To appear in: J. Blochowiak, C. Grisot, S. Durrleman-Tame and C. Laenzlinger (eds.) *Formal models in the study of language,* Berlin: Springer.

BUILDING AND INTERPRETING AD HOC CATEGORIES: A LINGUISTIC ANALYSIS [*]

Caterina Mauri (caterina.mauri@unibo.it)

The aim of this paper is to examine in a systematic way the linguistic expression of a particular type of categorization process, namely the construction of *ad hoc* categories. Based on a 60 language-sample and corpus data from English and Italian, it will be shown that the strategies used to refer to *ad hoc* categories are mobilized from a variety of different grammatical areas, ranging from connectives to special plural forms and derivational affixes. We will first provide a detailed semantic analysis of the constructions under exam, and then move to the examination of the morphosyntactic and functional patterns of variation attested in our data. Though highly differentiated, the pool of strategies employed to make reference to ad hoc categories shows systematic correlations between specific morphosyntactic features, different degrees of context dependency and different types of abstraction processes (e.g. leading to the construction of a set, a frame or a class). We will conclude with a preliminary analysis of how *ad hoc* categories are built and used in discourse. Corpus data will lead us to propose a shift of attention from ad hoc categories themselves to *on line* categorization, namely the process through which categories are abstracted from specific exemplars in context, regardless of their common or *ad hoc* nature.

1. INTRODUCTION

The main purpose of this paper is to provide a linguistic analysis of how so-called *ad hoc* categories are conveyed in discourse. The construction of ad hoc categories is a basic cognitive and communicative process, which, despite its close connection with verbalization, has not been systematically investigated from a linguistic point of view. The discussion will be based on data from a 60-language sample, integrated with corpus data of Italian and English.[1]

The identification of ad hoc categories is traced back to the work of the psychologist Lawrence Barsalou (1983, 1991, 2003, 2010), who coined the term and conducted several experiments aimed at unfolding the internal structure of these categories. According to Barsalou (2010: 86), ad hoc categories are novel categories constructed spontaneously to achieve goals relevant in the current situation (e.g., constructing [TOURIST ACTIVITIES TO PERFORM IN BEIJING] while planning a vacation). They are constructed spontaneously because they do not reside as knowledge structures in long-term memory. Ad hoc categories are contrasted with numerous well-established categories associated with familiar words (e.g., [CAT], [EAT], [HAPPY]). Barsalou (1983) showed that ad hoc categories are highly context-dependent, but once constructed, they function as coherent categories with internal structures

[*] I would like to thank Andrea Sansò, Mira Ariel, Alessandra Barotto and Cristina Lo Baido for crucial discussions on data and their interpretation, and Giorgio Francesco Arcodia for helping me with the glosses of Chinese and Japanese examples.

[1] The research described in this paper has been possible thanks to a Grant provided by the Italian Ministry of Research and University (MIUR) within the SIR program (SIR project *«LEAdhoC : The linguistic expression of ad hoc categories »*, prot. RBSI14IIG0, coordinated by Caterina Mauri.

and typicality gradience, much like stable categories. In his most recent paper on this topic (2010: 87), Barsalou argues that "much further study is needed to understand the role of ad hoc categories in cognition" and an important issue to explore is "how productive conceptual and linguistic mechanisms produce ad hoc categories".

However, while in psychology a number of further experimental studies has followed Barsalou's theory of ad hoc categories, his findings had only a small echo in linguistics. The aim of this paper is to show that the linguistic expression of this special type of categories, whose main distinguishing property is the dependence on context for their construction, deserves attention, because it may shed new light on the communicative role of such categories and on their functional properties, thus providing insights also to psychological research in this domain.

This paper is structured as follows. In section 2 the notion of ad hoc category is defined in purely semantic terms (2.1), discussing its relations to context and to vagueness (2.2). Some methodological remarks concerning the language sample and the identification of the relevant constructions are made in section 2.3. Section 3 is devoted to the discussion of cross-linguistic data: first the attested strategies are exemplified and discussed in detail (3.1), then we move to the observed patterns of variation (3.2), showing non-random correlations between specific morphosyntactic properties and specific ways of abstracting the categories. Section 4 is devoted to the analysis of ad hoc categories in discourse, in order to discuss the reasons and the communicative aims underlying this specific categorization process. Finally, section 5 contains some conclusive remarks and the prospects for future research.

2. AD HOC CATEGORIES: DEFINITIONS AND METHODOLOGY

2.1 Ad hoc categories and associative reasoning

Many of our cognitive categories are stable, others are ad hoc. Crucially, ad hoc categories are context-dependent and people construct them to achieve their communicative goals. For example, constructing the category "activities one can perform on a sunny Sunday afternoon" can be instrumental to achieving the goal of inviting a friend to spend the Sunday afternoon together. Cognitive psychologists (Barsalou 1983 and onwards) have shown that, for ad hoc categories, the category concepts, concept-to-instance associations, and instance-to-concept associations are much less established in memory than for common categories (e.g., [FRUIT], [FURNITURE]). In addition to these differences, however, ad hoc categories have been shown to possess graded structures (i.e., typicality gradience) as salient as those structuring common categories. This appears to be the result of a similarity comparison process that imposes graded structure on any category regardless of type (Barsalou 1983).

