Book of abstracts

Word Classes
Nature, typology, computational representations
Second TRIPLE International Conference
March 24-26, 2010

TRIPLE
Tavolo di Ricerca sulla Parola e il Lessico
Research Desk on Word and Lexicon
E-mail: triple@uniroma3.it
E-mail: http://host.uniroma3.it/laboratori/triple
## Contents

Luca Alfieri  
*Adjectival typology and modifier encoding strategy* [PRESENTATION]  
p. 5

Giorgio Francesco Arcodia  
*On some issues pertaining to the Chinese adjective* [POSTER]  
p. 5

Olga Boriskina  
*Noun cryptotypes as covert word classes* [POSTER]  
p. 6

Bruno Cartoni  
*Between adjectives and nouns: formalisation and use of relational adjectives* [PRESENTATION]  
p. 6

Denis Creissels  
*Between nouns and adjectives: the emergence of a new word class in Southern Bantu languages* [PRESENTATION]  
p. 7

Eva Dekany and Aniko Csirmaz  
*Hungarian classifiers* [PRESENTATION]  
p. 8

Annibale Elia, Mario Monteleone, Daniela Vellutino, Simona Sabatino and Federica Marano  
*Phrase and semantic structure for terminological multword automatic recognition* [POSTER]  
p. 8

Claudia Fabrizio  
*What's a meaning of a noun converted into verb? A semantic analysis of Italian denominal verbs from conversion* [PRESENTATION]  
p. 9

Livio Gaeta  
*Thoughts on decategorization* [PRESENTATION]  
p. 9

David Gil  
*Word classes in Roon* [PRESENTATION]  
p. 11

Maurizio Gnerre  
*Crossing boundaries from function to content words in Huave and other languages* [PRESENTATION]  
p. 11

Nicola Grandi, Malvina Nissim and Fabio Tamburini  
*Noun-clad adjectives. On the adjectival status of non head constituents of Italian attributive compounds* [PRESENTATION]  
p. 13

Maarten Janssen  
*Half an article* [PRESENTATION]  
p. 14

Elisabetta Jezek  
*Verbs of creation* [PRESENTATION]  
p. 14

Charitini Kallergi  
*Total reduplication in Modern Greek: word-class features and word-class change* [PRESENTATION]  
p. 16

Françoise Kerleroux  
*On the common inherent polysemy of event deverbal nouns and property deadjectival nouns* [PRESENTATION]  
p. 17

Efi Lamprou and Jean-Louis Vaxelaire  
*The subclass of proper nouns: how can we define it?* [POSTER]  
p. 17

Peter Lauwers  
*Between adjective and noun: category/function mismatch, constructional overrides and coercion* [PRESENTATION]  
p. 19

Christian Lehmann  
*The nature of parts of speech* [INVITED PRESENTATION]  
p. 19

Alessandro Lenci  
*Carving Italian verb classes from corpora* [INVITED PRESENTATION]  
p. 20

Beth Levin  
*What is the best grain-size for defining verb classes?* [INVITED PRESENTATION]  
p. 20
Martina Nied and Anna Pompei
Phrasal verbs, prefixes and word order [PRESENTATION] p. 21

Sebastian Nordhoff
Similarity thresholds and the number of word classes [PRESENTATION] p. 22

Anna Orlandini and Paolo Poccetti
Polifunzionalità e transcategorialità delle particelle di coordinazione in latino [POSTER] p. 22

Rossella Pannain and Anna Riccio
What word class do number words belong to? A syntax-semantics interface analysis [PRESENTATION] p. 23

Arantzazu Pascual Ortiz
Where have all the pro-complement verbs gone in Italian? [PRESENTATION] p. 23

Paolo Ramat
Reflections on word classes [PRESENTATION] p. 24

Irene Russo
Classifying adjectives: relational meanings, bifunctional lexical items and qualifying exploitations [POSTER] p. 25

Pawel Rutkowski
Demonstrative pronouns vs. personal pronouns in Polish Sign Language [PRESENTATION] p. 25

Galit W. Sassoon
A cognitive root for a word class distinction [PRESENTATION] p. 26

Raffaele Simone and Francesca Masini
Light nouns [PRESENTATION] p. 27

Anna Sores
On the borders of neglected classes: from preverbs to “satellites” via adverbs and particles [POSTER] p. 27

Julie Louise Steele
Tinkering with syntactic categories [PRESENTATION] p. 28

Giulia Venturi
Word classes and morphosyntactic annotation schemes. Shadows and lights in the existing part-of-speech tagsets for the Italian language [PRESENTATION] p. 29
Luca Alfieri

Adjectival typology and modifier encoding strategy [PRESENTATION]

Presented here is a cross-linguistic typology of the Modifier encoding strategy. It is based on a balanced variety sample of 82 languages and it also represents part of a larger project on the typology of part of speech systems.

It is widely agreed upon that languages differ both in the extent to which they structure the linguistic material by purely grammatical criteria and in the level at which they do so. We claim that the second parameter is the more fundamental in a word class typology. According to such perspective, we will not classify languages depending on the language-particular features they employ to establish their word classes, rather we will typologize according to the level (or layer) of linguistic complexity on which languages grammaticalize their word classes, and particularly their Modifier category.

Depending on the level at which Modifiers are encoded, we find:

• **Stem-lexeme-based Modifiers**, i.e. primary undervived adjectives, as in Latin, Dyirbal and Burushaski.
• **Derived-stem-based Modifiers**, i.e. no primary adjectives but secondary stems derived from verbal roots, as in Vedic, Yimas, Arabic and Babungo.
• **Phrase-based verb-prominent Modifiers**, i.e. no primary adjectives and verbal types settled in the Modifier slot, as in Lao, Kharia and Samoan.
• **Phrase-based noun-prominent Modifiers**, i.e. no primary adjectives and nominal types settled in the Modifier slot, as in Quechua, Pitjantjatjara and Hungarian.
• **Phrase-phrase-based Modifiers**, i.e. no primary adjectives and nominal Modifiers with a linker, as in Hausa and, to a lesser extent, in New Persian.
• **Clause-based Modifiers**, i.e. no primary adjectives and subordinated clauses employed in Modifier function, as in Koasati, Hua and Bororo.
• **Phrase-clause-based Modifiers**, i.e. no primary adjectives and relative participles employed as Modifiers, as in Garo, Urarina and Sumerian.
• **Superstem-based Modifiers**, i.e. no primary adjectives and incorporated verbal roots in Modifier function, as in Chukchee and Nivkh.

The paper is aimed at demonstrating that the definitional features of the adjectival class may be largely, although not completely, viewed as a surface effect of the level of language complexity on which Modifiers are grammaticalized.

Giorgio Francesco Arcodia

On some issues pertaining to the Chinese adjective [POSTER]

It has often been claimed that Chinese does not have an independent class of adjectives, which are seen by some authors as a sub-class of verbs (cf. Li & Thompson 1981, Hengeveld 1992, Tang 1998, among others). According to Li (1990:177, fn. 2), Chinese adjectives are verbs as they can function as the predicate of a sentence without requiring any copula, as in the example below:

(1) 她 真 聰明
    tā zhēn cōngmíng
        she really clever
        ‘she is really clever’

However, the requirement of a copula for predication does not seem to be crucial for a cross-linguistically valid definition of adjectives (see Dixon 2004). In our talk we shall first provide evidence for the status of adjectives as an independent word-class in Mandarin Chinese, on the basis of the differences in behaviour remarked by Paul (2005 and forthcoming). Secondly, we shall show that the category of adjectives in Mandarin Chinese has a high degree of internal complexity. A long-recognised distinction in Chinese linguistics is that between predicative adjectives (“verb-like” adjectives, in Dixon’s (2004) terms) and non-predicative adjectives (non-verblike; cf. Lü & Rao 1981); the latter category includes those property-denoting words which can neither fill a nominal slot nor be stand-alone predicates but may act as modifiers of a noun, as 良性 liángxing ‘positive, benign’ (cf. Li Y. 1996). We shall also discuss the category of “non-attributive” or “predicative-only” adjectives (Deng, Wang & Li 1996), i.e. adjectives which can never be used as modifiers of a noun, but only as intransitive predicates. To conclude, we shall make some remarks on the boundary between adjectives and inchoative verbs, which is apparently a blurred one in Mandarin Chinese (Liu forthcoming).

Selected references


Li, Y. (1990), On V-V Compounds in Chinese, in “Natural Language and Linguistic Theory” 8, 177-207.

Li, Y. (1996), 非谓形容词的词类地位 (Lexical category of non-predicative adjectives), in “Zhongguo Yuwen” 1.

Liu, C.S.L. (forthcoming), The positive morpheme in Chinese and the adjectival structure, in “Lingua”.


---

Olga Boriskina

Noun cryptotypes as covert word classes [POSTER]

Grammatical research in pre-Corpus Linguistics was concerned chiefly with study of grammatical categories which bear morphological marking, i.e. outward representation in morphemes, called phenotypes in terminology of B.L. Whorf. (1956, p. 68-72). The field studies of covert word classes (cryptotypes) in world languages didn’t have a remarkable impact on Linguistic Theory. Corpus Linguistics has broadened the horizons of grammatical research to make it open to exploring, describing and comparing cryptotypes of world languages.

A noun cryptotype is defined as a covert class of nouns which is marked by lexical selection rather than morphemic tags. We approach identification of noun cryptotypes via verb capacity to classify (sort out) nouns and relate them to cryptotypes, as well as noun capacity to select verbs to co-occur with. In other words, verb syntactical valency and semantics are considered to be the key to noun cryptotypes identification. For example, the noun life is classified by the subject or object valency of the verb stems to flow, to pour, to ooze, to stream, to leak, to splash, to splatter, to sprinkle, to flood and thus related to the EL cryptotype “LIQUIDUS”. The conceptual abstraction which is represented by the noun life appears to be categorized in the English-speaking culture as a liquid with the features typical of the bodies of water. The corpus data show that life is also categorized in English as a living creature, who moves; as a human being, who gets things by hand, in order to possess and give away; as a human being, who can talk; and as an object, which is small enough to fit the hand. Accordingly, apart from the cryptotype “LIQUIDUS” the noun life is related to such cryptotypes as “MOVENS”, “HOMO TENENS”, “HOMO LOQUENS”, “RES PARVA”. In a word, the noun is distributed among five cryptotypes. The fact that the noun is attributed to more than one cryptotype can be accounted for by the Law of Mythological thinking ‘Tertium datum’, which permitted multiple categorization of an entity. The presentation provides the criteria for the distribution of abstract nouns in six cryptotypes of the English language.

The Cryptotype classification of nouns however alternative it might seem is in fact within the mainstream linguistic framework of Conceptual Metaphor Modelling. Our approach refers to the ideas of conceptual metaphor theory (Lakoff and Johnson, 1980, Lakoff 1987, Johnson, 1987), family resemblance theory (Wittgenstein, 1953), primary metaphors theory (Grady, 1997) and central mappings (Kövecses, 2002, 2005). On the whole, the Method of Nomina Abstracta Cryptotype Distribution is the continuation of the theory of Metaphoric Mapping by reawakening Whorf’s approach to folk categories.

