Studi e Saggi Linguistici

SSL electronic version is now available with OJS at *www.studiesaggilinguistici.it*
Web access and archive access are granted to all registered subscribers

La rivista *Studi e Saggi Linguistici* può essere acquistata al prezzo di € 40.00 (estero € 60.00) presso le Edizioni ETS, Pisa, Piazza Carrara 16-19, 56126 Pisa.
È possibile abbonarsi a *Studi e Saggi Linguistici* tramite versamento della cifra indicata sul conto corrente postale 14721567, intestato a:
Edizioni ETS Pisa, Piazza Carrara 16-19, 56126 Pisa
tel. 050 29544; 503868
fax 050 20158
www.edizioniets.com

L'editore non garantisce la pubblicazione prima di sei mesi dalla consegna in forma definitiva di ogni contributo.

STUDI E SAGGI LINGUISTICI

LII (2) 2014

rivista fondata da

Tristano Bolelli

Edizioni ETS

SSL

STUDI E SAGGI LINGUISTICI

www.studiesaggilinguistici.it

SSL electronic version is now available with OJS (Open Journal Systems)
Web access and archive access are granted to all registered subscribers

Registrazione Tribunale di Pisa 12/2007 in data 20 Marzo 2007
Periodicità semestrale
Direttore responsabile: Alessandra Borghini

ISBN 978-884674096-0

Sommario

Saggi

Recensioni

Categories, features and values in the definition of a word class

Paolo Ramat

Abstract

The model suggested by Pullum (1994) in order to define the word classes (i.e. the traditional 'parts-of-speech' [PoS]) seems to be the most adequate among the many models that have been proposed. A word class is a set of linguistic objects having common traits which, on their turn, are implemented by particular values (see fn. 1 for the graphic devices to distinguish between 'categories', 'features', and 'values'). Thus, VERBS may share with NOUNS and PRONOUNS features such as *GENDER* and *NUMBER*. For instance, a possible implementation of the trait – technically called 'feature' – *MOOD* can be *subjunctive* while a possible implementation of the feature *GENDER* is *feminine*. Some values may be shared by different categories or features – though not all at the same time. Word classes are not water-proof boxes and lexemes may leak from a class to another one. For instance, the *-ing* forms show different functions according to different contexts (for a 'construction grammar' approach see Croft, 2007: 421, who speaks of «overlapping categories of formatives which represent their diverse distributional behaviour»): *These cars want washing*$_{VB}$ *carefully* versus *These cars want careful washing*$_{N}$. Category shifts (i.e. transcategorizations) are caused by the morphosyntactic constructs they appear in. Engl. *bar* is properly the imperative of the verb *to bar* "to prevent", but in a sentence like *Everyone is leaving bar Ernst*, it has prepositional function.

Riassunto

Fra i criteri proposti per una definizione delle classi di parole (alias 'parti del discorso' – 'parts-of-speech' [PoS] della tradizione classica) il modello di Pullum (1994) sembra essere il più adeguato. Una classe di parole è un insieme di oggetti linguistici che hanno tratti comuni, i quali a loro volta sono realizzati da valori particolari. Per esempio, una delle realizzazioni possibili del tratto *MODO* è il valore *congiuntivo* e una delle realizzazioni possibili del tratto *GENERE* è il valore *femminile* (vd. nota 1 per la differenziazione grafica tra 'valori' 'tratti' e 'categorie'). Alcuni valori possono essere condivisi da categorie e tratti diversi – ma non tutti allo stesso tempo. VERBI (VBS) possono condividere tratti come *GENERE* e *NUMERO* con NOMI e PRONOMI. Le classi di parole non sono compartimenti stagni e i lessemi possono passare da una classe ad un'altra. Per esempio, le forme inglesi in *-ing* hanno funzione diversa a seconda del contesto ('construction grammar'; cfr. Croft, 2007: 421: «overlapping categories of formatives which represent their diverse distributional behaviour»): *These cars want washing*$_{VB}$ *carefully* versus *These cars want careful washing*$_{N}$. I passaggi di categoria (le transcategorizzazioni) dipendono dai costrutti morfosintattici in cui i lessemi si trovano ad agire. L'ingl. *bar* è propriamente l'imperativo del verbo *to bar* "to prevent", ma in una costruzione quale *Everyone is leaving bar Ernst* ha funzione preposizionale.

Keywords: Parts-of-speech, categories, features, values.

Received: September 2013
Accepted: December 2013

1. *Introduction*[1]

A few general remarks are necessary before tackling the problem of the word classes or part of speech (PoS).