Stable categories can typically be expressed by fairly short conventional linguistic means (e.g., *queen, eagle*, etc.). Ad hoc categories, instead, do not come with ready-made linguistic labels (words - e.g. *furniture, clothing* - or small phrases - e.g. *grocery stores, vegetarian food*), but are typically described by means of complex expressions (e.g., *tourist activities to perform in Rome*, *clothing to wear while house painting*, etc.). Their identification in discourse nonetheless crucially depends on verbalization, i.e. the linguistic strategies that speakers systematically employ to refer to the process of ad hoc category building, as exemplified in (1) where *and so on* leads the hearer to go on constructing the category [ACTIVITIES TO PERFORM IN ROME].

(1)*We are in Rome for the weekend. We have plenty of things to do, you know: [visit the Colosseum, stroll through the Gardens of the Villa Borghese, go to the Trevi fountain,* ***and so on…*** *] everything in two days! But we'd love to meet you for a coffee.*

Despite the central role played by verbalization in the process of ad hoc categories construction, little attention has been devoted to the linguistic constructions that are systematically employed to convey this process. Taylor (2003), in his study on linguistic categorization, does not even consider this type of categorization.

The notion of ad hoc categories has recently been employed in research on lexical pragmatics, within Relevance Theory (Wilson & Carston 2007, Carston 2010), accounting for why the meanings of words must be pragmatically adjusted and fine-tuned in context, so that their contribution to the proposition is different from their lexically encoded sense. They thus focus on the semantic processes of narrowing and broadening that underlie the normal interpretation of words. The Relevance Theoretic approach crucially shows how the interpretation of potentially every word is dependent on context, thus ultimately suggesting that every abstract category conveyed by lexicon is necessarily translated into a more concrete category, anchored in the situational context.

Though closely connected to Wilson and Carston's research, the process at issue in this study focuses on a different phenomenon: we are concerned with the *linguistic strategies* specifically encoding, i.e. overtly signaling, this process, rather than with the cognitive mechanisms underlying the interpretation of lexicon. Furthermore, the construction of ad hoc categories starts from the context and requires an abstraction over concrete exemplars, rather than going from an abstract category and looking for its actualization in the context.

Examples 2) and 3) show instances of ad hoc categorization. In 2), starting from *water*, *teas* and *smoothies* the hearer infers the higher-level category [HEALTHY DRINKS], while in (3) the same list of exemplars leads to abstract a different class of entities, namely [NON-STIMULATING DRINKS], on the basis of a different context. In 2), the context relevant aspect is health, while in 3) the beverage list is opposed to *coffee*, driving the hearer to focus on the absence of a stimulating effect.

2) *It is necessary to drink a lot of [water, herbal teas, smoothies, and the like] to be healthy.*

3) *[Water, herbal teas, smoothies, and the like] are useless in the morning. I need just coffee.*

Languages show specific strategies to achieve the abstraction process and anchor it to the context. These strategies are characterized by the explicit mention of one or more exemplars, used as a starting point to infer some high-order entity. Depending on the relationship between the exemplars, the inferential process can lead to the construction of a *set,* if the exemplars are elements that co-occur in combination (ex. 4)), to the construction of a *class*, if the exemplars are equivalent alternatives (ex. 5)), or to the construction of a *frame,* if the exemplars are actions that occur within a narrative scheme (ex. 6)).

(4) Set [ingredients of a recipe]:
I need flour, milk, yeast and so on.

(5) Class [quiet activities that may appeal to a 6 year old boy]:
You can read a book, make a drawing or something.

(6) Narrative frame [actions that happen in the restaurant]:
You order, wait for food, urge the waiter because you are hungry, then wait again and so on.

The context that needs to be accessed in order to properly construe the inferential process can be of different types. Following the taxonomy proposed by Croft and Cruse (2004, pp. 102-103), we can distinguish between *linguistic context*, which includes preceding discourse, immediately adjacent co-text, and the type of speech, *physical context*, which includes elements selected on the basis of perception, *social context*, including the relationships between the interlocutors, and *encyclopedic knowledge*, including information related to the speaker, the listener, to their background and habits, etc.

Despite variation in the type of categorization (leading to build a set, a class or a frame) and in the type of context necessary to construe the abstraction, we can define a semantic core that invariably characterizes the linguistic strategies employed to convey ad hoc categorization. They indeed systematically make reference to three orders of elements:

a) one or more *explicit exemplars* of the category;
b) additional *implicit members X*, associated with the exemplars by virtue of a shared property P that is relevant to the context;
c) a *superordinate category C*, which includes both explicit exemplars and additional members X.