The approach discussed in the paper is targeted at a better understanding of metaphor in Modern languages on the one hand: the qualitative analysis of a noun’s cryptotype distribution indicates the hierarchy of metaphors or image schemas the noun represents and their relative value to the English-speaking culture. On the other, we believe that the description of cryptotype modelling of abstract nouns can contribute to computational modelling of their combinability.

The results of the research are presented in diagrams and illustrated, the implications of the findings are discussed.

---

Bruno Cartoni

Between adjectives and nouns: formalisation and use of relational adjectives [PRESENTATION]

Relational adjectives, i.e. adjectives that are derived from nouns (like presidential from president), take part of numerous phenomena that need to be addressed in Natural Language Processing (NLP). However, they are hardly explicitly recorded as such in NLP resources.

These adjectives are closely related to the noun they are derived from. They have special syntactic characteristics (they can not be used in a predicative position – *this election is presidential – and they are not gradable – *very presidential) (see (Mélis-Puchulu 1991; Daille 1999; Fradin 2008) for a complete description). Moreover, in a Noun-Phrase (NP), they set up a semantic relationship between the head noun of the NP and the base noun of the adjective. When prefixed, the prefixation process applies on the semantic of the base noun (e.g. anti-constitutional means something that is against the constitution). In some case, the prefixation can be applied indistinctly on the base noun or on the relational adjective (like in French traitement anticancer or traitement anti-cancéreux).

In a formalisation perspective, and following a lexicam approach to morphology (Fradin 2003), the coinage of
relational adjective can be formalised as followed:

\[
\begin{array}{c|c|c}
\text{INPUT} & \text{OUTPUT} \\
(G) & X & XsEx \\
(F) & X/ & /XsEx/ \\
(SX) & \text{cat:an} & \text{cat:a_rel} \\
(S) & X^* & X^* \\
\end{array}
\]

In this rule, the output (the relational adjective) keeps the semantic representation \(X^*\) of the base noun, and only the formal part of the constructed lexeme shows formal feature (\(sfx = \text{specific suffix of relational adjective}\)).

Relational adjectives are frequent in many languages (esp. romance), but their coinage is not necessarily parallel in all the languages. Indeed, a short study on bilingual French/Italian Dictionary Garzanti (2006) shows that more than 1,000 relational adjectives do not have a French equivalent and should be translated with a NP like in \textit{credizio} = \textit{du crédit}.

Many NLP applications are confronted to relational adjectives and could take advantage of a specific representation. For example, in the analysis and the recognition of terminological variants, it could be very useful to make a link between the term \textit{thérapie comportemental} and \textit{thérapie du comportement}. This link can be automatically made if the lexical processing step ensures the recognition of \textit{comportemental} being the relational adjective of \textit{comportement}.

In a multilingual perspective, it would also be very useful to analyse the Italian prefixed lexeme \textit{pre-congressuale} as being a prefixation of the base noun \textit{congresso} [eng: congress], in order to match the French equivalent \textit{pré-congrès}, coined on a noun, because \textit{congrès} does not have a relational adjective in the French morphological system.

In this submission, we would like to claim that relational adjectives are specific adjectives that need to be recorded as such in computer lexicon, and that many applications could benefit of them. Beside the fact that they are not qualifying adjectives, the most important information that needs to be recorded is the base noun they are coined on. Acquisition of such information requires a morphological analysis of all the relational adjective of a language. This can be done with morphological analysers such as Derif (Namer 2009), that can rapidly acquire the requested information. Implementation of such lexical information and its extensive uses in terminological variation identification are currently under process and complete results will be presented. Similar experiments on deverbal adjectives are also underway.

**References**


---

**Denis Creissels**

\textit{Between nouns and adjectives: the emergence of a new word class in Southern Bantu languages} [PRESENTATION]

In Bantu languages, a class of adjectives can generally be delimited on the basis of the following two criteria: adjectives show the same noun class prefixes as nouns, but noun stems are compatible with a limited number of class prefixes, and this constitutes a lexical property, whereas adjectives have forms for all noun classes, and their prefix is determined by agreement rules. However, Southern Bantu languages also have a type of noun modifiers, traditionally called “relatives”, the status of which needs for clarification.

So-called relatives have semantic properties quite comparable to those of adjectives, and are introduced by the same attributive linker, but their agreement morphology is different, as illustrated by ex. (1) and (2) from Tswana.

\[(1)\quad \text{mo-nna} \quad \text{jo} \quad \text{mo-lele} \quad \text{‘a tall man’} \quad (-\text{lele} \text{is an adjective stem})
\]

\[
\begin{array}{c|c|c|c}
\text{CL1-man} & \text{CL1.ATTR} & \text{CL1-tall} \\\n\end{array}
\]

\[(2)\quad \text{mo-nna} \quad \text{jo} \quad \text{o-bothale} \quad \text{‘a clever man’} \quad (-\text{bothale} \text{is an ‘relative’ stem})
\]

\[
\begin{array}{c|c|c|c}
\text{CL1-man} & \text{CL1.ATTR} & \text{CL1-clever} \\
\end{array}
\]

In the South African bantuist tradition illustrated by the classic grammars of Zulu (Doke), Southern Sotho (Doke & Mofokeng) and Tswana (Cole), noun dependents are classified according to the set of agreement prefixes they select. This leads to grouping together ‘relative’ stems like \textit{bothale}, which otherwise have nothing in common with verbs, and relativized verb forms – ex. (3).
This classification does not distinguish between word level and phrase level, leaving entirely open the question of the precise status of the ‘relatives’ that are not relative clauses. I will argue that they constitute an emerging class of adjectives originating from a predicative use of nouns in which predicate nouns do not identify an entity, but rather qualify it. For example, the stem of the relative modifier o-bothale (ex. (2)) is also an abstract noun (bothale ‘cleverness’), and the stem of the relative modifier a-molelo ‘hot’, as in metsi a amolelo ‘warm water’, is also the concrete noun molelo ‘fire’.

On the basis of such examples, an alternative analysis could be considered, according to which a subclass of nouns would have the additional property of occurring in predicative and attributive constructions expressing qualification, with an agreement prefix added to their inherent prefix (the class 14 prefix in bo-bothale, the class 3 prefix in mo-lelo).

However, in a synchronic perspective, this analysis is contradicted by the fact that some relative stems are formally distinct from their cognate noun(s), for example -boi ‘fearful’ and bo-boi ‘fear’: it is impossible to use *boi as a noun, or *-boboi as the stem of a ‘relative’ modifier.

Eva Dekany and Aniko Csirmaz
Hungarian classifiers [PRESENTATION]

Numerals classifiers figure prominently in Hungarian but descriptive work on them is close to nil and no formal analysis exists either. Beckwith (2007) identifies 2 types of Cls in the language (sortal and group) but does not describe their distribution. This talk is the first step in filling in this gap. We show that there are many more Cls in Hungarian than identified by Beckwith and present criteria for delineating the 2 classes from each other. We also examine how they relate to measure-terms (5 boxes of candies).

Beckwith recognizes 7 sortal Cls (2 seim cukor ‘2 CL candy’ 2 candies) and 6 group Cls (2 falka kutya ‘2 CL dog’ 2 packs of dogs). We have identified 12 more sortal and 21 more group Cls. We present the following distributional tests to differentiate between the two types of Cls.

- adjectives predcing the CL modify the CL itself with group Cls but the noun with sortal Cls (3 nagy kötég újság ‘3 big bundle newspaper’ 3 big bundles of newspapers vs. 3 nagy seim dió ‘3 big CL nut’ 3 big nuts)
- group Cls can be both preceded and followed by the same adjective or adjectives with contradictory meaning, but sortal Cls cannot (2 big packs of big/small dogs vs. *2 big sortal-CL big/small candies)
- deleting the noun from the template these are 2 CL Ns preserves the meaning of the phrase only with group Cls

These tests also reveal that there are striking typological paralles between Cls in Hungarian and Asian Cl-languages, eg. Mandarin (cf. Zhang 2009).

Measure-terms exist in all languages. In Hungarian they share some properties with both types of Cls but they form a category of their own. They pattern like groups Cls for adjectival modification and like sortal Cls for Ndeletion in the template these are 2 CL Ns. They differ from both classes in that they cannot form compounds of the form CL-N with the noun they modify.

An important disagreement between typologists is whether sortal Cls are identical to measure-terms and they form one category or not (Senft 2000, Aikhenvald 2000). Arguments for treating them differently come mainly from semantics. Our results show that in Hungarian there is syntactic evidence against conflating these categories into one.

References
Zhang, N. (2009), Syntactic properties of numeral classifiers in Mandarin, Ms, National Chung Cheng Universtiy.

Annibale Elia, Mario Monteleone, Daniela Vellutino, Simona Sabatino and Federica Marano
Phrase and semantic structure for terminological multiword automatic recognition [POSTER]

By studying word combination typology description, this study focuses on the processes of multiword creation, mainly of those multiwords which have been lemmatised in a specialised domain electronic dictionary called Dizionario Informazione Europea (DIE). This dictionary includes 3126 frozen complex lexical units which have been retrieved from community institutional glossaries. It is a linguistic resource functional to the training of natural language processing software. Phrase and combinatory properties of multiword internal structures have been determined on syntactic basis. Different combination typologies of simple words into complex lexical units have been observed, i.e.: bi-gram, tri-gram, fourth-gram, fifth-gram, and so on, up to more extended structures, composed by more than ten
elements. Such typologies have been determined formalising the constituents of each morphosyntactic internal structure. DIE formalism also includes references to word external structures, insofar as it copes with synonymic, hiponymic, and hyperonymic lexical variants, with orthographical variants and with variants depending on sociolinguistic features, which are useful to the completion of terminological simplification processes.

References
De Bueris G. et al. (2008), Le polirematiche dell'italiano, in De Bueris G. & Elia A. (eds.), Lessici elettronici e descrizioni lessicali, sintattiche, morfologiche ed ortografiche, Salerno, Plectica.
Vellutino D. (2009), La comunicazione pubblica per la promozione delle pari opportunità, in D’Antonio, V.& S. Vigliar (eds.), Studi di Diritto della Comunicazione. Persone, Società e Tecnologie dell’Informazione, Padova, CEDAM.
Vietri S. (2008), Dizionari elettronici e grammatiche a stati finiti, Salerno, Plectica.

Claudia Fabrizio
What’s a meaning of a noun converted into verb? A semantic analysis of Italian denominal verbs from conversion [PRESENTATION]

The aim of my contribution is to investigate the subclass of Italian denominal verbs coming from a morphological process of conversion. These verbs represent an interesting case-study inside the continuum between Noun and Verb, especially because of their apparently unpredictable semantics. As a matter of fact, in spite of the morphological identity of the output (lacking in any Wortbildungsbereichung of derivational suffixes), figliare ‘to pup’, astrologare ‘to behave as an astrologer’, salmonare ‘to give something the colouring of a salmon’, uccellare ‘to go fowling’, show radically different meanings. No morphological clue guides the interpretation of these items, which appear morphosemantically opaque.