Let us begin with a wonderful quote from the *Grammatica o Arte de la lengua general de los Indios de los reynos del Peru* by Fray Domingo de Santo Tomas, published in 1560 (see Jacquesson, 2008: 195, n. 3):

El plural deste pronombre [quechua] ñoca es ñocánchic o ñocáyco que quiere dezir 'nosotros'. Y es de notar que entre ñocánchic y ñocáyco ay dos differencias: una intrínseca, de parte de la significación dellos; otra extrínseca, de parte del verbo que les corresponde [...]. La primera es que, aunque ñocánchic y ñocáyco significan 'nososstros', el ñocánchic significa 'nosotros', connotando & inclyendo en sí la persona con quien hablamos [...] ñocáyco [...] quiere dezir 'nosotros', connotando que se excluye de aquella pluralidad la persona o personas con quien hablamos. (*Grammatica*, [1560[1]] 1994: 34-35)

"The plural of this [quechua] pronoun ñoca is ñocánchic or ñocáyco meaning 'we'. Note that between ñocánchic and ñocáyco there are two differences: the first one has an intrinsic nature and pertains to the meaning [of the form] itself; the other one is 'extrinsic' and depends on the verb the forms are connected with [...]. The first resides in the fact that, although both ñocánchic and ñocáyco mean 'we', ñocánchic refers and includes the person we are speaking to [...] ñocáyco [...] means 'we' and excludes from the plurality the person(s) we are speaking to".

The opposition between inclusive and exclusive 1st Plural Pronouns is clearly grasped by Fray Domingo, whose mother tongue is Spanish, a language where the opposition is not grammaticalized. This corresponds to a semasiological approach to linguistic problems: given a certain form, what is its meaning? Probably, a native speaker of Spanish would not have formulated the abstract onomasiological question "How is the difference between inclusive and exclusive expressed?". This is to say that it is necessary to distinguish between formal categories and semantic (conceptual) functions.

[1] List of abbreviations: ACC = Accusative; ADJ = Adjective; ADm = Admirative; ADP = Adposition; ADV = Adverb; alien. = Alienable; CONV = Converb; DAT = Dative; FEM = Feminine; GEN = Genitive; Imperf = Imperfect; inalien. = Inalienable; Indic. = Indicative; MASC = Masculine; N = Noun; NTR = Neuter; Opt = Optative; PART = Partitive; PLUR = Plural; PoS = Parts of Speech; PREP = Preposition; PRON = Pronoun; PPRON = Personal Pronoun; PRT = Particle; Sing = singular; Subj = Subjunctive; TAM = Tense, Aspect, Mood; VB = Verb. I am using *italics* to refer to *values*, *SMALL CAPS* in italics for *FEATURES*, and SMALL CAPS in roman type for CATEGORIES.

Accordingly, linguists have to start empirically from linguistic forms, the only data they have at their disposal. Extensive cross-linguistic comparisons have uncovered linguistic properties which are not shared by all languages. For instance, the opposition between 'alienable ~ inalienable possession' is expressed in German and Polish but it becomes evident only via a cross-linguistic comparison, as there are languages, like Italian, that do not make such a distinction: see German *mein* vs. *mir* ((1a) vs. (1b)) and Polish *mu* vs. *go* ((2a) vs. (2b)), but just *mi* in Italian (3a-b).

German

(1) a. *Ich wasche* ***mein***$_{(\text{POSS})}$ *Auto*
I wash my.ACC car
??*Ich wasche* ***mir***$_{(\text{PRON})}$ *das Auto*
I wash me.DAT the car
"I'm washing my car" [alienable]

b. *Ich wasche* ***mir*** *die Hände*
??*Ich wasche* ***meine***$_{(\text{POSS})}$ *Hände*
"I'm washing my hands" [inalienable]

Polish (Frajzyngier, 1997: exs. (7)-(8))

(2) a. *Dali mu w czapę* (ACC)
they.gave to.him.DAT PREP hat>head
"They hit him on the head" [inalienable]

b. *Uderzyli* ***go*** *po kieszeni*
they.hit him.ACC PREP pocket
"They hit him in the pocket" [alienable]

Italian (Heine, 1997: 17)

(3) a. ***Mi***$_{(\text{PRON})}$ *lavo le mani*
to.me.DAT I.wash the hands
*?*Lavo le mie*$_{(\text{POSS})}$ *mani*
I.wash the my hands
"I'm washing my hands" [inalienable]

b. ***Mi***$_{(\text{PRON})}$ *lavo la macchina*
to.me.DAT I.wash the car
??*Lavo la mia*$_{(\text{POSS})}$ *macchina*
I.wash the my car
"I'm washing my car" [alienable]