Let's see an example from spoken English, taken from the BNC:

7) [..] *a lump sum to be held by the town council, to be used as a form of grant, or financial support for [low-income families, students, unemployed, etcetera], on production of the relevant proof, erm, Depending on the individual's needs*, [...]. (BNC, hyjS_meeting)

The speaker who utters (7) refers to:

a) three explicit exemplars: *low-income families, students, unemployed;*
b) additional implicit members X, sharing with the exemplars the property P 'in need of financial support', identifiable by drawing on the linguistic context and encyclopedic knowledge;
c) a superordinate category C [PEOPLE WHO ARE IN NEED OF FINANCIAL SUPPORT], which includes explicit exemplars and implicit members X.

In order to think about further members X and thus abstract the class of persons to which the speaker is making reference, the property P 'in need of financial support' has to be inferred from the context in a sufficiently non-ambiguous way. In case the inferential processes leading to the identification of property P cannot be correctly set up, the utterance in 7) could not be properly interpreted.

Let's see in detail what are the inferential processes that enable to abstract sets, classes and frames starting from exemplars. First of all, the use of a linguistic expression such as *etcetera* (cf. also *and so on, or something* in examples 4) and 5)) induces to recognize the existence of additional exemplars X besides the explicit ones, guiding the hearer to mentally open an empty folder, where such further items can be 'saved'. A linguistic expression like *etcetera* therefore has a clear referential function and indicates the presence of a variable X, whose identity has to be saturated on the basis of the specific context. We can then define the first inferential process activated by this type of constructions *saturation* of a variable X, not identified but to which explicit reference is made.

To fulfill the saturation process and assign a value to the variable X, you need a second inferential process, namely associative reasoning (cf. Recanati, 2004). The members X to

which reference is made must in fact be associated, or associable, to the explicit exemplars on the basis of some shared properties. This property can be intrinsic, that is, it may define the nature or function of the elements themselves (e.g. financial need in 7), the absence of harmful substances in 2)), and in this case associative reasoning leads to the identification of a relationship of *similarity* (cf. Joosten 2010: 32); or it can be extrinsic, based on the contiguity of elements recurring within specific schemas (eg. co-occurrence in the schema 'baking a cake' in 4), or contiguity in the schema 'eating at the restaurant' in 6)), and in this case the elements are associated with each other by virtue of their contiguity with a frame.

Crucial to associative reasoning is therefore the identification of the specific property P that is relevant in a given context. As we noted in examples 2) and 3), it is indeed the case that different contexts lead to select different properties for the same explicit exemplars: 'healthy' is identified as the relevant property in 2), therefore Sprite is not a possible value of X (cf. 8)); 'devoid of stimulating substances' is instead the relevant property in 3), and in this case Sprite becomes a possible value of X (cf. 9)).

8) *It is necessary to drink a lot of [water, herbal teas, smoothies, and the like] to be healthy.*
X = orange juice
No! X = Sprite
Class [healthy drinks]

9) *[Water, herbal teas, smoothies, and the like] are useless in the morning. I need just coffee.*
X = orange juice
X = Sprite
Class [non stimulating drinks]

Since the specific property P, detected through associative reasoning, determines the inclusion or exclusion of certain values of X, it also has direct consequences on the result of the *abstraction* process, which constitutes the third inferential process involved in the construction of ad hoc categories. The process of abstraction indeed leads to build a category C that includes both the explicit exemplars and the possible values of X.

To sum up, we identify three inferential processes involved in the exemplar-driven construction of categories: saturation, associative reasoning and abstraction. Linguistic strategies that realize this function are characterized by a common semantics, which consists of reference to: *i*) one or more explicit exemplars, *ii)* additional implicit elements X, associated to the exemplars by virtue of a shared, contextually relevant property P, *iii*) a category C that includes both explicit exemplars and the implicit elements X.

2.2. Context dependence and vagueness

The distinguishing property of ad hoc categories is the dependence on context for their construction and interpretation, which is not shared by common categories, whose denotation is stable enough to include a core of exemplars inferable without accessing to the speech situation. Given the centrality of context in the definition of the object of analysis, it is worth dwelling on the various respects under which the construction of ad hoc categories may be analyzed as being an intrinsically pragmatic phenomenon.

First of all, it lies at the intersubjective discourse level, where the speaker guides the hearer in the interpretation of the mentioned exemplar(s) mainly as arrows to the category, rather than as bearing an independent (and discourse relevant) reference. To do this, the speaker uses overt, dedicated strategies, which are the object of our research and can be analyzed as

bearing a clearly *procedural* value. Second, the associative inferences through which the category may be constructed are necessarily anchored in and dependent on the specific speech situation, including knowledge relative to the interlocutors, to the temporal and spatial conditions of the speech event, and to the shared background. The hearer has to have access to such contextual information in order to identify the correct associations that lead to inferring further potential members of the category.

The pragmatic component is inserted into a more basic cognitive function, such as the ability, which is at the same time a necessity, to categorize the world, i.e. to consider jointly and aggregate into sets entities and situations that have something in common. Only, the reasons to collect explicit and implicit exemplars into the same set are not necessarily semantic in nature, nor connected to their frequent association in everyday life, but rather depend on highly specific and contextual goals.