My hypothesis is that verbalizations obtained with a conversion trigger a detaching of nominal coefficients on the verb (specular to the detaching of verbal coefficients on the noun, cfr. Simone 2006); therefore, aspects of the meaning of the noun detached on verb can guide the semantic interpretation of the resulting verbal lexeme. The theoretical frame I adopt here is the theory of Qualia Structure (Pustejovsky 1995). According to this pattern of analysis, I assume that the semantics of denominial verbs correlate especially with the kind of quale (formal, telic, costitutive, agentive) activated during the process of conversion. For instance, in figliare the agentive quale is activated; in astrologare the telic one; in salmonare the costitutive one; in uccellare, at last, the formal one. The detaching of a certain quale could depend on the semantic types (natural, complex, mixed) nominal lexemes belong to (Pustejovsky 2001).

Furthermore, possible correlations between Qualia Structure and syntactic behaviour of the verb (transitivity, argumental structure, distribution of Qualia among internal and external arguments) will be taken into account.

The treatment of the more complexes cases will lastly give a hint to focus on the tricky relationship between lexicon, derivational semantics and encyclopaedia.

References

Livio Gaeta
Thoughts on decategorization [PRESENTATION]

While it is not yet well understood whether word classes should be considered a universal property of languages – in the sense of an absolute language universal like for instance the occurrence of vowels – in our lives as theoretical and descriptive linguists we are normally confronted with a plenty of examples in which word class properties are changed, namely with recategorization. In a way, this might be taken as a good news for the issue of word classes, because, if something gets recategorized, it must have belonged to a category before. However, there are several ways of
understanding recategorization. Here, I will accurately distinguish between decategorization and recategorization (cf. Malchukov 2004), because one may also expect that losing certain word class properties does not necessarily amount to acquiring a new set. The paper will focus on this latter aspect, namely with decategorization, which will be mainly discussed on the basis of data from German.

Clearly, decategorization very much depends on our universalistic stance towards word classes: if we assume the latter to be universal, then decategorization will be massively present in typically isolating languages like Chinese, because words may commonly occupy different syntactic positions independent of their morphological properties, which are quite poor, if not absent. For this reason, I will rather follow a different approach (cf. Vogel 1996, 2005), which keeps the syntactic dimension (for instance being a predicate or an argument) distinct from the morphological properties of a word (basically, its inflectional behavior). While the first dimension can aspire to being universal, the second one surely is much more language-specific. In fact, inflectional behavior not only refers to the lexical properties of inflecting according to a certain morphological pattern. Much more, inflectional behavior characteristically fixes the word class of a certain word. Thus, for instance German nouns inflect differently from verbs, and this has consequences on their syntactic behavior. Accordingly, something must happen to a German noun in order to get used as a verb: typically, it must undergo any kind of affixation (we can mao-isieren a society or a party, or we can ver-merkel-n it), or simply conversion, as in öl-en ‘to oil’. The latter can be much more difficult to identify, because no affixal modifications take place: however, we can rely again on inflectional behavior to solve the question, which confirms the idea that inflectional behavior actually fixes the word class of a certain word. Notice that such morphological processes usually imply a recategorization.

With these premises, we can now face the question of decategorization, namely of those cases where we observe the loss of word class properties without having at the same time a recategorization, i.e. the attribution of new word class properties. If we understand word classes in prototypical terms, then it is clear that decategorization will be found at the borders.

A first type is given by classical instances of incorporation in which nouns lose their word class properties being inserted into wider units headed by verbs. In German, such incorporation-like decategorization usually results from the reverbalization of compounds (cf. Wurzel 1998):

1. a. reverbalization of complex nominalized infinitives: das Bau-sparen ‘lit. the building-savings’ > bau-sparen ‘to save through a building and loan association’;
   b. reverbalization of complex deverbalized perfect participles: Platin-beschichtet ‘lit. platin-coated’ > platinbeschichtet ‘to platin-coat’;
   c. reverbalization of action nouns: Notlandung ‘emergency landing’ > notlanden ‘to have an emergency landing’;
   d. reverbalization of agent nouns: Testfahrer ‘test operator’ > testfahren ‘lit. to test-operate’.

2. Empfehlenswert ist, dass man den Scooter vorher kurz testfährt / *Teste fährt
   ‘It is advisable to shortly test the scooter before’.

A second type is more complex and has been referred to as conversion, but in a looser way. We can distinguish two cases: A first case is given by nouns used as adjectives for color terms like flieder ‘lilac’, sand ‘sand’, etc., where no inflectional behavior is acquired. Thus, we simply have decategorization. However, the syntactic dimension comes into play which forces these units into syntactic patterns: thus, a syntactic recategorization takes place, even though it basically leaves the words as such decategorized. Notice that while cases like flieder are sporadic, the syntactic recategorization can be, and usually is, quite productive as shown by nominalized infinitives and adjectives: das Streben ‘the ambition’, das Gesehen-werden ‘lit. the being seen’, das Schöne ‘the beauty’, etc. This has been often referred to as syntactic conversion, distinct from morphological conversion. In our terms, we can say that we have a syntactic recategorization in the absence of a morphological recategorization.

A second case is given by instances where decategorization results from a different syntactic process, namely (ascriptive) predicate formation:

3. a. Der Typ ist echt Banane. ‘The guy is really weird (lit. banana)’
   b. Eure Musik ist echt Hammer. ‘Your music is really great (lit. hammer)’

This also leads to decategorization in that the noun loses its inflectional possibilities (cf. Die Typen sind echt Banane / *Bananen). Notice that this happens to be the case not only for a relatively small number of nouns employed in such a predicative function, but for the complete set of adjectives, whose inflectional behavior is suspended when used predicatively. We can say that they are strongly decategorized. In contrast with other languages like English or Italian decategorization affects German adjectives massively, in that the boundary with adverbs is blurred.

Finally, we have a third type which results from the combination of the former two, incorporation and predicate formation:

4. a. Diese Frau ist Klasse. ‘This woman is wonderful (lit. class)’
   b. Der Bassist spielt echt Klasse. ‘The bassist plays really great (lit. class)’.
The decategorized noun can be not only employed in predicate formation, but also as incorporated noun (4b). This multifunctionality characterizes typical isolating languages like Chinese. In the light of the distinction made above between morphological properties and syntactic dimension, this does not come out as a surprise, but rather confirms the necessity of keeping the two levels distinct. In its essence, decategorization, in the absence of a morphological recategorization, turns out to enhance multifunctionality, which is typical for the isolating language type. Notice that the latter may also happen as a consequence of processes of incorporation via the reversalization of a compound (cf. Klassenspiel > klasse spielen), which normally characterize the polysynthetic type. Thus, an isolating trait turns out to be enhanced by a polysynthetic process. These different sources and their different outcomes make the classical distinction of language types in isolating, agglutinating, fusional and polysynthetic languages debatable and in search of a finer characterization.

References

David Gil

Word classes in Roon [PRESENTATION]

This paper presents an exploration of word classes in Roon, a hitherto undescribed language belonging to the South-Halmahera-West-New Guinea subgroup of Austronesian, spoken by some 1100 people on the eponymous island just off the Wandamen peninsula, in the Cenderawasih bay in West Papua. Its focus is on patterns of grammatical agreement and what such patterns reveal about grammatical categories of various kinds.

Agreement markers in Roon are based on the independent pronouns, but occur as prefixes attached to their hosts, the targets of agreement. Agreement markers bear features of person (1st, 2nd, 3rd, further distinguishing in the 1st non-singular between inclusive and exclusive), number (singular, dual and plural), and, within 3rd person, gender (animate, inanimate). The primary syntactic domain of agreement is clausal, between the subject as controller and associated activity expression as target.

However, agreement also occurs phrase-internally. In particular, the definite marker -ya also agrees with its host, displaying an agreement paradigm that is completely identical to that of words denoting activities, with the same grammatical features expressed with the same forms. Morphologically, the definite article -ya ‘the’ thus seems to be a regular verb. Syntactically, however, its behaviour differs in important respects from that of other activity-denoting expressions. First, whereas activity-denoting expressions such as i-farar ’3SG:ANIM-run’ have the properties of independent words, forms such as =i-ya ’3SG:ANIM-the’ are not independent words but rather enclitics. Secondly, whereas forms such as i-farar may constitute complete sentences, either on their own or with construction with an argument, =i-ya can not: thus romau=ii-ya ’child=3SG:ANIM-the’ cannot occur as a complete sentence with a meaning such as ‘The child is definite’.

Agreement in Roon thus highlights the importance of distinguishing between categories of different types: morphological, syntactic, and semantic. This paper presents a preliminary inventory of the morphological, syntactic and semantic word classes of Roon, showing how they cross-cut and overlap in ways quite different from those familiar from other languages. Morphologically, the main distinction is between forms that host the agreement markers (eg. i-farar ‘run’ and i-ya ‘the’) and words that do not (eg. Amos-i ‘Amos’ and ba EXCL). Syntactically, the primary distinction is between expressions that may stand alone as complete non-elliptic sentences (eg. i-farar ‘run’ and ba EXCL) and expressions that may not (eg. i-ya ‘the’ and Amos-i ‘Amos’). Thus agreement in Roon shows, in a rather novel way, how morphological and syntactic categories are logically independent of each other, as well as of various semantic categories such as thing, property, activity and so forth.

Maurizio Gnerre

Crossing boundaries from function to content words in Huave and other languages [PRESENTATION]

In many languages (excluding those with a highly synthetic typology) two macro-classes of words are found: content (or ‘full’) words, and function words. The first includes sub-classes such as verbs, nouns, adjectives, or adverbs; the second, independent pronouns, demonstratives, adpositions, marking certain syntactic relations, and others possible
sub-classes. Languages differ among them more in the distinctions they recognize inside the function macro-class than in the content one. The two macro-classes are frequently characterized also as open and closed sets, respectively. Beyond the semantic-functional, and the open/close criteria used to define them, a third criterion is grounded on the “internal transitions” between sub-classes. In fact, inside the first macro-class transitions among its sub-classes are frequently found, to the point that in many languages clear-cut distinctions among verbs/nouns/adjectives seem to be rather artificially casted. Inside the second macro-class, on the opposite, sub-classes can be usually recognized as self-contained, with much rarer, or nonexistent, transitions among them.

When we watch at the boundaries between the two macro-classes, we find that in several languages these are, or have been, crossed in different ways, with shifts from one to the other. These shifts are sometimes unevenly distributed: while many shifts from content to function words have been described as examples of grammaticalization (e.g., nouns and verbs grammaticalized as adpositions), opposite shifts, from function to content words are, at least apparently, less frequent, or have been observed more rarerly. For these cases we are entitled to use, most of the times, the complementary concept of ‘lexicalization’, or ‘de-grammaticalization’ (Heine & Kuteva 2007). To safely show the first type of shift, diachronic or comparative data are usually needed. On the other hand, to show lexicalizations (or de-grammaticalizations) of the second type of shift, synchronic observations alone could provide useful insights.