2. *Categories, features, values*

We can now move to discuss the nature and status of word classes. A class is a set of physical or mental objects that are considered as having common features. A word class is a set of linguistic objects that are considered as having common FEATURES which, in turn, are implemented by particular *Values* (see Pullum, 1994; Ramat, 1999). It goes without saying that the analysis starts from real lexical items and proceeds towards more and more general concepts. For instance, from the notion of *Subjunctive* we pass to the notion of MOOD and from MOOD to the general, categorial notion of VERB, even if in some languages (e.g. in Hausa) MOOD can be marked on PROS. PoS (i.e. word classes) are characterized by particular values and features. Accordingly, every lexeme will present its specific profile or matrix. The binary analysis *à la* Jakobson can help us draw the matrix, and if we adopt the semasiological point of view – i.e. if we start from a linguistic form, and ask what it means – we can extract a matrix which shows the features and values that characterize that linguistic form.

For instance, the Latin form *amabantur* "they were loved" has the following morphological matrix (see Ramat, 1999)[2]:

(4) Morphological matrix of Latin *amabantur*
Values: [+ indic. + imperf. + imperfective + passive + 3rd pers. + plur.] →
FEATURES: [+ mood + tense + aspect + diathesis + person + number] →
CATEGORY: VERB

The corresponding singular *amabatur* "he/she/it was loved" shows a different matrix as regards its *Values*: [+indic. +imperf. +imperfective +passive +3rd pers. –plur./+Sing]. The same holds for the infinitive *amare*, which presents another pattern with the value +*Infinitive* which, in binaristic terms, means –*Indicative*, –*Subjunctive*, –*Optative* etc. But the features PERSON and NUMBER are lacking. On the contrary, the feature TENSE has to be kept, as there are other infinitives (*amavisse, amaturum esse, amaturum fuisse*). As I said before, each entry, i.e. each lexical form, has its own scheme.

This analysis is based on purely morphological characteristics:

[2] For the prototypical matrix of the VERB category see below in this section.

Kumyk

(5) *Bular, kölnü* ***görüp*** *...* *čemodanny ačyp,* *... šišlany*
they lake.ACC see.CONV suitcase.ACC open.CONV bottle.ACC
čyğaryp... *čemodanğa* *sala*
take.out.CONV suitcase.DAT put.PRES
"They see the lake...open the suitcase...take out the bottles...put them back into the suitcase"

The Kumyk (Turkic) sentence has many converbal forms and just one finite verb (*sala*). On a more syntactic and functional basis, the first converb, *görüp*, has been analysed as follows (van der Auwera, 1998): *gör* "see"$_{VB}$ + the converbal mark üp, with the *Values*: [+dependent, -argumental, -adnominal, -finite, +embedded] → *FEATURE*: "mood" → CATEGORY: VERB. The *Values* in van der Auwera's (1998) analysis extend to the syntax since they include ±dependent and ±embedded. Even the pragmatics is (at least partially) included because of the presence of ±argumental.

Verbs may share with Nouns and Pronouns (and hence also with Adjectives) features like *GENDER* and *NUMBER*. That is why PoSs are not waterproof boxes and shared features may make easier for lexemes to shift from a word class to another. See, for instance, the Engl. *-ing* forms in the following examples, whose meaning change according the sentence construction they enter:

(6) *These cars want washing*$_{VB}$ *carefully* vs. *These cars want careful washing*$_{N}$

As Croft (2007: 421) argues, we are faced with «overlapping categories of formatives representing their diverse distributional behaviour» (I will return to this point at the end of this paper).

Broad range comparisons have shown that in a category some features and values are more common than others and, conversely, some other features and values are quite rare (e.g. the verbal value *Polarity* or the nominal value *Obviative)*. Accordingly, we can sketch the prototypical features of a VB, a N, an ADJ, and so on. The notion of 'prototype' (also including the concept of 'family resemblances' viewed as a continuum) is crucial for this typological approach. If we accept this approach the often asked question "where is the cut-off point between two categories ?" is not appropriate: there are not clear-cut boundaries between adjacent categories (think for instance of ADVERBS and ADJECTIVES: the same form may function as ADJ or as ADV, as in Germ. *Sie ist schön* "she's beautiful" and *Sie singt schön* "she

sings nicely". On the phenomenon of transcategorization see Ježek and Ramat (2009), where many cases are discussed. NOUN and VERB are the most diffused categories, probably they are universal categories (see below). But the distinction between the two is not that clear-cut from the morphological point of view[3].