Does context dependence equate to vagueness? The discussion on the inferential processes necessary for the interpretation of ad hoc categories has shown that there is at least one element that should not be vague, or better, should be sufficiently unambiguous to allow associative reasoning: such element is the property P. The identification of the relevant property P is necessary in order to bring about associative reasoning and, ultimately, abstraction. Once the property P is detected, in fact, the speaker is able to assess the inclusion or exclusion of possible values for X, constructing the category. What property has to be selected as relevant is thus dependent on the specific context, but, given the specific context, the identification of the property should not be vague.

However, this does not hold for all the possible values of X. A sufficiently unambiguous identification of the property P may be compatible with non specific, vague values of X. In other words, the category can be inferred correctly even if the values of X are undefined, and if the process of saturation does not lead to identify each and every possible additional member that could be associated to the explicit exemplars in a relevant way.

Does this lead an inherent vagueness of ad hoc categories? Voghera (2012: 354-358) discusses general extenders such as *and so on* within what she calls 'intentionally vague categories'. Also Channell (1994) and Overstreet (1999) analyze these linguistic strategies in relation to their vague semantics - Channell even calls them ‘vague category identifiers’. The analysis conducted so far may help us to describe in a more accurate way to what extent we can ascribe ad hoc categories to vagueness.

It is true that you cannot predict in advance which items will become part of the category, and this leads to vagueness. Furthermore, the additional elements to which reference is made, that is, the possible values of X, can remain non-specific - it is indeed not necessary to produce the exhaustive list of all the potential members of the category - and this produces vagueness. However, in both these cases vagueness concerns the identity of the members of the category, not the identity of the category as such.

The identity of the category is guaranteed by the contextually relevant property P, whose identification allows to discriminate between potential members and elements that should be excluded. In other words, a sufficiently unambiguous identification of the property P allows to delimit the borders of the category, and ultimately to identify in a sufficiently unambiguous way the category itself.

To sum up, the identity of the members of the category may be vague, but the category as such must be defined unambiguously, on the basis of the property P relevant to the specific context. If P cannot be identified, as in 10), where the explicit exemplars cannot be traced back to a common property or a common pattern, then not only we observe vagueness in the identity of additional elements X, but it becomes impossible to identify the category.

10) ??? *It is necessary to buy [water, a deer, three pencils and the like] to be safe.*

2.3 Methodology for cross-linguistic analysis

In order to analyze how languages encode the cognitive and communicative process described in the preceding sections, we will first take a cross-linguistic perspective, which will then be complemented by a glance into intra-linguistic variation, as it is attested in corpus data.

Cross-linguistic data constitute the core of the analysis, while the discourse perspective is, for the purpose of this paper, limited to the discussion of some relevant examples taken from Italian (Corpus of spoken Italian LIP '*Lessico dell'italiano parlato*') and English (spoken section of the *British National Corpus*). The language sampling procedure aims to be compliant with current standards in typological research. Due to the highly discursive nature of the phenomenon, descriptive grammars have been integrated by a metalinguistic questionnaire for language experts and, wherever possible, by the analysis of naturally occurring texts. At this preliminary stage, the study is based on a convenience sample consisting of 60 languages.[2]

Due to their heterogeneous nature, linguistic strategies encoding ad hoc categories are not generally addressed in specific chapters of grammatical descriptions. This limitation, however, does not prevent the cross-linguistic identification of the relevant phenomena: modern comprehensive grammars often contain a more or less large corpus of (mostly oral) texts, in which the relevant structures can be identified; moreover, the existing literature, though episodic, provides hints as to the areas of grammars that are most likely to provide means for encoding ad hoc categories (number, morphological processes such as reduplication, connectives, discourse particles, etc.). In the absence of significant data (but also in order to complement limited data sets) resort to language experts has been essential to the aims of this research.

3. THE LINGUISTIC EXPRESSION OF AD HOC CATEGORIES

Being the internal structure of ad hoc categories based on typicality gradience (Barsalou 1983), the linguistic strategies expressing them often involve the explicit naming of one or more exemplars, that the addressee processes as pointers for conjuring up an ad hoc category (cf. also Ariel and Mauri 2016).