In ombeuiyüts (San Mateo Huave, one of the four varieties of a genetically isolated language, spoken in four varieties by 12,000 persons in Southern Mexico), the macro-class of function words is relatively small not only, at it should be expected, when compared to the other macro-class of the same language, but also when compared to the same classes of many other languages. In ombeuiyüts, this macro-class includes a rather reduced sub-class of adpositions, a typological Mesoamerican areal feature. Although ‘internal transitions’ among function sub-classes are cross-linguistically not so frequent, in Huave we find at least some examples of ongoing grammaticalizations, involving the demonstrative aag ‘this’, at the base of three ‘concomitance’ adpositions: n-aag ‘instrumental’, an-aag and aw-e-aag, both (Singular) comitative (Cuturi & Gnerre 2005). These are bi- (or tri-?) morphemic adpositions.

Two other adpositions, synchronically underrived, and very frequent in spoken language, are mono-morphemic wix and ti-, both carrying spatio-temporal locative meanings: ‘on, over, at’, and ‘in, inside, to’, respectively (examples in (2), (3), and (6)). Ti- frequently shows up in a derived form ti-ul (< ti-al where -al is a ‘stative’ (Stat.) morpheme). So, the five Huave adpositions mark predicative relations (in the sense of VanValin 2005) i.e.: spatio-temporal locatives, comitative and instrumental ones. Non-predicative relations are codified as Arguments of verb forms, without any explicit adpositional marking. Ti- is implemented also in a set of locative constructions preceding names of body parts: ex. ti-pech (< o-pech ‘back, shoulders’), ‘in/on the back, behind’ (but also expanded in ti-xi-pech ‘behind me’, where – xi- ‘my’). Both wix and ti- can be used in adposition stranding constructions, as in 1): (all examples are from naturally-spoken language; relevant forms underlined):

(1) xe-müm a-püüch mats sa-na-meay wix
    Poss.1.mother Atmp.3.-open folding bed Fut.1Sg.-sleep Adp.
    ‘My mother opens the folding bed for me to sleep on it’

The two adpositions, are also implemented in ‘phrasal verbs’, contributing to produce a wide set of idiomatic lexical units: in (2) the modal ndaj ‘(to be) possible’ followed by wix conveys an adverbal/adjectival meaning of ‘last’ (usually with a temporal value):

(2) oxej ndaj wix sa-mb ti-ul plas
    tomorrow possible Adp. Fut.1Sg.go Adp-Stat. market
    ‘Tomorrow is the last day I will go to market’

The same two adpositions, however, are the base for many (at least twenty) derived forms, which are lexicalized (or de-grammaticalized), becoming content words, crossing in this way the boundary between the two macro-classes. Through prefixes, suffixations, and reduplications, they assume not only several adverbial meanings, but also some adjectival-adverbial ones. Here I provide examples with only some glosses: al-ea-wix ‘equally, equal’, ka-wix ‘up’, al-ka-wix ‘it is up’, wix-an, wix-ay ‘exactly at the same time’, wix-ay-an ‘fine, right’, wix-wix-an ‘at every moment’, al-ti-l-ül, si-tiül/si-tel ‘in case…’, ti-ul-an, ti-l-l-ül, ti-l-ül-an, ti-ul-ti-ul-an ‘always at the same place’. Another very productive derivation from ti-, ti-ul, basically ‘down’ (<ti-ul, lit. ‘in ground’), not an adposition, is also expanded in at least six forms. The prefixes and suffixes most frequently implemented in all these derivations are: al- (Stat.), ka- ‘locative’ (fossilized), and two intensifiers (Int.): -Vy and –an, the first used mostly in verb forms, the second in both verb and noun forms. Some examples of these forms are provided in (3)-(7):

(3) wix-ay ta-xom-ay-ej ijpw-ül-yiy-ej wix a win
    ‘The two of them met where the turtle was, exactly at the same time’

(4) wix-wix-an lea-aw xi-meats xa-noj
    Adp.Temp.-Int. remember-Pas.1Sg. Pos.1Sg.-heart Pos.1Sg.-husband
    ‘I worried at every moment for my husband’
(5) ti-ül-til-ül-an a-omb
Adp.Adp.-Int. Atmp.3-lay egg Det. turtle where Atmp.3-lay egg
‘The turtle lays its eggs always in the same place (one on top of the other)’

(6) ta-mbaag... alax-wix a chaw ti-ül xax teat Eucario
‘Mr. Eucario distributed the atole in the cups in equal shares (among everybody)’

(7) wix-áy-an axix? Wix-áy-an!
Adp.-Int.-Int. salty? Adp.-Int.-Int!
‘Has this been salted to the right point? It is fine!’

In (3)-(5) we find derivations with adverbial meanings, while in (6)-(7) we find adjectival-adverbial meanings: ‘equally’ and ‘right, fine’, respectively. These derivations can trigger further reflections on the lexicalizations and grammaticalization of adpositions. In first place they contradict well-known generalizations such as “adpositions do not take part in derivational morphology, as either inputs to or outputs of word formation rules.” (Baker 2003: 305). This observation is not valid, however, only for Huave. In much more familiar, ancient and modern languages, such as Latin and English, we find adpositions that through suffixations become not only adverbs, but also adjectives, and even nouns.

Examples from Latin are: interior, interim < inter; superior < super; exterior < extra < ex.
Examples from English are: inner/ inmost/ innermost; outer/ outermost/ outing; upper/ uppermost.

We could observe that in these examples, most of the implemented suffixes share basic intensifying meanings (but not so for Lat. interim, Eng. outer, in the boxing jargon carrying a nominal meaning, and outing), as in Huave -Vy and –an. These data open unexplored perspectives on some aspects of the ways in which boundaries between function and content words are crossed, and, ultimately, on the lexical-grammatical continuum. In fact, the relative weakness of boundaries between the two macro-classes is strictly related to the never-ending negotiation between lexical and grammatical dimensions of language.

References

Nicola Grandi, Malvina Nissim and Fabio Tamburini
Noun-clad adjectives. On the adjectival status of non head constituents of Italian attributive compounds [PRESENTATION]

Typologically, word classes are described as prototypes determined by the simultaneous occurrence of three semantic classes (objects, properties, actions) and three pragmatic functions (reference, modification, predication), which are cross-linguistically universal. However, the constructions realizing them are language-specific. In this picture, every linguistic item/construction that performs a modification by property can be labelled as an adjective/adjectival. In alleged Italian NN compounds such as as riunione fiume or parola chiave, the second constituent (N2) emphasises a property of the head (N1) thus performing a typical adjectival function. Therefore, a noun encoding the pragmatic function of modification by property can be expected to exhibit shifts towards the class of unmarked adjectives, both formally (displaying a morpho-syntactic behaviour typically associated with adjectives, e.g. agreement with N1), and semantically (expressing features normally carried by adjectives that are frequently associated with the head nouns; cf. riunione fiume / riunione lunga).

In this paper we report preliminary results of a corpus study aimed at testing whether such properties indeed appear to surface in occurring data, in support of such typological claims. We selected some nouns that typically fill the N2 slot in appositive compounds of the kind described above, and observed their behaviour when featuring in meaningful constructions in two corpora of Italian: a balanced one (CORIS, 120-million words), used for the investigation of semantic features, and a larger but less controlled one (ItWac, web-derived, 2-billion words), which we exploited for finding possibly unexpected behaviour from a morphosyntactic perspective.

As for semantic features, for each selected N2 (e.g. fiume), we exploited corpus evidence to obtain the sense features carried by N2 (such as fiume ~ lungo). Specifically, for each N2:

• we extracted all N1 N2 constructions
• we singled out all nouns filling the N1 slot
• we extracted all adjectives co-occurring with N1s in a variety of constructions
We then assumed that the set \( A = \{ \text{ADJ}1, \ldots, \text{ADJ}n \} \) formed by those adjectives occurring most frequently with most N1s would be indicative of the implied, used sense of N2. At a second stage we checked whether uses of “N1 N2” and uses of “N1 ADJ1” (where ADJ1 \( \in A \) are indeed comparable from several perspectives (e.g. frequency).

On the formal level, we tested whether some of the typical morphosyntactic features of adjectives are associated with N2 (agreement in number with N1, any sort of adverbial modification, and superlative form). Our preliminary findings show that not all features are equally exhibited, thus indicating that (i) although the function performed by N2 is modification, only some of the expected shifts are actually observed, and that (ii) there are clear degrees of adjectivehood/nounhood, so that some features are acquired/lost with much more ease/difficulty. In the presentation we will discuss the details of the corpus study and the resulting observations.

Maarten Janssen
Half an article [PRESENTATION]

Word classes do reflect the combinatorial behaviour of words: *like* is a verb and not a noun because it cannot occur in direct object position. And subclasses in a sense really provide a motivation for grammaticality judgments: *My car is *greener than yours* is wrong because *green* (as opposed to *greener*) is not a comparative adjective.

But there are many words for which the assignment of a differentiating POS tag is problematic. In this article, we provide an in-depth analysis of the English word *half* as a “measure out phrase” in sentences like:

(1) I ate half of the apple

Most dictionaries classify *half* as a noun, but it is at best not a typical noun; it behaves more like quantifiers such as *all* and *some*. But it is different from those in that you cannot say *I ate all/some/*half apples*, whereas you can use it in a number of constructions in which you cannot use most other quantifiers: *half the apple*, *half eat an apple*, *a half red apple*, *half an apple*, *half these apples*. In fact, there is no other word in the English languages that behaves exactly alike. And there are many other words that behave in a similarly unique way.

If word classes are to be predictive, then *half* would have to be assigned a class of its own, which is a highly inelegant solution. And we show that it is also not viable to assign the difference in combinatorial behaviour between *half* and words like *some* to other factors such as semantic blocking or grammatical features.

Our claim is that the problem arises from the simple fact that words like *half* do not actually *have* a word-class. Rather, the combinatorial behaviour is encoded in the lexical element of the word itself. The reason why *half* behaves uniquely is because it happens to have a unique set of lexically assigned grammatical patterns assigned to it.

This solution lies in the line of the recent work by Hanks & Pustejovsky (2006) who argue in favour of a pattern dictionary. But where the pattern dictionary is motivated by a desire to capture the semantic difference between the interpretation *shoot a person* and *shoot a bullet* in terms of patterns rather than word-senses of the verb *shoot*, we argue that a similar approach is even more fitted to account for words that do not properly belong to any particular word class.

It should be noted that especially for nouns and adjective, the distributional behaviour appears *not* to be lexically determined, but rather be a proper word-class determined, despite all the problems related to their exact delimitations. Therefore, we argue that it is not so much the case that there are no word classes, but rather that not all words belong to a word class, or at least not in a very significant sense of the word.

Elisabetta Jezek
Verbs of creation [PRESENTATION]

Verbs of creation come in two main sorts: those that denote the coming into being of the referent of their direct internal argument as a result of the event in question (e.g. build) and those that denote the coming into being of an entity that does not surface into an argument to the verb (e.g. translate).