Moreover, if we accept the notion of 'prototypical representative(s)' of a word class the choice of the defining properties of the class members is not arbitrary but it comes out from a cross-linguistic large comparison.

We may represent the matrix of the VERB category as follows (cf. Pullum, 1994: 480, adapted; Ramat, 1999; I shall comment below this scheme):

(7)	CATEGORY		*FEATURES*	*values*	
	VERB	⇒	±*TENSE*		
			±*MOOD*	⇒	[±*indic*_]
			±*ASPECT*		[±*subj*]
			±*DIATHESIS*		[±*opt*]
			±*NUMBER*		[±*adm*]
			±*PERSON*		[±*sing*]
			± .		[.]
			± .		[.]
			± .		[.]

[3] Already GIVÓN (1979: 320-322) observed that there is a continuum NOUN-ADJECTIVE-VERB as far as the temporal stability is concerned: Japanese ADJs may have verbal endings such *waka-i* "to be young", *wakak-at-ta* "to have been young" with the same *-ta* past suffix we find in *tabe-ta* "I ate". Prototypical nouns referring to physical and psychic objects (e.g. "dog", "book", "liberty", "justice", etc.) are, according to GIVÓN (1979), the most stable category: in a temporal perspective they don't change and a dog is always a dog, regardless of his time location. Prototypical verbs, on the contrary, are most unstable as they refer to events ("to eat", "to run", etc.) which need to be time-located via tense markers. What matters for the present discussion is that nominal forms of the verbs such as participles, gerunds, infinitives and converbs are placed between the two poles. For instance, verbal nouns may show verbal behaviour and govern an accusative; see the Vedic example:

(i) *hántāhám* *pr̥thivī́m*$_{\text{ACC}}$ (*RV* X 119, 9)
destroyer.I earth
"I will destroy the earth".

Finally, remember that 'participle' is the Latin translation of Greek *metochḗ* "participation" (scil. of verbal and nominal nature): cp. Finnish:

(ii) *Pekka* *uskoi* *Jukan* *luke-va-n* *kirja-a*
P. thought J.GEN read-PRT- GEN book-PART
"Pekka thought that Jukka was reading a book" (lit. ...of Jukka reading a book)

The dots set up in column indicate that other features may concern a specific (but not prototypical) verbal matrix (see below).

Simone (2008) quotes examples of *TENSE* markers detached on NOUNS («tempo distaccato sul Nome») as in Turkish:

(8) Turkish
mühendis-ti-m
engineer-PAST-1Sing
"I was an engineer"[4]

Tense, Aspect and Mood (TAM) are typically verbal features. Nonetheless, according to the principle that linguistic features may be shared by different categories, we can expect *TENSE* to be also expressed in Ns, ADJs (as in the case of Japanese adjectives quoted in fn. 3) or in ADVs (see, for instance, *the president-to be*). We may propose the following implication: if a language marks *TENSE* on Ns, ADJs or ADVs, it will mark *TENSE* also on VBs with more than chance probability. On the other hand, it is hard, though theoretically not impossible, to think of Ns as having the values of *Subjunctive* or *Optative* depending on the grammatical context they enter. Special phonological or morphosyntactic 'ad hoc' markers might be used (see Frajzyngier's quote below). I am inclined to agree with Vogel and Comrie (2000: ix). when they affirm that «there are typological restrictions with regard to the conceptualisation of semantic features and morphosyntactic structures». But we have to recognize that Vogel and Comrie's 'typological restrictions' are not logical deductions but empirical observations that may admit counterexamples. At any rate, «[a]s typological research has accumulated, it has become evident that [also] inflectional categories and syntactic functions do not vary at random but are drawn from very restricted sets» (Anward, Moravcsik and Stassen, 1997: 168).

As I said above, broad range comparisons have shown that the matrix of

[4] 'Nominalized predicates' (i.e. non-finite complement clauses, as the forms with *-DIK* and *-(y) AcAK*, participles and the verbal nouns in *-mAk, -mA* and *-(y)iş*) have sentential properties, i.e. the nominalized predicate can govern further arguments within embedded clauses and may bear mood and tense/aspect markers on a certain degree. On the other hand these nominalised predicates may bear case markers like nouns:

(i) *"Ali'-nin gel-eceğ-in-i" söyle-di-m*
Ali-GEN come-FUT-3Sg-ACC say-PRET-1Sg
"I said that Ali is going to come" (lit. of A. his future coming I said)
(COŞKUN, 2010: 203)

the category VERB usually contains the features ±*TENSE*, ±*ASPECT*, ±*DIATHESIS*, ±*MOOD*, ± *PERSON*, ±.... As already noted, dots mean that other features may concern the verbal matrix, as, e.g., *GENDER* (but they do not belong to the prototypical matrix of VERB). The ± symbol means that none of these features is obligatory: a language may lack the feature *DIATHESIS* or not possess tense distinctions, which may be expressed by other lexical means (e.g. by adverbs meaning "in the past, now" etc.). In its turn the feature *MOOD* will be implemented by values such as ±*indicative*, ±*subjunctive/irrealis*, ±*optative*, ±.... Once again the list of possible verbal values is not complete: *evidentiality* is an important modality speakers use in their utterances. But only a limited number of languages does have a grammatical (morphological) marker for it, as the Turk. *-mIş* conjugation. In Aikhenvald's words,

> Linguistic evidentiality is a grammatical system (and often one morphological paradigm) [...]. Saying that English has 'evidentiality' [...] is misleading: this implies a confusion between what is grammaticalized and what is lexical in a language. (Aikhenvald, 2004: 6, 10)

It should be evident that our matrixes refer to forms, in a semasiological approach: what does the form *amabantur* express, what features and values are morphologically represented? From the onomasiological, i.e. functional point of view, it is evident that English is perfectly capable to express evidentiality, as in (9):

(9) *The queen would have said that*..... or *Allegedly/Reportedly, John is ill again* (cp. Ramat, 1996)

Once we have stated via the semasiological procedure (the 'way up') that the opposition 'alienable ~ inalienable' is relevant for some languages (see example (1)), we can reverse our procedure and ask via an onomasiological approach (the 'way down') how language X behaves with regard to this opposition.

3. *Functions and forms (or onomasiology and semasiology)*

We have distinguished between universal semantic functions and their language-specific implementations. Likewise, we have now to distinguish between definition and implementation of categories, values and features. If a

phenomenon a (say, the opposition 'alienable ~ inalienable') is present in a language L_1, then there are no logical arguments to exclude that a may also appear in L_2, L_3, ...L_n. Contrary to a diffused opinion (Lazard, 1992, Haspelmath, 2007, among others), a categorial definition cannot be language-bound. As already said above, on the basis of a large sample of languages with their own linguistic structures we may arrive at a prototypical definition of a category; or state what a prototypical member of the category, e.g. ADJECTIVE, should look like. A quote from Coseriu may clarify what is meant here:

> ...si l'on définit universellement un adjectif, ceci ne signifie aucunement que l'on attribue l'adjectif à toutes les langues, puisqu'*une définition n'est pas un jugement d'existence*. (Coseriu, 1974: 49; my emphasis)

This means that one cannot foresee a priori the categories and their number a given language will have. A universal definition of ADJ does not entail a general theory of prototypical (morphological) categories. On the contrary, it is based on empirical observation.

Typology has shown that the roles and functions of 'Agent' and 'Predicate' may be expressed by means of very different strategies in different languages. On the other hand, to predicate something about someone or something (i.e. about entities or states of affairs conceived in our mind) does in fact belong to the basic cognitive activities of our brain. Therefore, there is an almost general agreement to the effect that NOUN (N) and VERB (V) are the most basic grammatical categories, although, as we have seen above, there is a continuum between the two categories[5]. Similarly, we may ask *a priori* whether PERSONAL PRONOUN (PPROs) constitute a universal category. Many languages make use of affixed morphemes to express the notion of person, so that we can imagine a language that lacks this grammatical category. Nonetheless, the necessity of distinguishing between "me", "you" and "she, he (it)"[6] is certainly primary in all languages. Once again, we differen-

[5] See DRESSLER (2008: 115 f.); RAMAT (2009). Note that I am not speaking of grammatical/syntactic roles, such as Nominative or Accusative; moreover, N+V does not mean that N must be the grammatical SUBJECT; in fact, we have constructs like Lat. *Caesarem*$_{(ACC)}$ *pudet alicuius rei*, where the NOUN of the Experiencer is in the Accusative, versus English *Caesar*$_{(NOM.SUBJ)}$ *is ashamed of something* or Italian *Cesare*$_{(NOM.SUBJ)}$ *si*$_{(REFL)}$ *vergogna di qualcosa*, where Caesar is in the Nominative as grammatical subject.

[6] JACQUESSON (2008: 89) observes that "me" and "you" are not true pronouns as they do not substitute a noun, like the 3rd person "(s)he, it" (which is actually a demonstrative) and their referent constantly changes according to the conversational roles. This is correct, but does not change the terms of the present discussion.

tiate between language specific implementations and general pre-linguistic concepts, which refer to mental operations of the cognitive domain.