A cross-linguistic perspective shows great variation in the types of strategies that may be used with this function. Constructions encoding ad hoc categories indeed range from more transparent discourse-level strategies such as English *or stuff like that*, to synthetic, less transparent means such as the Japanese non-exhaustive connective *-ya*, dedicated plurals (so-called associative and similative plurals, see Daniel 2000 and Moravcsik 2003), derivational affixes or special types of reduplication (e.g. Turkish *m*-reduplication). Morphosyntactic

[2] The selection of languages has been driven by two criteria: genealogic diversity and the availability of descriptive grammars containing relevant information for our study. The following languages have been selected (in alphabetical order): Cavinena, Central Pomo, Chemeuevi, Comanche, Cupeno, Czech, Classical Japanese, Darma, Diu, Diyrbal, Dogon, Dom, Dutch, English, Finnish, French, Galo, Georgian, German, Haitian Creole, Hakha Lai, Hausa, Hebrew, Hindi, Hungarian, Hup, Italian, Jamsay, Japanese, Kannada, Koasati, Kokota, Kuuk Thayorre, Lao, Lavukaleve, Lezgian, Malayalam, Makalero, Mandarin Chinese, Mani, Maori, Marthutunira, Mongsen Ao, Nenets, Nootkan, Paumarì, Russian, Scolt Saami, Skou, South Efate, Spanish, Ughele, Tagalog, Tahitian, Tamil, Tokelauan, Tongan, Turkish, West Greenlandic, Yurakare.

variation is the object of sections 3.1 and 3.2 below. In section 4 it will be argued that some variation can also be observed in the functions of ad hoc categorization in discourse.

3.1. Cross-linguistic variation

3.1.1 *Special plurals*

There are languages where the construction of ad hoc categories may be achieved through a morphological strategy that has been classified in the literature as being a special type of plural. Corbett (2000: 101-111) and Daniel (2000) name this special form *associative plural*, when it applies to animate referents, and *similative plural*, when it applies to inanimate referents. Great terminological variation can be observed in the (rather poor) literature: elliptical dual (Delbrück 1893: 137), approximative plural (Jespersen 1965: 192), *representativnaja množestvennost'* (representative plural). Daniel and Moravcsik (2005) argue that "associative plural constructions consist of a noun X (typically of human reference, usually a person's name or a kin term) and some other material, most often an affix, a clitic, or a word. The meaning of the construction is **'**X and other people associated with X'." Associative plural constructions are identified by two semantic properties: *referential heterogeneity* and *reference to groups*.

Referential heterogeneity distinguishes between associative plurals and additive plurals. An example of additive plural is *girls*: it denotes a set where every member is a girl and therefore the set is argued to be referentially homogeneous (every referent of the plural is also a referent of the stem). An example of associative plural is Japanese *Tanaka-tachi*: it denotes a heterogeneous set, because it does not refer to more than one person named Tanaka, but rather to a group of people only one of whom is named Tanaka. According to Daniel e Moravcsik (2005), referential heterogeneity is a property that characterizes also the so-called *similative plural* (e.g. Telugu *puligili* 'tigers and such'), which denotes "a class of objects sharing similar features, rather than a group of closely related associates". The second semantic property characterizing associative plural is reference to groups, in that it denotes sets with a clear internal cohesion, which can be described as groups of individuals. Under this respect, associative and similative plurals are connected to *collective plurals* (see also section 3.1.2 on derivation), which are however usually referentially homogeneous.

From a morphosyntactic point of view, Daniel and Moravcsik (2005) analyzed as associative plurals dedicated affixes, clitics, determiners, plural markers and pronouns, connectives and constructions with plural verb forms. Such a formal variety questions whether the very label of 'plural' is appropriate and suggests that a re-examination of the constructions classified as associative plurals may ultimately lead to the identification of a larger class of strategies employed for ad hoc categories. Let us now provide some examples and than focus on the relation between associative plurals and ad hoc categories.

11) Hungarian
   a. *János-**ék***
      János-**ASSOC**
      'János and the others'
   b. *János-ok*
      János-PL
      'more than one person called János'

12) Dogon (Niger-Congo, Corbett 2000: 111)

ibɛ *ya-ɛ-w* *yo,* *isu* ***mbe*** *nie* ***mbe*** *bawiɛ*
market go-AOR-2SG if fish PL oil PL buy.IMP.2SG
'if you go to the market, by fish, oil and other such things.'

In examples 11) and 12) we can see a case of associative plural and a case of similative plural, respectively. The suffix *–ék* in Hungarian may only follow animate nouns, preferably proper nouns, while the plural marker *mbe* in Dogon may follow inanimate objects. It has to be noticed that in 12), reference to 'other such things' is only conveyed through the repeated plural marker *mbe* and there is no analytic strategy comparable to the English one. In example 13) from Diu, the marker *tud* has the function of additive plural if it occurs before the noun it refers to (13a), while it has a similative plural value if it occurs in post-posed position (13b):

13) Diu (Indo-Portuguese, Cardoso 2009)
a. *mĩ* ***tud*** *amig*
1SG.POSS PL friend
'my friends'
b. *el t-iŋ* *vay* *nə* *ũ* *jungle pu* *traz-e* *koys, aros* ***tud***
3S IPFV-PST go.INF LOC one jungle PURP bring-INF thing rice **SIML**
'He went to the jungle to bring certain things, rice and so on.'

Examples 12) and 13b) denote highly similar situations: in both cases there is a subject that reaches a place (market or jungle) where he can find a specific type of objects. In order to interpret the similative plural construction, and thus to identify the type of objects the speaker has in mind, it is necessary to use the mentioned exemplars (fish and oil in (5), rice in (6b)) as starting points for an inferential reasoning based on similarity. The context allows for the identification of further potential alternatives to the mentioned items, and thus to the construction of a goal-derived, ad hoc category.

