Such roles can be seen in the proverbs expressions "... *limpapeh rumah nan gadang, sumarak dalam nagari, hiasan di dalam kampuang, nan tahu jo malu sopan, hiasan kampuang jo halaman, langkok ka koto jo nagari, sampai ka balai jo musajik, panyusun sumarak rumah tango, laku bahieh budi baiak...*"(Madjo Indo, 1999.) These proverbs and sayings describe that *Bundo Kanduang* is a symbol of honor and magnificence, she is not only physically decoration in the society but she has to comprehend the manners and customs of Minangkabau and must have self concept. Semantically, the diction used have deep meaning and tend to have connotative and figurative meaning that imply all of the rules and aspects of life in Minangkabau traditions. It is very important for people, in this case Minangkabau women to read the various proverbs and sayings comprehensively in order to make them understand how important their functions in running the daily life based on the customs and tradition of Minangkabau. Hopefully, the great values implied in the proverbs and sayings can be applied in all of aspects of life in developing the nation and Indonesian country.

Symposium: Sign Language Typology

Ulrike Zeshan

This symposium is intended to familiarize typologically-minded linguists with the newly emerging linguistic sub-discipline of sign language typology. The results of sign language research are still not being sufficiently integrated into typological linguistics, despite their obvious potential for opening up new perspectives in the field. Sign language typology seeks to apply typological considerations to the cross-linguistic study of signed languages and to integrate results from sign languages research into linguistic typology. We propose to provide a comprehensive introduction to this field of study in four sections. The first presentation will set the background for the other topics with basic information about the communities where sign languages are used. The second presentation focuses on linguistic diversity within the group of sign languages and describes cross-linguistic case studies in sign language typology. In the third presentation, we present a particularly interesting instance of linguistic diversity across sign languages, the sub-group of sign languages used in village communities. Finally, the last presentation uses a cross-linguistic typological framework to explore some of the unresolved theoretical problems facing sign language linguistics.

1. Sibaji Panda: "Deaf communities and sign languages"

This presentation provides relevant sociolinguistic background in sign language research. The language situation of various sign language communities around the world, different types of signed communication systems, the development and main current trends in sign language research, the current stage of sign language documentation around the world, and issues of applied sign linguistics, language planning and language policy are explored in detail. In addition to what is known about deaf communities in industrialized countries, the presentation focuses on the deaf community in India and its sign language as an example from a developing country with a complex linguistic situation (Sethna, Vasishta & Zeshan 2005). Since the presenter is a deaf Indian, all issues are considered from an "insider's" point of view. This presentation will be in International Sign with simultaneous interpretation into spoken English.

2. Ulrike Zeshan: "Aims and methods of sign language typology: Case studies in interrogatives, negatives and possession"

Sign language typology involves both the typologically informed description of individual undocumented sign languages, and typological studies of a particular linguistic domain across a large sample of sign languages. The cross-linguistic comparison of signed languages then leads to a theory of variation across sign languages, based on inductive higher-level generalizations that emerge from the data. We can ask what range of typological variation we find across sign languages in various domains of grammar and why, and what characterizes sign languages as a linguistic type distinct from spoken languages. To exemplify the methodology of large-scale cross-linguistic studies on sign languages, two case studies in sign language typology are presented: A summary of a completed typological study on interrogatives and negatives across 37 sign languages (Zeshan 2004a, 2004b), and first results from an ongoing study on possession, involving data collection through visual stimulus materials and through questionnaires. The design and methodology of the projects is discussed, and analysis of some of the results presented, demonstrating the range and kind of typological variation, with videotaped examples from various sign languages.