Canonically, only verbs belonging to the first group are considered creation verbs. However, this is a restrictive view that fails to capture a number of interesting generalizations about the way creation events are encoded in language, for instance the fact that both the build-verbs and the translate-verbs may exhibit result nominalizations which are able to refer to the created entity (e.g. building, translation). Examples of creation verbs of both groups are in (1):

(1) verbs of creation:
   a. verbs denoting the coming into being of the referent of their direct internal argument:
      built a house, write a book, compile a list, dig a hole, cut a slice, bake a cake
   b. verbs denoting the coming into being of an entity that does not surface as an argument:
      translate a book, paint the landscape, photograph a sculpture

The two classes in (1) may be further subdivided (cf. Bisetto & Melloni 2007 a.o.). For instance, Piñón 2008 claims that verbs in (1a) fall into three (semantic) subclasses, depending on the semantic character of their direct internal
argument: those denoting the creation of a physical object (2a), those denoting the creation of an event (2b) and those denoting the creation of an abstract entity (2c).

(2) (from Piñón 2008)
   a. Sarah painted a picture of the Hungarian parliament building
   b. Sarah sang a sad song
   c. Rebecca fabricated a story

Verbs of creation pose interesting challenges for verb aspectual classification. Though traditionally analyzed as accomplishments, they differ from “regular” accomplishments inasmuch as at the level of Event Structure, the subevent denoting the activity (e1) is co-extensive with the gradual coming into being of the effected object and the temporal extent of the event is delimited by the (state of) existence of the object. In this perspective, creation verbs have been analyzed as incremental theme verbs (cf. Dowty 1991).

The analysis proposed here aims at recasting the notion of incrementality associated to events of creation within a scalar approach, following a suggestion in Hale, Kennedy & Levin 1999 and Beaver 2008 a.o.

According to these scholars, the incrementality of creation events may be re-interpreted in scalar terms as the property of the effected object to go through gradual changes in its spatial extent during the execution sequences of the event. In this view, it is the gradable property spatial extent of the incremental theme, and not the incremental theme itself, that instantiates the homomorphism between the event and the object.

The analysis proposed attempts a formal modeling of the Event Structure of creation verbs taking as a starting point the enriched view of Event Structure proposed in Pustejovsky (2000), where it is assumed that creation predicates introduce the mode of opposition (Opposition Structure, OS) that the object undergoes through the event, namely between /the object not existing/ (e3) and /the object existing/ (e2).

![Fig. 1. Enriched ES of build](image)

While the enriched ES in Fig. 1 captures the relation between the event and the created object (by introducing reference to the ¬existence subevent, corresponding to e3 in the figure), it still does not capture the incrementality associated to creation verbs, which in our view can be conceived as a sequence of ordered changes along a scale that measures the gradual progression of states of existence of the argument’s referent throughout the event.

We will propose to model the incrementality associated to creation acts by integrating intermediate result states (sub-result states) in the enriched ES representation and in particular in the development of the process subevent e1 (see Caudal 1999 for a previous treatment of incrementality in a similar spirit – although in Caudal 1999 the notion of incrementality is opposed to the notion of scalarity, instead of integrated in it). Our proposal draws on Pustejovsky 2009, where Processes (such as e1 in Fig. 1) are analyzed dynamically as indefinite sequences of Transitions through States, with a change in value over some property of y in each State (see Naumann 2001 for a treatment of processes in a similar spirit).

Also, we will argue that a richer encoding of the end result state is needed to account for the different entailments associated with the result subevent (e2) of build-verbs and translate-verbs. In particular, we will propose that while “Mary built a house” introduces references to a single result-state predicate (e.g. exist (house)), “Mary translated the book into Italian” encodes two end states, instead of one: the state of the book being (fully) translated and the state of existence of the translation.

The overall aim of our presentation is to put forth a fine-grained representation of the event structure of creation verbs that is able to capture the different entailments associated to creation events, such as the existence of the object before the event is over (as disclosed by our ability to refer to created objects even during the early stages of the creation, as noted in Parsons 1990:174 ff.). Also, the analysis aims at examining the class of syntactically covert creation (translate-verbs) in more detail and at providing a preliminary taxonomy of the main verb classes falling into this group.

---

1 A different attempt to model the incrementality of creation acts starting from enriched event structures as in Fig. 1 can be found in Melloni & Jezek (2009), where the proposed formal solution is subevents co-identification (drawn from Levin Rappaport 1999).
References
Pustejovsky, J. (2009), Compositionality and the parameterization of meaning, Oral presentation at the 5th International Workshop on Generative Approaches to the Lexicon, Pisa, ILC-CNRS, Sept. 17-19.

Charitini Kallergi
Total reduplication in Modern Greek: word-class features and word-class change [PRESENTATION]

Total reduplication (TR, i.e. reduplication of whole words, such as Italian bello bello ‘very beautiful’ from bello ‘beautiful’, Dressler & Merlino 1994: 516) has been discussed as a possible language universal (Moravcsik 1978, Stolz 2004). This process has both lexical and grammatical functions, the latter mostly referring to its ability to change the class of the word on which it operates in some languages (see e.g. Dressler & Merlino 1994). Although the idea of word classes primarily refers to orthographic words, it is relevant to TR since the latter is often discussed as a word-formation process (Inkelas & Zoll 2005) and especially as a derivational one (Botha 1988). The paper explores the relation of TR to word classes, with special reference to Modern Greek (MG) and with respect to two specific issues: first, the distribution of the meanings of TR across word classes in MG and other languages, and a possible correlation between word class and reduplicative meaning; second, whether TR in MG changes the grammatical class of the words it applies to. Regarding the first issue, typological data lead to the observation that there are probably not any general, i.e. cross-linguistically valid, characteristics of word classes that restrict reduplicative meaning. Rather, it seems that particular lexical and morphosyntactic features of sub-classes within major classes, such as gradability and plural marking, are responsible for the rise of certain meanings; e.g., nouns marked for plural in MG express a special type of distributivity, that of scattered arrangement in space/time (topus topus place-PL.ACC place-PL.ACC= ‘in different places here and there’), which does not arise for reduplicative nouns in the singular. Concerning the second issue, it appears that in MG there is no absolute word-class change. However, the constituents of reduplicative expressions of certain types appear to lose some of their categorial features, e.g., in the V-V reduplicative construction with the interpretation of ‘long duration/continuation’, verbs are restricted in the form of second person singular of the imperative (pes pes say-2.SG.IMP say-2.SG.IMP= ‘by constantly saying’). As these are arguably on their way to grammaticalization (in the sense of Lehmann 2002), there may be a case for gradual class change.

References
Dressler, W. & L. Merlino-Barbaresi (1994), Morphopragmatics: diminutives and intensifiers in Italian, German, and other languages, Berlin, Mouton de Gruyter.
The **polysemy pattern** displayed by the deverbal nominals, or nomina actionis, is a well-studied phenomenon, investigated in Grimshaw (1990), Pustejovsky (1995), Jezek & Melloni (2009). So we know **how** these nominals are polysemic, as their distinct readings (event vs results objects) depend upon their construction (along definite article, argument-taking capacity, exclusive singular marking, aspectual adjectives). But we don’t know **why** these morphologically derived nominals are inherently polysemous.

Our aim is to explain:
- why do natural languages manage this maximal difference in ontological status to be borne by unique linguistic forms;
- why other deverbal nouns (denoting agent, instrument, localisation) are monosemic, yielding always an object denotation and never any event reading.

Our explication will rest upon 4 reasons:
1. Lexical categories (N, V, A) are combinations of values of two types: semantic denotational class and pragmatic discursive function. Nouns, for instance, combine the function of referring with the semantic class of denominations of objects, as their prototypical realization, instantiated by simplex lexemes.

<table>
<thead>
<tr>
<th>Semantic class</th>
<th>N</th>
<th>A</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic Function</td>
<td>Object</td>
<td>Property</td>
<td>Event</td>
</tr>
</tbody>
</table>

Table 1. Prototypical correlations defining lexical categories

2. Each simplex (or simpler) lexeme is supposed to be the base of a complex one, obtained by permutation of the pragmatic function and maintaining the semantic class:

<table>
<thead>
<tr>
<th>Base Lexemes</th>
<th>blanc</th>
<th>Voir</th>
<th>Main</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derived Lexemes</td>
<td>blancheur</td>
<td>vue, vision</td>
<td>manuel</td>
</tr>
<tr>
<td></td>
<td>Blanchir</td>
<td>Visible</td>
<td>manier, manipuler</td>
</tr>
</tbody>
</table>

Table 2. Function-changing Morphology

3. Hence deverbal nouns (event referring) and deadjectival nouns (property referring) are non prototypical, and as such, submitted to the attraction force of the major class of prototypical nouns. That’s the reason why N like chauffage or éclairage, saleté or grosseur, are polysemic, selecting an object reading: *Tu as fermé /porté à réparer le chauffage?; Les saletés sont toutes remontées à la surface; Ce chat a une grosseur dans le cou.*

4. Hence the unique direction of this semantic shifting: “If there is a semantic shift in zero coding of an occurrence of a word (i.e. flexibility) in a part-of-speech construction, even if it is sporadic and irregular, it is always towards the semantic class prototypically associated with propositional act function” (Croft 1991: 73).

The semantic flexibility of these two subclasses of Nouns depends on the prototypical character of categories, – which are the first thing we learn (at the elementary school) and the last one we understand, along Baker’s (2003) paradox.

**References**

---

**François Kerleroux**
*On the common inherent polysemy of event deverbal nouns and property deadjectival nouns* [PRESENTATION]

**Efi Lamprou** and **Jean-Louis Vaxelaire**
*The subclass of proper nouns: how can we define it?* [POSTER]
These differences are not surprising because, as far as the morphosyntax of proper names is concerned, the extent to which there are overt markers of namehood, and the nature of the markers, varies from language to language, so that simple formal definitions of names prove to be non-generalizable.

Two options are opened to us: first of all we can consider that the definitions of this class will be different from one language to another, which seems odd to us. Otherwise, we can try to find a definition which can be appropriate for every language; this can be difficult but rewarding.

Our main problem is that, from a multilingual point of view, some of the criteria used to define the class of proper nouns are simply irrelevant. Several authors consider that one distinguishes the proper name from the common noun thanks to the initial upper-case letter (Ullman 1964; Sag & Wasow 1999) and according to Frontier, “il suffit d’une majuscule pour transformer un nom commun en nom propre” (1997: 27). Even if this almost true for languages like English and French (the preceding words prove that it is not that easy: the capitalization of French does not transform the name of this language in a proper name), it certainly is not for languages that do not use a European alphabet.

Another criterion comes from the fact that many people think that proper names cannot be translated. The untranslatability is even the main criterion for Matczak (1991) because the majority of the proper names of his French corpus are not translated into Polish. The question of translation is more complex because one does not translate words but texts and a word that is not translated does not become automatically a proper noun.

It is a logical consequence of the concept of untranslatability to consider, in the manner of Alain Rey (1982) that proper names do not belong to any language. Baker (1993) adds that “proper names are words that are foreign to any language”. We will try to show that morphosyntactically each proper name belong to a specific language.

Another criterion is that proper names do not show any inflection: Sørensen (1958) for instance wrote that an *Alp is impossible but also that the Martins (the family) canno be a proper name because of its inflection. This conclusion is not intuitive and we will prove that in the Italian Mario / Maria, –o is the mark of the masculine and –a of the feminine.