The construction of an ad hoc category may appear less straightforward in 11), given the presence of a proper name: how can we construct a category taking a proper name as the main exemplar? For associative plurals it is indeed the case that more than a heterogeneous category, we are constructing a set of referents including the exemplar (*János*) and other persons having in common some relationship with the exemplar. Yet, the inferential processes leading to the correct identification of the set are the ones identified in section 2.1, namely saturation of further members X, associative reasoning based on a contextually determined relationship with the exemplar, and abstraction of a set comprising the exemplar and additional X.

Interestingly, the role played by the exemplar in associative plurals is slightly different than in similative plurals. In associative plurals the exemplar is the pivot of the set and can be also analyzed as being itself a 'property' shared by all the other members: *János* is a member of the set, but also what all the others have in common: they may be *János*' friends, relatives, colleagues, etc., what the other members share is that they entertain the same relationship with *János*. Daniel and Moravcsik (2005) call it the *focal referent*, as opposed to the *associates.* In similative plurals, on the other hand, the exemplar(s) (fish, oil and rice in the examples above) cannot be analyzed as a property shared by the other members, but simply as exemplars that the speaker considers sufficiently relevant to allow for the abstraction of an ad hoc category. As we will see, a similar opposition is also attested in certain types of derivational strategies.

3.1.2 *Derivational strategies*

Some of the strategies attested to convey the process of ad hoc categories construction can be analyzed as derivational. Let us start by looking at example 14) from Kuuk Thaayorre. As argued by Gaby (2006), "speakers may add the suffix *=yuk* 'STUFF' to a noun in order to speak in general terms about a 'kind of thing', or to generalise their reference to include things normally associated with the denotatum of the noun in question, or to indicate reference to type rather than token (i.e. 'not a specific bull' in (14d), and 'not a particular dance' in (14e)):

14) Kuuk Thaayorre (Australian, Pama-Nyungan)
a. *minh ulp ngancnhan reeka-rr, ngat=**yuk** reeka-rr*
meat(ACC) DEM:ADR.PRX 1sg:EXCL give-PL.PF fish(ACC)=STUFF give-PL.PF
ngancnhan
1sg:excl
'[they] gave us some meat and fish or whatever'
b. *kuta=**yuk** yuuw yat*
dog(NOM)=STUFF far go:PL.PF
'there were no people, even the dogs not there' (lit. 'dogs and stuff had gone off')
c. *pormpr=**yuk** koop thiik-nhan*
house(ACC)=STUFF all break-GO&:NPST
'all the houses and things will be broken [in a cyclone]'
d. *bull=**yuk** thaangk-m peln*
bull(ACC)- STUFF climb-PL.IPFV 3PL(NOM)
'they would ride bulls [in the rodeo]'
e. *ngancn wuuc=**yuk** thowol-nam ulp nhangun*
1sg(NOM) dance=STUFF perform-PL.PF DEM:ADR.PRX 3sgDAT
Jesus-ak
Jesus-DAT
'We were doing those dances for Jesus [at Christmas]'

Let us take a look at the diachrony of this morpheme. Etymologically, *=yuk* is derived from the generic noun *yuk* (Gaby 2006: 642), which denotes the class of trees and stick-like objects (e.g. cigarettes), but also a somewhat eclectic collection of (typically elongated) 'things' (including cyclones, planes, microphones, etc.). It has an intrinsically classificatory value, which may be the feature that favored its reinterpretation as a type/category marker, since the very notion of type and category is connected to the ability to classify.

A slightly different derivational strategy is attested in Italian, where the affix *–ame* may be attached to a variety of roots (inanimate, animate, see Poletto and Penello 2005, and even proper nouns, see examples 15) and 16) below) to derive collective nouns. When it follows proper nouns, the meaning of the collective is very close to that of associative plurals discussed above. Let us see some examples.[3]

15) a. *Dire che la Boldrini è uguale a Mastella, al figlio di Bossi o al **berlusconame** è una violenza ideologica che non porta da nessuna parte*
'to say that Boldrini is the same as Mastella, as Bossi's son or as ALL THOSE PERSONS HAVING TO DO WITH BERLUSCONI (INCLUDED BERLUSCONI HIMSELF) / BERLUSCONI &

[3] Examples (15), (16) and (17) do not have glosses, because glosses are not relevant for the discussion, centered around the interpretation of the derived forms.