3. I Gede Marsaja: "Village-based sign languages"

Village-based sign languages are used in villages with a high incidence of hereditary deafness over several generations. The socio-cultural situation in such villages, where most of the hearing population is also fluent in the village sign language, has been documented in several regions around the world (e.g. in the Urubu-Kaapor tribe in northern Brazil, in a village in Ghana, in an Arab Bedouin tribe in Israel), but detailed linguistic documentation of some of these languages is only just beginning. Initial results indicate that these sign languages are in many ways different from what researchers have come to expect to find in urban sign languages, and that they have the potential to call into question many of the "sign language universals" that have previously been posited. One of the first village-based sign languages to be studied in detail is the Kata Kolok ("deaf language") in a village in northern Bali (Branson, Miller & Marsaja 1999). After an overview of other village sign languages, the presentation mainly focuses on Kata Kolok, based on extensive fieldwork by the presenter. Some interesting linguistic structures are described as an example of how village sign languages add another perspective to linguistic diversity in the framework of sign language typology.

4. Waldemar Schwager: "The sign and the word – current theoretical problems in sign linguistics"

Although sign language research has made much progress over the past 30 years, many basic theoretical problems remain unresolved. This presentation explores some of these issues in a cross-linguistic and typological perspective, looking at both the sign unit across sign languages and the word unit across spoken languages. Sign language linguistics so far lacks an explicit workable definition of the morpheme, with the result that the analysis of some kinds of complex constructions (e.g. the so-called "classifier constructions", e.g. Emmorey 2003) is highly controversial. Even more critical is the distinction between the phonological and the morphological level in sign languages because the use of iconicity appears to lead to a large number of meaning-bearing units of analysis at the phonological level (Fernald & Napoli 2000). Finally, it is also problematic to distinguish the equivalent of phonetic features versus phonemes in sign languages for most formational aspects of signs. Taking the sign unit as the basis of the argument, this presentation pursues possible avenues towards solving some of these theoretical problems, which in turn should lead to a deeper understanding of the similarities and differences between sign languages on the one hand and spoken languages on the other hand. This presentation will be in International Sign with simultaneous interpretation into spoken English.

References

- Branson, Jan, Don Miller & I Gede Marsaja (1999): Sign language as a natural part of the linguistic mosaic: the impact of deaf people on discourse forms in North Bali, Indonesia. In: Winston, Elisabeth A. (ed): *Storytelling and conversation: discourse in deaf communities*. (Sociolinguistics in Deaf Communities; 5) Washington, DC : Gallaudet University Press.
- Emmorey, Karen, ed. (2003): *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Erlbaum.
- Fernald, Theodore B. & Donna J. Napoli (2000): Exploitation of morphological possibilities in signed languages: Comparison of American Sign Language with English. In: *Sign language and linguistics* 3: 1, pp. 3-58.
- Sethna, Meher, Madan Vasishta & Ulrike Zeshan (2004): Implementation of Indian Sign Language in educational settings. *Asia-Pacific Journal of Disability and Rehabilitation* Vol. 15. No.2, pp. 15-39.
- Zeshan, Ulrike (2004a): Hand, Head and Face – Negative Constructions in Sign Languages. *Linguistic Typology* 8:1, 1-58.
- Zeshan, Ulrike (2004b): Interrogative constructions in sign languages – Cross-linguistic perspectives. *Language* 80.1, 7-39.

Semantic Parameters for Typological Identification of Derivational Categories of the Ket Verb.

Marina Zinn

Ket is a language isolate spoken by a small scattered group of indigenous people of North Siberia, its verb morphology displays templatic structure, lack of purely inflectional categories, idiosyncratic selection of agreement model (=actant conjugation) in the domain of lexicon (a typologically unique situation). The derivational relations between the so called infinitive and the finite verb as the infinitive displays far simpler and non-templatic morphology, moreover it is not always solely the base for finite verb form derivation as the template matrix itself is a marker of verb formation and the derivational meaning (the same function as affixes in concatenative morphology). Main morphological mechanisms for verb derivation are 1) preselection of positions and corresponding affixes in the template as well as the configuration of positions according to derivation category, 2) recycling inflectional positions (e.g. actant markers) for derivation, 3) conjugation switch 3) incorporation, 4) verbal root selection (removing the input root or shifting it to the incorporate position).