One of the most widespread assumptions concerning proper names is that they are Names without determiner (cf. Chomsky, 1965). According to Bolinger (1980), “a noun does not lack a determiner because it is a proper name, it is a proper name when it lacks a determiner.” If this definition is quite effective in English, it is much more problematic in other languages: one can add a determiner with a first name in dialectical varieties of French, Italian or German, and they are required in Portuguese or Greek. In Greek, even the names of city are preceded by a determiner, one will thus speak about “The Athens”.

Other authors try to find semantic arguments. Read (1974) notices that the verb that the verb which accompanies The United States is often in the singular in American English. It looks for him as a characteristic desemantisation of the proper name, but The United States will be a full proper name when it will be used without a determiner. As it is not the case in French (the verb will always have a plural inflection with “les Etats-Unis”), would it mean it is not a real proper name in this language? It is another unacceptable conclusion. Several linguists do not wish to find a descriptive content in proper names. For instance, Brondal (1948) thinks we should exclude from this class syntagms like la tour Eiffel but also New-port, Deutsch-land, Cam-bridge, and derived words like (French) Ita-lie, Turqu-ie.

The will to present all proper names as non-descriptives leads to the absurd situation presented by Kałuza (1968), where London University (proper name) is non descriptive whereas the University of London (definite description) is.

Finally, some authors think names are not nouns like Anderson (2007) who says they are part of the determinatives class (along pronouns and determiners). According to Hjelmslev (1971), numerals and proper nouns belong to the class of pronouns. We will defend the classical view of the proper nouns as a subclass of the noun and define them as monorefential frozen lexical forms.

References
Brondal, V. (1948), Les parties du discours — Partes orationis, Copenhagen, Einar Munksgaard [1928].
Sørensen, H.S. (1958), Word-classes in Modern English — with special reference to proper names, with an introductory theory of grammar, meaning and reference, Copenhagen, G.E.C. Gad Publisher.
This contribution examines some non-canonical usages of lexical items in French. As a result of their use in a construction typical of another word class than the one they are usually associated with, they exhibit mixed word class properties. Thus, this paper addresses two central topics of the conference, viz. constructions and continua between word classes.

Starting from two case-studies based on original data taken from informal French:

(a) (a1) le simple et le beau ‘the simple and the beautiful’
   (a2) voilà comment on fait du très beau avec du simple ‘that is how you can make beautiful things with simple things’

(b) des costumes très ‘théâtre’ ‘very theater-like costumes’

it will be shown that both alleged cases of respectively (a) nominalization and (b) adjectivization display categorial deficiency with respect to the target category. Cases such as (a), in which a descriptive adjective is preceded by the singular definite article (a1) or the partitive article (a2), exhibit stable restrictions on modification, determination and anaphoric uptake, and at the same time, can still be submodified by a degree adverb (Lauwers 2008). These observations attest the categorial hybridity of such NPs, which are both nominal and adjectival, and lead us to reject all accounts based on null elements or ‘full’ recategorization (morphological conversion, lexical relisting, etc.). Instead a syntactic (constructional) account based on constructional overrides and coercion (Michaelis 2003) that captures both restrictions and meaning effects, e.g. massification over entities sharing the same property (= a) and extraction of stereotypical characteristics of the nominal referent (= b). Finally, it will be argued that these creative/productive overrides are subject to lexicalization, yielding additional (idiosyncratic) meaning shifts and (progressively) full membership of the target category.

On a theoretical level, this study brings together several research lines that have not yet been related to each other, such as coercion (which, as will be argued, can also exhibit morpho-syntactic effects) and several accounts concerned with category/function mismatch, both in the Anglo-saxon and French (European) linguistic tradition (distorsion catégorielle; translation; Kerleroux 1996; Tesnière 1959; Werner 1993).

References

Christian Lehmann
The nature of parts of speech [INVITED PRESENTATION]

The categorization of linguistic signs has two independent motivations: It has a cognitive basis in the types of concepts coded in language; and it has a structural basis since elements of a semiotic system of some complexity must be arranged in classes. Consequently, each of the categories of parts of speech has both semantic and structural aspects. The two sets of features are essentially incommensurate, since the semantic features are partly based on extralinguistic cognition, which the structural features are essentially based in the combinatorial potential of the signs. Consequently, the two sets of features are largely independent of each other. Their combination in a language yields sets of parts of speech which deserve the name of part-of-speech system only to some extent.

Yucatec Maya has at least 19 parts of speech with a different distribution: verb, verboid, auxiliary, noun, adjective, possessive classifier, quantifier, pronoun, article, enumerative, numeral, adverb, adverboid, free particle, enclitic particle, ideophone, interjection, preposition, conjunction. Naturally, most of them have their subclasses. These classes may partly be systematized by certain criteria that structure part-of-speech systems universally. There remains a rest of language-specific categorization which derives from the particular combinatorial potential of these word classes in Yucatec.

References
Ansaldo, U., J. Don & R. Pfau (eds.) (2008), Parts of speech, in “Studies in Language” (Special Issue) 32 (3).


Alessandro Lenci
Carving Italian verb classes from corpora [INVITED PRESENTATION]

I will report on current research aiming at building verb classes on the grounds on their distributional behavior. State-of-the-art computational methods have been applied to extract from a large corpus of Italian statistical information about the argument structure of verbs, their selectional preferences and argument alternations. These data have been used to identify classes of verbs sharing similar argument distributions, with the goal of investigating on a large scale the correlation between the paradigmatic organization of Italian verbs and their combinatorial properties. Various results of this analysis will be presented, together with their consequences for the study of the syntax-semantics interface.

Beth Levin
What is the best grain-size for defining verb classes? [INVITED PRESENTATION]

Fillmore’s study, “The Grammar of Hitting and Breaking” (1970), demonstrates the importance of semantically coherent verb classes as descriptive devices for understanding the organization of the verb lexicon and for capturing patterns of shared verb behavior. Much subsequent work confirms and extends the findings of his study, although other work provides support for positing verb classes of other grain-sizes. Specifically, some researchers suggest splitting Fillmorean classes into smaller classes (e.g., Boas 2006, commenting on Levin 1993), or even treating each verb as representing a unique class.

I propose that there are three particularly useful levels of grain-size for verb classification. Taking the English verb run as an example, at a Fillmorean, or medium-grained, level of classification, it is a manner of motion verb, contrasting with verbs in other manner verb subclasses, e.g., manner of speaking verbs or surface contact verbs. From a coarse-grained classification standpoint, run is a manner verb (and not a result verb like go). Turning next to a fine-grained classification, among manner of motion verbs, run lexicalizes a manner of motion typically intended to cause directed displacement towards a goal (as opposed, say, to roam, stroll or wander) (Allen et al. 2007, Jones 1996).

Rappaport Hovav & Levin (1998, in press) argue that the coarse-grained manner vs. result verb dichotomy is important for establishing the basic behavioral properties of verbs; this distinction receives independent support in studies of child language acquisition. Researchers also argue for finer-grained verb classifications based on the non-uniform behavior of verbs in various Fillmorean classes. Boas (2008), for example, classifies 20 English manner of motion verbs into four groups, represented by the verbs walk, parade, stagger and totter. These groups form a cline with respect to whether their members allow certain grammatical properties; Boas proposes that the number of properties verbs of each type show is inversely correlated with the specificity of the manner they lexicalize: verbs showing the most properties lexicalize the least specific manner.

This talk explores the role of finer-grained classifications by asking the following question: Does recognizing finer-grained classes among, say, manner of motion verbs, obviate the need for recognizing the larger, Fillmorean class? I argue that it does not. I show that membership in a Fillmorean class is necessary for showing certain facets of behavior, even if it may not be sufficient; however, the finer-grained manner detail plays a part in determining sufficiency. Drawing on Levin, Song & Atkins’ (1997) study of English verbs of sound emission (e.g., creak, rumble, whir), I show that it is not simply the degree of manner specificity that matters, but rather the particular facets of manner and their interactions with each other. Thus, even if finer-grained verb classifications based on manner specifications are critical to our understanding of verb behavior, Fillmorean classes remain equally important.

References


---

**Martina Nied and Anna Pompei**

*Phrasal verbs, prefixes and word order [PRESENTATION]*

This talk would aim at analyzing a sub-class of verbs, phrasal verbs, from a syntactic point of view. In particular, the analysis looks at the relationship between phrasal verbs, prefixed verbs and word order. It has been often observed that phrasal verbs are linked to prefixed verbs, either semantically (Iacobini & Masini 2007), or as two distinct steps in the process of grammaticalization (Cuzzolin 2005). What is proposed here is that phrasal verbs and verb affixation can actually be part of a continuum that goes from syntax to lexicon, i.e., from ‘constructions proper’ to ‘phrasal words’ (Simone 2006), until univerbation is reached, first by means of compounding, and subsequently by derivation. This process is clearly shown in Homeric poems, where the peculiar textual tradition enables the coexistence in the same verses of very early phrasal verbs - characterized by more (2) or less (1) cohesion - and of new prefixed verbs containing the same particles (3):

1. ἐβως ἐς ἰλιον ἕλε
   ‘he came to Ilions again’ [v 175]
2. ἐς δη ἔθον μηνηρις ἤγηρες
   ‘then the suitors came in’ [a 144]
3. ἐβως κληριν ἐοφιλε
   ‘he went inside his tent’ [Π 254]

Homeric Greek is an OV language, which implies that the particle is usually placed to the left of the verb in all phases of the grammaticalization process. On the contrary, in Old English – where word order shifts from OV to VO – the particle may be placed both to the left and to the right of the verb in main clauses, but only to the left in subordinate clauses, which are conservative. This is the case in the following passage, where there is a phrasal verb with a particle to the left in a relative clause:

4. Wæron da Noes suna de of dam arce eodan
   ‘And the sons of Noah who went out of the ark were...’ [Heptateuch, Gen. IX.18]

German shows even more interesting data. In main clauses (VO), particles are post-posed to the verb, at the end of the clause, whereas in subordinate clauses (OV) the same verbs are prefixed.

Considering these and other Indo-European languages, the talk aims at:

- showing that the position of particles in phrasal verbs is constrained by word order;
- analyzing the phonological and morphological reasons for the full grammaticalization of particles in OV languages, as opposed to the blocking of the process before the full compounding in VO languages.

**References**


One of the major questions in word class research is: “How many word classes are there in language X?” As for this question, the field is divided between lumpers and splitters (Croft 2001). Splitters take even small distributional differences as constitutive of membership in different classes, while lumpers are willing to ignore differences if on the overall view the members of a proposed word class are more similar to each other than to lexemes outside the class.

The range of different distributional classes proposed ranges between 1 for North American or Austronesian languages (e.g. Broschart 1997) to 12,000 for French (Gross 1979). The interesting thing is that, although the interpretation of these facts may vary, there is general agreement about the distributional properties of a Lexeme Lx in a given language Lg. It is clear that the question is not an empirical one about the distributional properties of lexemes, but a methodological one of what counts as similar.