4. *Cross-linguistic differences*

The differences among languages, though large, are not unlimited. Heine (1997: 47 ff.) has shown that the strategies for expressing possession are limited in number: for instance there exists, the Action Schema 'X takes/holds Y' meaning "X has, owns Y" as in Port. *O menino tem fome* "the child is hungry" (lit. "takes hunger": abstract 'possession'), *Temos outros vestidos* "We have other dresses" (concrete possession); or the Location Schema 'Y is at X's place' meaning "X has, owns Y" as in Turk. *Ben-de araba / otomobil var,* lit. "Me-LOC car is", i.e. "I have a/the car"; etc.

Speaking of cognitive contents and cross-linguistic differences, even 'relativist linguists' such as Lazard, sceptical about the very existence of universal categories, admit that there exist invariant notions in all languages (Lazard, 1992: 431): "to be hungry" is one of these notions, whose linguistic implementations can be different according to the language. Moreover, the same language may have more than a single strategy: Portuguese may use the comitative schema ('X is with Y'), *O menino está com fome,* along with the already mentioned *O menino tem fome*. Both constructions refer to the same mental, cognitive map consisting of semantic functions (see Croft, 2001) and the Portuguese example shows that one and the same 'slot' of a cognitive map may include different linguistic strategies (synonymic expressions) in the very same language.

5. *Word classes and constructions*

Hengeveld (1992: § 4.5.1.) distinguishes between 'specialized' and 'non-specialized' languages: the former are those in which every lexeme gets a particular function with its own morphological markers. Latin, with its formal distinction between VERB, NOUN, (P)PRO, ADJ, and ADV, belongs to the specialized, i.e. differentiating type. The more specialized a lexical class, the less it is necessary to mark this class by means of syntax (or morphology). For instance, English *-ly*-adverbs can occur in (almost) every linear ordering since they are unmistakably marked as adverbs (with very few exceptions, such as *lovely*):

(10) a. *Unfortunately she didn't arrive in time*
b. *She, unfortunately, didn't arrive in time*
c. *She didn't arrive in time, unfortunately*[7]

Adverbs not marked via the *-ly* suffix are more bound to the sentence structure.

According to these criteria Samoan is a 'non-specialized flexible' language[8]. On the contrary, German or Russian are 'specialized languages' as VB, N, ADJ and – partially – ADV have specific markers. As a matter of fact, the same lexeme may belong in Samoan to different word classes: VB, N, ADJ and Manner ADV – which actually means that it does not belong to any class. Similarly, in Tongan the function is assigned to a lexeme only by its syntactic frame. A lexeme *per se* does not belong to any word class. If it appears in a TAM-phrase it has predicative function, but if it appears in an ART-phrase it has referential function (see Vogel, 2000: 264-276). The previously quoted English *-ing* forms «are 'underspecified' for 'word classes' and their specification takes place on the syntactic level by phrase markers, e.g., articles» (Vogel, 2000: 274).

We are faced here with the well-known phenomenon of 'transcategorization', as in the English example *bar* quoted above in the abstract (see Ježek and Ramat, 1999). Forms such as Engl. *bar, except*, Ital. *tranne* "except", etc. separated from the verbal paradigm they belonged to, became autonomous items, frozen forms now registered in the lexicon as lexemes. In this sense, they can be considered less grammatical, though they still belong to a closed class of lexemes (namely ADPositions) that serve grammatical functions. In other cases the transcategorization concerns not just a single morpheme but a whole construction. This is the case of prepositional phrases as *instead of* studied by Schwenter and Traugott (summarized in Traugott, 2003: 636-638).

[7] Of course, the pragmatic effect of the three sentences is different, depending on the focussing strategies of the adverb.

[8] Non-specialized languages divide into two major subgroups: 'flexible' and 'rigid' languages. Flexible languages are those in which a single PoS may be used in different functions; rigid languages are those in which for certain functions a single PoS may be lacking and which combine different functions in one and the same lexeme form.