In this paper the author makes an attempt and describe Ket verb derivation as a system pairing the semantic parameters in play signaled by means of morphological mechanisms. The main fact in the focus of attention is an attempt to describe the combination, interlinking with the morphological form and interrelation of semantic parameters in the verb derivatives grouped into derivational (and speaking more generally grammatical) categories. As we proceed we outline the derivational patterns and the specific features of the categories; the basic component (and typically linked subcomponents) is put forward as the basis for cross-linguistic category identification. In addition we point out the cases of interweaving of the categorial semantics yielding to the new meaning shades production, and even cross-categorial replacement. The data under study is the finite verb derivatives grouped into the main categories according to the similar meaning, and according to the similar derivational pattern (the combination of input and output semantic and morphological constructions).

The approach to the semantic form is viewing the complex event structure as the hierarchy of super/subevents that are interrelated (the relations relevant here are CHANGE TO, CAUSE, POSS; types of events {action} and {state} are postulated. The arguments mapped onto the event structure are finally reflected in the semantic structure as S (subject) and O (object) that is bringing forth the rank in the argument hierarchy. Besides, it is significant for diathetic shift of the Ket verb whether the features [+/- control], [cause], [instrument] that are commonly viewed as inherent to the notions like agent or instrument, can be transferred from the argument of one rank to the other.

See below the outline of the paper and derivational categories briefly introduced.

A. Main semantic process: event structure derivation

I. Derivational categories derived from infinitive or nominal. Main morphological process: incorporation. **Inceptives** (4 patterns for intransitive verbs only 1) CHANGE TO¹⁶ {action} + infinitive incorporation, 2) CHANGE TO {action} [-control] + infinitive incorporation, 3) CHANGE TO {action} POSS – incorporating the possessum noun, 4) CHANGE TO {action} + event noun incorporation), **Translative** (CHANGE TO {state}, 4 patterns, state is specified by incorporation of noun, adjective, adverb, and one pattern overlapping with causative formation), **Causatives** (2 types: CAUSE {action} differ for iterative/momentaneous split in derivational patterns, and CAUSE {state} – Type 2 causative in addition serves as inceptive for transitive verbs, displays 4 patterns, differing in transitivity and iterativity).

II. Derivational categories derived from finite verb Main morphological process: verbal root selection **Iterative/Distributive** (multiple event/multiple object), the most pervasive category, combines with causative formation.

B. Main semantic process: argument transfer Main morphological process: conjugation switch, recycling actant markers for derivation. Categories are further classified according to valence change: mainly valence decrease which display the transfer of the object to subject $Sx + Ox \rightarrow Sx + Sy$ ¹⁷ (**Reciprocal, Reflexive, Autocausative, “Autoinstrumental”**) plus removing the input subject $Sx + Ox \rightarrow \emptyset y + Sy$ (**Anticausative** Sx [cause] + $Ox \rightarrow \emptyset y + Sy$ [cause], **Resultative** Sx [cause] + Ox {action} $\rightarrow \emptyset y$ [\emptyset] + Sy {state}) or deleting the object (Deobjective $Sx + Ox \rightarrow Sy + \emptyset y$); then a small number of valence increase: adding an **a core or peripheral argument** (**Causative**¹⁸ $/Sx$ {action}/ $\rightarrow Sy$ [cause] + $/Oy$ {action}/ and **Instrumental applicative** $Sx + Ox/ \rightarrow /Sx + Ox/ + [Oy$ Instrument]) or interpretation (inverse type: Involuntary causative [-control] feature shift). The possibility of second step valence change are briefly illustrated.

¹⁶ Here CHANGE TO relation connects zero superevent to the starting subevent, that is the change is in the status of event existence

¹⁷ The left part of the equation shows input construction, the right part – output construction, the succession of the arguments is preserved the same in the right part of the equation as in the left, \rightarrow shows the direction of derivation, S – subject, O – object, x – input argument or output argument preserving its status, y – output argument (if there is a change in its status), ‘ \emptyset ’ stands for deletion of the argument or a feature, that was given in the left part of an equation in this place. slashes enclose the part of the input construction preserved in the output, even if the functional status of the arguments changed, additional comments are given in round brackets.

¹⁸ The superevent controlled by subject S causes the subevent – the domain of the output object O (the argument demotion from the S of the non-causative input event)