Mathematical approaches can help in deciding that question. For n distributional properties and m lexemes, we obtain an n x m matrix. This matrix defines an n-dimensional space, in which we compute the distance between lexemes (represented as vectors) and obtain a distance matrix M. We then cluster the lexemes according to their distances. The resulting clustering is evaluated for intra-cluster and inter-cluster dispersion. We define a threshold \( \theta \) for the maximum dispersion we are willing to tolerate within a cluster and a threshold \( \tau \) for the minimum dispersion we require between clusters. The number of word classes is then a function \( F(M, \theta, \tau) \) of the distance matrix and the two thresholds.

Interpreted in this mathematical framework, the question of whether there are 1 (Peterson 2005) or 2 (Evans & Osada 2005) word classes in the Munda languages boils down to the best values to choose for \( \theta \) and \( \tau \). Evans & Osada clearly advocate a value of \( \theta=0 \), while the value in the “lumpers’” case will be higher. The methodology proposed in this talk can help shift the discussion from the dead-end of a binary in/out model for word class membership towards a discussion of which dispersion thresholds are sensible in the domain of word class research. The talk will be illustrated with clusterings from English and projected clusters for different constellations frequently discussed in word class typology.

References
Peterson, J. (2005), There’s a grain of truth in every “myth”, or, Why the discussion of lexical classes in Mundari isn’t quite over yet, in “Linguistic Typology” 9 (3), 391-405.

Anna Orlandini and Paolo Poccetti
Polifunzionalità e transcategorialità delle particelle di coordinazione in latino [POSTER]

Una classificazione categoriale delle particelle di coordinazione che distingua in modo netto le congiunzioni di coordinazione (copulativa, connettiva, disgiuntiva e avversativa), gli avverbi, i connettori, i marcatori discorsivi, è difficile da farsi.

Come esiste un continuum tra subordinazione e coordinazione (cf. Simone 2009), tra scrittura e oraleità, esiste altresì una continuità transcategorialità per le particelle coordinanti. La loro polifunzionalità ne è la prova. Noi analizzeremo il funzionamento delle particelle latine et (‘e’), at (‘ma’), atque (‘eppure’), aut, vel (‘o’) e dei loro corrispondenti sabellici.

Gli “effetti di senso contestuali” che facilitano l’interpretazione delle particelle come marcatori discorsivi sembrano caratterizzare particolarmente la lingua parlata. Abbiamo quindi cercato di avvicinarci all’oraleità, per quel che è possibile trattandosi di lingue morte. Il nostro corpus si compone di dialoghi di commedie, di discorsi di Cicerone che rappresentano un grado diverso di oraleità, in quanto redatti in forma scritta e destinati alla lettura, e di iscrizioni italiche, testi scritti ma considerati molto vicini alla lingua parlata.

Un’attenzione particolare è riservata alla posizione delle particelle nella frase. Si è notato, in particolare, che la posizione in inizio di frase o di enunciatum incrementa la polifunzionalità e la possibilità di scambio tra le particelle. L’altro parametro è il livello semanticale al quale le particelle operano. A livello “content” (cf. Sweetser 1990) le particelle di congiunzione e di disgiunzione funzionano per lo più come operatori logici. Le strutture a cui appartengono sono vero-funzionali. A livello semanticale superiore (“epistemico”, “interpersonal level”), i marcatori discorsivi di coordinazione segnalano una relazione di coerenza tra le proposizioni che introducono ed non sono vero-funzionali.

Particolarmente interessante è l’uso di particelle di coordinazione dinnanzi ad una principale (“paraipotassi”):

(1) *computa, et tibi plus do quam accepi* (Petron. 45,13)

‘fai il conto e io ti do più di quanto ho ricevuto’

La natura di questo et è quella di un avverbiale deittico (‘ed ecco che’). Serve a rinforzare la coesione discorsiva tra frasi semanticamente autonome: La sua origine deve essere rintracciata nell’antico dittico correlativo indo-europeo (in latino *cum ... tum* ; *cum ... atque; sic ... et* ) ; negli enunciali condizionali la particella coordinante enfatizza l’apodosi.
Terms indicating numbers and quantities are commonly classified by pedagogical or reference grammars as ‘Numerals’. Classifying a word by such metalinguistic label implies stating that its content and its distributional characteristics are those shared by the general body of membership of that word class. As is well known, it is not always the case that all members of a particular word class share the same set of properties.

The first part of the study consists of a cross-linguistic analysis of syntactic and semantic properties of number words in a sample of genetically and typologically representative languages, selected from an analytic database of number words in languages of the world (AUNIN\(^1\)), so far including 180 languages. Based on the data, the question is addressed whether ‘numerals’ can be regarded as a word class of their own. The analysis will identify the different morphosyntactic strategies for the expression of the conceptual contents of number words, thereby showing how number words could rather be regarded as specific subclasses of other classes (Greenberg 2000: 770). For instance, in Maori (Polynesian, New Zealand; Harlow 2007:112), numerals are verb-like in that they regularly occur with a tense/aspect marker, usually e, and are characteristicly used as relative clauses to modify nouns (as in (1) below):

(1)  ōku  whare [e rua]relative clause
     Pl.-of-1Sg. house  TA two
     ‘my two houses’, ‘my houses which are two (in number)’

The second part of the study focuses on whether a syntax-semantics interface approach (Van Valin 2005) can provide a useful descriptive and interpretative framework for investigating the organization of number words in sentences and the related sets of rules and constraints. The aim is to capture cross-linguistic generalizations about the ways in which number words contribute their quantitative reference, thereby redefining the class of number words as a constellation of morpholexical and morphosyntactic properties, that might, possibly, be shared with other lexical/grammatical categories.

References

Aranzazu Pascual Ortiz
Where have all the pro-complement verbs gone in Italian? [PRESENTATION]

In Italian, there are a number of verbal constructs consisting of a verbal nucleus and one or two clitic elements of pronominal or adverbal origin. Such verbs present a noteworthy change of meaning with respect to the base meaning of the lexical verb (e.g. voler-ci ‘to need’, aver-ce-la ‘to be hungry at’, prender-se-la ‘to be hungry at’, tirare su ‘levantar la moral’, buttare giù ‘deprimir’). Simone (1996) has classified these constructions as “verbi sintagmatici” or phrasal verbs, independently of whether the clitics are from adverbial or pronominal origin. De Mauro (1999) and Russi (2008) prefer instead to classify verbal constructs with pronominal clitic as forming the so-called class of pro-complement verbs.

\(^1\) The AUNIN (Universal Atlas of Numerals and Numerical Expressions) research project, designed and directed by Domenico Silvestri at the Università degli Studi di Napoli “L’Orientale”, financed as a PRIN 2000, includes among its results an electronic database (initially outlined in Pannain 2000) aimed at providing analytic data for cross-linguistic typological studies of numerals and numeral systems, such as the syntactic study by Riccio (2007).
This paper posits the question of whether these two types of verbal constructs, that is, those with either adverbial or pronominal clitics, pertain both to one unified verbal class or whether they must be considered as two different categories displaying distinct characteristics.

In this direction, I will propose the existence of just one verbal macro-class, and I will posit the question of whether this class should be treated as a pro-complement class or as a phrasal verb class, as proposed by Simone (1996), or alternatively, whether it would be better to regard the class as a Verb Particle Construction in order to accommodate such verbal constructs and other similar in Italian to the terminological and conceptual framework proposed for other languages.

To such a goal, I take into account morphological, semantic and pragmatic characteristics of these verbs and will provide information about frequency data from both written and oral corpora of standard Italian. I also propose a number of criteria for determining membership to the macro-class so that it becomes a convenient category to be used accordingly in both dictionaries and corpora of Italian.

To conclude, I additionally point out a number of issues involving grammaticalization and lexicalization, which played an important role in the emergence of such constructs.

References
De Mauro, T. (1999), Dizionario della lingua italiana Torino, Paravia

Paolo Ramat
Reflections on word classes [PRESENTATION]

Among the criteria suggested for the definition of word classes (‘parts-of-speech’ [PoS], according to the classical tradition) the model proposed by Pullum (1994) seems to be the most fitting. A word class is a set of linguistic objects that are considered as having common features, which, on their turn, are implemented by particular values. Thus, e.g., one of the possible implementations of the feature Mood is Subjunctive and one implementation of the Gender feature is Feminine. Some of these features can be shared by other objects, but not all of them at the same time. So, we observe that Verbs (VBs) may share the features Gender and Number with Nouns and PRO(nouns). That’s why PoSs are not water-proof boxes and lexemes may shift from a word class to another. See, e.g., the Engl. -ing forms whose meaning can change according the sentence construction they enter: These cars want washing, carefully versus These cars want careful washing, (point 4 of the list: ‘continua between classes’, or ‘overlapping categories of formatives which represent their diverse distributional behaviour’, Croft 2007: 421). In Samoan, a language where VB, N, ADJ, ADV do not show any specific morphological marker, the function is assigned to a lexeme only by its syntactic use. German or Russian are, on the contrary, ‘specialized languages’ where VB, N, ADJ and – partially – ADV have each particular markers. Categorial shifts (i.e. transcategorizations) depend on the morphosyntactic constructions a lexeme is involved in (topic 2: ‘word classes and constructions’): cf. the Engl. preposition bar ‘except’, actually the imperative of the verb to bar, when used in a construction like Everyone is leaving bar Ernst.

On the other hand, it is difficult to think of Ns having the values of “Subjunctive” or “Optative”: “there are typological restrictions with regard to the conceptualisation of semantic features and morphosyntactic structures”, Vogel & Comrie (2000: ix).

Broad-range comparisons show that some features and values are more common than others and that, on the contrary, some other features and values are quite rare (e.g. the verbal value ‘Polarity’). Accordingly, we can sketch the prototypical figure of what is a VB, a N, an ADJ, etc. If we adopt the semasiological point of view, i.e. if we depart from a linguistic form and ask what does it mean, we can extract its matrix which shows the features and values that characterize it. The binary analysis a la Jakobson can help us in drawing the matrix. For instance the Latin form amabantur ‘they were loved’ will show following morphological matrix:

VALUES: [+indic. +impf. +imperfective +passive +3rd pers. +plur.]; →
FEATURES: [+mood +tense +aspect +diath. +person +number] →
CATEGORY: VB
(see Ramat 1999)

From a more functionally oriented point of view the Kumyk (Turkic) converbial form görüp can be analysed as follows: gör ‘see’, vb + ip converbial mark: VALUES: [‘+dependent’, ‘-argumental’, ‘-adnominal’, ‘-finite’, ‘+embedded’] → FEATURE: ‘mood’ → CATEGORY: VB (see van der Auwera 2000).

References
Polish Sign Language (hereinafter PSL) is an understudied natural sign language used by the Deaf community in Poland. The aim of the present paper is to discuss the grammatical properties of the word class of pronouns in that language.