CO. is an ideological violence that does not lead anywhere'
(http://forum.gamberorosso.it/viewtopic.php?f=6&t=155954&start=100)

b. *[penso] che la principale esigenza del paese sia in questo momento liberarsi politicamente di berlusconi e sradicare il* ***berlusconame*** *dalla società*
'[I think] that the most important need for the country in this moment is to get rid of Berlusconi at the political level and to eradicate THE WAY OF ACTING AND THINKING THAT BERLUSCONI INTRODUCED from the society'
(http://forum.bbfc.it/index.php?topic=4465.35;wap2)

16) *Io non sarei d'accordo, loro, il* ***grillame****, invece dovrebbero esserne molto felici: quando parlano dicono solo stupidaggini o cose insensate.*
'I wouldn't agree, while they, GRILLO & CO., should be very happy about it: when they speak they always say stupid things or nonsense.'
(http://www.ilgiornale.it/video/interni/grillino-sorial-attacca-napolitano-boia-986642.html)

17) *Io non ero abituata a tutto quel* ***bambiname****, e soprattutto non conoscevo nessun gioco*
'I was not used to all those CHILDREN AND SO ON (ALL THE THINGS CONNECTED TO CHILDREN), and above all I didn't know any games' (R. Calabrò *Di matrigna ce n'è una sola*)

In example 15), the derived noun *berluscon-ame* refers to two distinct types of sets. In 15a) it behaves as a collective noun denoting the set of 'people having to do with Berlusconi', functioning in a similar way as associative plurals. In 15b), instead, the collective noun denotes a heterogeneous set comprising persons, situations, attitudes that share Berlusconi as a common denominator. We could paraphrase 15b) as 'Berlusconi and everything that goes with it', meaning to include also parties, luxury dinners, corruption, a specific way of talking to people in TV, etc. In the two cases at issue, the noun Berlusconi plays two different roles: in 15a) it is what Daniel and Moravcsik (2005) called the focal referent for associative plurals, and the category it allows to construct is restricted to animate members; in 15b) Berlusconi has to be taken as an exemplar of a heterogeneous category including persons, situations and attitudes, but also as a property shared by all members of the set. In both cases, knowledge concerning Italian politics and access to the Italian context is required in order to identify the similative relations to infer, in order to correctly abstract the category.

Example 16) is similar to example 15a), in that B. Grillo, the leader of a political movement called Movimento Cinque Stelle, receives the suffix *–ame* in order to derive a collective noun meaning 'Grillo & co.', 'Grillo and associates', just like an associative plural. Example 17) shows instead a different case, where the root denotes an animate entity ('children') but not a proper noun. In this example *bambin-ame* is a collective noun means 'children and so on', referring to an ad hoc category including children, what children usually do, laughing, screaming, playing, generating disorder etc. In other words, it denotes a heterogeneous set of persons and situations having to do with children, whereby the specific types of associative inferences allowed are determined by the context.

The derivational strategy with *–ame* may have a pejorative reading, as in 15) and 16), but may also be neutral, as in 17). It may also apply to inanimate entities, as in *scatolame* 'set of boxes of various types', or *legname* 'set of pieces of wood of various dimensions'. It shows many parallels to associative and similative plurals, especially as far as the role played by the named exemplar is concerned.

3.1.3 *Reduplication*

A number of languages employ so-called *echo word formation* to construct ad hoc categories. It is a special type of reduplication in which the reduplicated segment does not completely coincide with the original word or phrase. Let us see an example from Turkish, where *m*-reduplication is commonly employed to generalize the concept denoted by a specific word or phrase in order to include similar objects, events or states of affairs. It is attested especially in colloquial Turkish. It may applied to animate, inanimate 18a) and proper nouns 18b), but also to phrases 18c):[4]

18) Turkish (Göksel and Kerslake 2005: 91-92)
   a. *Eve çat kapı bir alıcı geldi,* ***odaları modaları*** *dolaştı.*
      'Today a potential buyer came without notification, and looked at the ROOMS, ETC.'
   b. ***Soner'i monerí*** *görmedim.*
      'I didn't see SONER OR OTHERS.'
   c. *Ben adam* ***tarih hocasıymış marih hocasıymış*** *anlamam. Fransız tarihini ondan daha iyi biliyorum.*
      'I don't care if he is a HISTORY TEACHER OR ANYTHING ELSE. I know more about French history than him'

A similar construction is attested in Tamil and Hindi, where "the semantics of 'echo reduplication' involves a notion of generality or at least indefiniteness" (Keane 2005: 241). As can be observed in example 19), a category of possible nonsense utterances has to be construed in order to understand the speaker's aim. However, in order to imagine potential alternatives, the hearer has to have access to the context and to some shared background on what could be a possible lie in that specific situation.

19) Tamil (Dravidian)
   kumaarukku kuʈutteen-ŋŋu kimaarukku kuʈutteen-ŋŋu *poi kollaatee*
   Kumar.DAT give.PST.1SG-QUOT **echo** lie say.NEG.IMP
   'Don't lie that you gave it to Kumar OR SOME SUCH NONSENSE.'

According to Keane (2005: 240), reduplication is a common strategy to convey ad hoc categories in South East Asian languages. In Hindi (Indo-Aryan), for instance, from *aam* 'mango' it is possible to derive *aam vaam* to denote 'mango and similar fruit'. In Kannada from *kannu* 'eye' it is possible to derive *kannu ginnu* 'eyes and so on'.