6. *Conclusions*

I have already alluded to prototypical forms, such as *amabantur* or *görüp*, which show *values* and FEATURES typical to verbs. According to Frajzyngier (2004):

> the use of an element in a function other than the one that is its inherent property requires the use of other coding means, e.g. derivational morphology, linear order, adpositions, articles, or phonological means such as tone and intonation. (Frajzyngier, 2004: 264)

On the contrary, Radical Construction Grammar denies the existence of inherent properties which may define categories:

> Grammatical categories must be defined relative to specific constructions. [...] The categories are defined by the construction [...] not the other way around [and] there are categories for each construction and each constructional role in a language. (Croft, 2007: 409, 419, 421)

Radical Construction Grammar is too radical, though, in that it denies the existence of properties inherent to lexemes. But if we come back to the examples I have used in this paper, it is hard to deny that, for instance, *amare,* with its suffix *-are* which is characteristic of infinitival forms, is *per se* a verb. However, in the sentence *amare humanum est* "love is a human characteristic / property" the infinitive controls the agreement of the predicate (*humanum*$_{\text{NTR}}$ and not *humanus*$_{\text{MASC}}$ or *humana*$_{\text{FEM}}$) and consequently has the syntactic property of being the subject of this predicative sentence. The syntactic function and, at least partially, also the meaning of *amare* depend on the syntactic structure it enters. The conclusion is that both positions, the onomasiological functionalist on one side and the grammatical formalist (i.e., semasiological) on the other, have to be considered when dealing with word classes. A correct balance between the two approaches is needed:

> Grammatical categories are usually distinguished from grammatical relations insofar as the latter, but not the former, are defined in terms of the role played by particular linguistic elements vis-a-vis other elements or the clause as a whole. For example, categories such as verb are defined in terms of inherent properties of lexical roots, for example, ability to take tense inflection, while relations such as subject or object are defined in terms of the relationship between particular verb arguments and the verb or the clause as a whole. (Cristofaro, 2008: 476)

Acknowledgments

This article is the revised version of a lecture given at the *Università di Roma Tre* on the occasion of a workshop on "Word classes: Nature, typology, computational representations" (March 24-26, 2010). I am grateful to the organizer of the workshop, Raffaele Simone, and to Luca Alfieri for many insightful suggestions.

References

Aikhenvald, A. (2004), *Evidentiality*, Oxford University Press, Oxford.

Anward, J., Moravcsik, E. and Stassen, L. (1977), *Parts of speech: A challenge for typology*, in «Linguistic Typology», 1, pp. 167-183.

Coseriu, E. (1974), *Les Universaux linguistiques (et les autres)*, in Heilmann, L. (1974, *ed.*), *Proceedings of the 11th International Congress of Linguists*, Il Mulino, Bologna, vol. I, pp. 47-73.

Coşkun, H. (2010), *Embedding indirective (evidential) utterances in Turkish*, in Diewald, G. and Smirnova, E. (2010, *eds.*), *Linguistic realization of evidentiality in European languages*, Mouton De Gruyter, Berlin-New York, pp. 187-222.

Cristofaro, S. (2009), *Grammatical categories and relations: universality vs. language-specificity and construction specificity*, in «Language and Linguistics Compass», 3, 1, pp. 441-479.

Croft, W. (2001), *Radical Construction Grammar. Syntactic theory in typological perspective,* Oxford University Press, Oxford.

Croft, W. (2007), *Beyond Aristotle and gradience*, in «Studies in Language», 31, pp. 409-430.

Dressler, W.U. (2008), *Tendenze tipologiche nella struttura e nella acquisizione della flessione verbale*, in Bertinetto, P.M. *et al*. (2008, *eds.*), *Diacronia, teoria, tipologia. Atti del XXXI Convegno della Società Italiana di Glottolologia*, Il Calamo, Roma, pp. 115-132.

Frajzyngier, Z. (1997), *Is the external possessor construction a simulant?*, paper presented at the *Conference on External Possession and Noun Incorporation, Eugene (OR), September 1997.*

Frajzyngier, Z. (2004), *The principle of Functional Transparency in language structure and in language evolution*, in Frajzyngier, Z., Rood, D. and Hodges, H. (2004, *eds.*), *Linguistic Diversity and Language Theories*, Benjamins, Amsterdam-Philadelphia, pp. 259-283.

Givón, T. (1979), *On understanding grammar*, Academic Press, New York.

Grammatica = Fray Domingo de Santo Tomas ([1560[1]] 1994, *Grammatica o Arte de la lengua general de los Indios de los reynos del Peru*, Ediciones de Cultura Hispánica, Madrid.

Haspelmath, M. (2007), *Pre-established categories don't exist: Consequences for language description and typology*, in «Linguistic Typology», 11, pp. 119-132.

Heine, B. (1997), *Possession,* Cambridge University Press, Cambridge.

Hengeveld, K. (1992), *Non-verbal predication*: *Theory, typology, diachrony*, Mouton de Gruyter, Berlin-New York.

Jacquesson, F. (2008), *Les personnes. Morphosyntaxe et Sémantique*, Éditions du CNRS, Paris.