One of the most characteristic typological features of sign languages around the world is a close relationship between indexical pointing and pronominality. The focus of attention of this paper is on PSL constructions containing the so-called pointing sign (an index handshape directed to a point in the signing space). I argue that the indexical should be viewed as a grammaticalized linguistic element (pronoun), rather than a mere gesture. It may be treated as analogous to demonstratives in spoken languages. It is also used as a personal pronoun: “I” (directed towards the speaker), “you” (directed towards the person spoken to), “he/she/it” (directed towards one of a potentially infinite number of points in the signing space, previously associated with the referent in question). However, I assume that the
distinction between “demonstratives” and “personal pronouns” is not applicable to PSL. Instead, I argue that there is only one pronoun in PSL. I follow Diessel’s (2006) observation that, crosslinguistically, personal pronouns are derived from demonstratives, which means that the latter are more basic and belong to the universal set of core vocabulary. By proposing the above analysis, I attempt to take a position in one of the most important debates in contemporary sign linguistics. According to many researchers, personal pronouns in sign languages are typically inflected for person. Meier (1990) argues for a two-person pronominal system – first and non-first. Liddell (2003), on the other hand, assumes that the use of space for pointing is purely gestural (gradient) rather than linguistic. I follow Liddell (2003) in assuming that there are no reasons to distinguish first and non-first person in PSL; I show that even “first-person” pronouns may be subject to displacement, e.g. when telling a story about oneself; however, this does not mean that pointing is nonlinguistic. The pointing sign is a pronominal element but, unlike its counterparts in spoken languages, it is not associated with person features. The only formal feature associated with the pointing sign is that of referentiality (therefore, it resembles definite articles in spoken languages). The sign in question enters into a two-directional matching relation with the context (the signing space), thus being provided with a semantic value. Therefore, the signing space becomes a kind of grammatical feature, unparalleled in spoken languages.

References

Galit W. Sassoon

A cognitive root for a word class distinction [PRESENTATION]

Consistent results concerning artificial category learning are revealed only when studies are divided by category type.

Classification in rule-based categories depends on a single dimension or a simple enough conjunction or disjunction of dimensions. Conversely, in similarity-based categories information about instances’ degrees in multiple dimensions is integrated, typically by averaging. In accordance, rule vs. similarity tasks have different neural and developmental correlates. I propose that adjectives and nouns cluster with these two category types, i.e. typically trigger processing by two different cognitive systems.

Adjectives with multiple dimensions have a ‘dimension’ argument slot, which can be saturated (as in healthy with respect to blood pressure), or bound (as in are healthy in every respect). When the adjective occurs bare (as in Dan is healthy), the dimension argument is implicitly saturated or bound. In both cases, speakers and listeners have to pay selective attention to one of multiple dimensions or to conjunctions or disjunctions of dimensions. This suggests that adjectival dimensions are typically processed with the rule-based system (they combine via Boolean operations).

Conversely, nouns like bird typically do not have a dimension argument (as illustrated by the infelicity of #a bird with respect to flying/ #except for size; #generally a bird.) All ‘exceptional’ nouns (like an Italian or health) have corresponding adjectival entries. These facts suggest that nominal dimensions are integrated via weighted-mean operations at an early processing stage (are typically processed with the similarity-based system).

Based on existing literature, the talk aims to support this proposal by pointing at a link between findings pertaining to rule- vs. similarity-based categories and a variety of corresponding developmental, neural and distributional findings pertaining to adjectives vs. nouns as well as conjunctions.

References

26
Raffaele Simone and Francesca Masini

Light nouns [PRESENTATION]

It is commonly recognized that word classes are not monolithic but are better represented as complex objects including subclasses to be distinguished on various bases. Quite a lot of work has been done to this effect as regards verbs: suffice it to recall the full / light verb opposition in light or serial verb constructions and the differentiation of full verbs into semantic, syntactic and actional classes.

The category noun has also been submitted to classification. Traditionally established subclasses of nouns are, notoriously, common vs. proper nouns, count vs. mass nouns, collective nouns, concrete vs. abstract nouns, and so on (cf. Lyons 1977 and Rijkhoff 2002 for a discussion). However, scarce attention has been devoted to its articulation into finer subclasses as well as to the very possibility to construct “scales” of nouniness, although some attempts (such as the “nouniness squish” proposed by Ross 1973 and the scale of process nouns proposed by Simone 2003) are available.

In this contribution, focussing precisely on the latter issue, we propose a new kind of “nouniness” scale, based on the notion of Referential Force. Our claim is that it is possible to construct for nouns a scale with full lexical nouns on one side [+NOUN] and gradually lighter nouns on the other [-NOUN]. The area in-between both extremes includes nouns displaying a gradually lower Referential Force. This set includes several entities, such as qualifiers (e.g. Italian sacco ‘sack’ or English lot); approximators (e.g. Italian sorta or English sort) and so-called support nouns (e.g. Italian colpo ‘blow’ in expressions such as un colpo di telefono ‘a ring/call’, lit. a blow of telephone).

Some of such classes have been partially identified and discussed in some languages (e.g. English, cf. Traugott 2007, 2008; French, cf. Rouget 2000; Italian, cf. Masini 2007, Simone & Masini 2009). However, a full crosslinguistic characterization of them still has to be done; moreover, the category support nouns are virtually unknown (see Simone & Masini 2009 and Gross 1984 for a seminal study).

In languages such as English, French and Italian, these subclasses of nouns share an important syntactic property: they are bound to form a phrase with a subsequent noun. Indeed, these nouns occur as the first noun of a binominal construction where both nouns are linked by the preposition ‘of’ (e.g. Italian un sacco di ‘a sack of’, lit. a sack of telephone).

In our presentation, light noun constructions will be described in more detail. Special attention will be paid to the support noun construction.

References

Simone, R. & F. Masini (2009), Support nouns and verbal features: a case study from Italian, in “Verbum” XXIX, 143-172.

Anna Sores

On the borders of neglected classes: from preverbs to “satellites” via adverbs and particles [POSTER]

1. The paper proposes reflections on “neglected word classes” on the basis of Hungarian preverbs. We suggest to identify a non-universal word class of directional verb modifiers called satellites (according to Talmy 1985’s terminology). This type of morphemes can be found in several European languages like German (herauskommen), English (to come out), Italian (salire fuori), etc., but they are traditionally listed as particles, adverbs or prepositions. After demonstrating some common features of these morphemes that justify considering them as instances of a single type of word class, the paper presents possible typological developments.

2. Hungarian has more than 40 preverbs, e.g.:

(1) Péter kijöjt/lejöjt
Peter out-down-came
‘Peter came out/down’
Grammars (Kiefer & Ladányi 2000) identify preverbs as an autonomous word class:

- according to Bloomfield’s criteria of wordhood, they are free word forms;
- they are verb modifiers not being an argument of the verb;
- other properties: unmarked place before the verb, detachment in case of negation or focalization;
- other arguments: most of Hungarian preverbs are not homophonous with adpositions or adverbs like for example Engl. out, in, down or German ein, aus, etc.
- preverbs can be used isolated in an answer to Where do you go?

3. If Hungarian preverbs can be considered as the prototype, other languages have less typical elements which share some features with Hungarian preverbs. We can compare them according to the criteria commonly used to identify word classes: closed minor word class; morphologically invariable; syntactic role: verb modifier; semantically, preverbs, particles and adverbs form a heterogeneous class; the distributional criterion does not work crosslinguistically.

4. Conclusions and perspectives. Establishing a prototype of “satellites” based on European languages helps to capture some typological developments. Further studies should first consider languages having preverb-like “satellites” out of Europe, and then go on, in order to include other grammatical phenomena like serial verbs or converbs.

Our proposal contributes to the reflections on criteria identifying word classes: we propose here a flexible word class, according to traditional criteria, but meaning takes a more important place in the definition. In fact, if some features are shared by languages having “satellites”, the most relevant one is their common meaning.

References

Julie Louise Steele
Tinkering with syntactic categories [PRESENTATION]

By adopting a network theory approach, this paper provides a straightforward account of how non-general syntactic categories (evidenced by Croft 2001) emerge through the ‘tinkering’ or duplication process of reusing the same structure for a different function (Jacob 1977; see also Solé & Valverde 2006: Figure 1). Linguistic form may be represented as a network of word co-occurrences (BNC 2001) where a specific node represents a given word and a link represents a precedence relation between adjacent words in a sentence. The typology of the word co-occurrence network is derived from optimising node connectivity i.e. type frequency, and link length i.e. the isomorphism or cohesion between two adjacent words realized in terms of transitional probability (Steele submitted; Mathias & Gopal 2001).

Within this (near-) optimal network, certain areas may undergo a transition to a more isomorphic state that is more computationally friendly. These areas represent potential syntactic category members and their associated function words. Since token and type frequency may identify function words, the interception of frequency and connectivity distribution patterns within the network highlight function words and thus potential (non-general) syntactic categories (Steele submitted). For an isomorphic relation to occur (figure 1, right), an abstract element represents the set of members of a potential syntactic category i.e. \{N_1, N_2, N_3, N_4\} in figure 1 (left).

![Figure 1. Part of the network (left) is reduced to a more isomorphic form (right).](image)

However, this network transition is temporary and these areas move back to a more optimal and automorphic state. This movement induces an automorphic relation between the function word and the abstract element representing the syntactic category. Consequently, the relations of the function word are duplicated for the syntactic category. From this tinkering process, a syntactic category is created (figure 2B). As such, syntactic categories arise “for free”, in the spirit of Solé (2005), from the properties of the word co-occurrence network, with minimal theoretical prerequisites.
The present work aims at pointing out how the well-known controversial issue concerning word class categorization can be investigated through an empirical analysis of the existing morphosyntactic annotation schemes (i.e. PoS-tagsets) for a given language. In this study we will apply such a methodology to the case of Italian language.

This work stems from an analysis carried out by comparing the existing PoS-tagsets for the Italian language. It suggests that the divergences brought forth during that detailed comparison can shed light on the most controversial dimension of variations in categorizing parts-of-speech.

Even though the compared PoS-tagsets broadly agree on the main parts-of-speech (e.g. on the categorization of ‘verb’, ‘noun’, ‘adjective’), several divergences occur. Of the various disagreements found amongst the considered PoS-tagsets, the present work focuses on those which concern the word class defining criteria and the identification criteria of a word. In particular, two case studies which exemplify the aforementioned divergences are described. Namely, the analysis of i) how determination is differently categorized and ii) how clusters of words, such as multi-word expressions and clusters of two (or more) clitics and verb-clitic clusters, is carried out.

In the first case, the lack of agreement on exploiting a morphosyntactic or a functional criterion, or both of them, to categorize words which convey determination is proved by the fact that some PoStagsets include only a Determiner or an Article class, while some have both classes, others have a Pronoun-Determiner class, and so on.

In the latter case, the issue concerns whether each single word contained in a cluster of words is to be categorized or rather the whole cluster is to be considered as a unique word. In particular, the different treatment of clusters of clitics was revealed to be closely intertwined with a different categorization of clitics. These divergences can shed light on the reason why, according to Simone (1983), in literature clitics are described with difficulty.

These investigations are overtly devoted at finding empirical evidence for grounding the morphosyntactic annotation of corpora on linguistic reflections with a view to linguistic analyses.

According to Jurafsky & Martin (2005), “corpora that have been marked for parts-of-speech are very useful for linguistic research”. However, “the process of assigning a part-of-speech […] to each word in a corpus” (i.e. PoS-tagging) is strongly influenced by distinctions and decisions made at the word class categorization stage. Accordingly, grounding a PoS-tagset on linguistically-sounded choices can improve the PoS-tagging of large corpora for fine-grained analyses.

References