Ježek, E. and Ramat, P. (2009), *On parts-of-speech transcategorization*, in «Folia Linguistica», 43, pp. 391-416.

Lazard, G. (1992), *Y a-t-il des catégories interlangagières*, in Anschüz, S.R. (1992, *ed.*), *Texte, Sätze, Wörter und Moneme. Festschrift für Klaus Heger zum 65. Geburtstag*, Heidelberger Orientverlag, Heidelberg, pp. 427-434.

Pullum, G. (1994), *Categories, linguistic*, in Asher, R.E. (1994, *editor-in-chief*), *The Encyclopaedia of Language and Linguistics,* vol. 5, Pergamon Press, Oxford.

Ramat, P. (1996), *Allegedly, John is ill again: stratégies pour le médiatif*, in Guentchéva, Z. (1996, éd.), *L'énonciation médiatisée,* Peeters, Louvain-Paris, pp. 287-298.

Ramat, P. (1999), *Linguistic categories and linguists' categorizations,* in «Linguistics», 37, pp. 157-180.

Ramat, P. (2002). *La natura dell'infinito*, in Jansen, H. *et al*. (2002, *eds*.), *L'infinito & oltre. Omaggio a Gunver Skytte,* Odense University Press, Odense, pp. 409-417.

Ramat, P. (2009), *How universal are linguistic categories?*, in Scalise, S., Magni, E. and Bisetto, A. (2009, *eds*.), *Universals of Language Today*, Springer, Heidelberg, pp. 1-11.

Simone, R. (2008), *Coefficienti verbali nei nomi. Categorie del verbo,* in Bertinetto, P.M. *et al*. (2008, *eds*.), *Diacronia, teoria, tipologia. Atti del XXXI Convegno della Società Italiana di Glottologia*, Il Calamo, Roma, pp. 83-113.

Traugott, E. (2003), *Constructions in grammaticalization*, in Brian, D.J. and Janda, R.D. (2003, *eds*.), *The Handbook of Historical Linguistics,* Blackwell, London, pp. 624-647.

van der Auwera, J. (1998), *Defining Converbs*, in Kulikov, L. and Vater, H. (1998, *eds.*), *Typology of verbal categories. Papers presented to Vladimir Nedjalkov on the occasion of his 70th birthday,* Niemeyer, Tübingen, pp. 273-282.

Vogel, P.M. (2000), *Grammaticalisation and part-of-speech systems*, in Vogel, P.M. and Comrie, B. (2000, *eds.*), pp. 259-284.

Vogel, P.M. and Comrie, B. (2000, *eds.*), *Approaches to the typology of word classes,* Mouton de Gruyter, Berlin-New York.

Paolo Ramat
Piazzetta Arduino 11
27100 Pavia (Italy)
paoram@unipv.it

NORME PER GLI AUTORI

Le proposte editoriali (articoli, discussioni e recensioni), redatte in italiano o inglese o altra lingua europea di ampia diffusione, vanno inviate preferibilmente tramite il sistema *Open Journal System* (OJS) collegandosi al sito *http://www.studiesaggilinguistici.it* (ove sono indicate le procedure da seguire) oppure, in via eccezionale, per posta elettronica a Giovanna Marotta (mail: *gmarotta@ling.unipi.it*), utilizzando due formati: un file pdf anonimo ed un file word completo di tutti i dati dell'Autore (indirizzo istituzionale e/o privato, numero telefonico ed e-mail).

Nella redazione della proposta editoriale, gli Autori sono invitati ad attenersi scrupolosamente alle norme redazionali della rivista.

Le proposte di articoli e discussioni dovranno essere corredate da un breve riassunto anonimo in lingua inglese, della lunghezza di circa 15 righe o 1.000 battute (spazi inclusi) e da 3 o 4 parole-chiave che individuino dominio e tema dell'articolo.

I contributi saranno sottoposti alla lettura critica di due *referees* anonimi, e quindi all'approvazione del Comitato Scientifico.

Il contributo accettato per la pubblicazione e redatto in forma definitiva andrà inviato tramite OJS nei tempi indicati dal sistema, e alla Segreteria di Redazione per posta elettronica (Dott.ssa Maria Napoli: *maria.napoli@lett.unipmn.it* oppure Dott. Francesco Rovai: *francesco.rovai@unipi.it*), sia in formato word che pdf, includendo i font speciali dei caratteri utilizzati.

Edizioni ETS
Piazza Carrara, 16-19, I-56126 Pisa
info@edizioniets.com - www.edizioniets.com
Finito di stampare nel mese di dicembre 2014