Let us now examine the case of Lao (Tai-Kadai, Kam-Tai). According to Enfield (2007: 306), "echo formation typically involves four syllables […] The idea is to take a V-N sequence and repeat it, substituting the N in the repeated phrase with something semantically related (usually, a synonym or antonym). The result is that the first and third syllables are the same verb, and the second and fourth syllables are semantically related nouns." Example 20) is a simple statement in which we take the *song5* 'trousers' to refer to nothing other than its conventional referent, 'trousers'. The semantic effect of the echo formative in 21) is to put *song5* 'trousers' together with *sùa4* 'shirt', broadening the reference of both expressions, beyond the literal conjunctive reading 'trousers and shirts' and to a generalized notion of 'clothes', i.e., things of the kind that trousers and shirts are prototypical examples.

[4] Again, examples in (11) do not have glosses, because they were not available in the source.

20) *man2 pajø sùù4 song5*
3.B DIR.ABL buy trousers
'He (went and) bought trousers.'

21) *man2 pajø* ***sùù4 song5 sùù4 sùa4***
3.B DIR.ABL buy trousers buy shirt
'He (went and) bought CLOTHES (lit. TROUSERS AND SHIRT AND SO ON).'

There is also a generic, probably default echo-formative strategy, in which the complement element of the repeated phrase is replaced by the indefinite inanimate pronoun *ñang3* meaning 'something, what, whatever'. For example, example 22) above might be expressed as follows:

22) *man2 pajø* ***sùù4 song5 sùù4 ñang3***
3.B DIR.ABL buy trousers buy INDEF.INAN
'He (went and) bought TROUSERS AND SO FORTH.'

3.1.4 *Non-exhaustive connectives*

Another strategy that is attested across languages to convey ad hoc categories is what can be safely labeled *non-exhaustive connectives*. Haspelmath (2007: 24) briefly mentions this type of connectives and calls them 'representative conjunction'. According to him, "in this construction, the conjuncts are taken as representative examples of a potentially larger class". Stassen (2000: 5) calls them 'enumerative coordinators', while the label 'non-exhaustive' is well established in the literature on East Asian languages (Chino 2001, Zhang 2008). What distinguishes these connectives is the restriction to occur *only* in open-ended lists, i.e. non-exhaustive sets. In 23a) an example from Koasati shows the use of the non-exhaustive connective *-ó:t* to construct the ad hoc category 'humid places', starting from the two exemplars 'rivers' and 'swamps':

23) Koasati (Muskogean, Haspelmath 2007: 24)
a. *akkámmi-t ow-i:sá-hci hahci-f-**ó:t** oktaspi-f-**ó:t** kámmi-fa*
be.so-CONN LOC-dwell.PL-PROG river-in-**EX** swamp-in-**EX** be.so-in
'So they live in rivers and in swamps AND IN SUCHLIKE PLACES.'
b. *[asá:l-**o:t]** talibo:li-t sco:pa-t*
basket-**EX** make-CONN sell-CONN
'She made and sold THINGS LIKE baskets.'

In 23b) the same connective occurs after the first and only overt exemplar 'basket'. One may wonder where lies the difference between such a one-slot connective construction and the similative plural in 13b) or the derivational strategy in 14c). Actually, the difference is not functional, and even structurally we observe a comparable [exemplar-MARKER] scheme. The difference lies in the synchronic distributional properties of the marker, which lead the authors of the grammars to classify them as plural affixes, derivational markers or connectives, respectively. It does not come as a surprise, indeed, that a closer look at data shows frequent diachronic links between the different types of strategies. Let us see the case of Hakha Lai in 24):

24) Hakha Lai (Sino-Tibetan, Kuki-Chin)

a. *làwthlawpaa=ni*ʔ *vok-**teè-pool*** *ʔa-tsook*
farmer=ERG pig-COLL 3SG.SBJ-buy2
'The farmer bought pigs AND SUCH (e.g. other domesticated animals).'
b. *làwthlawpaa=ni*ʔ *vok-**teè*** *ʔaàr-**tee*** *tsoo-**tee*** *ʔa-tsook-hnaa*
farmer=ERG pig-COLL chicken-COLL cow-COLL 3SG.SBJ-buy2-PL.OBJ
'The farmer bought pigs, chicken, cows AND SUCH (e.g. other domesticated animals).'

As argued by Peterson and VanBik (2004: 351), the nominal suffixes - *teè* and *–pool* are used in constructions having the semantics of what Haspelmath (2004) calls 'representative conjunction'. However, their value is not clear (due to law text frequency) and they should probably be analyzed as collective markers. More examples on the diachronic connection between the various types of constructions are provided in section 3.1.5.

Example 25) provides an instance of the correlative coordinators *-a…-a* in Mandarin Chinese, which can only be used in non-exhaustive lists, as seen in 25a) and 25b). "If the conjuncts do not form an open set, as in 25c), the coordinators may not be used" (Zhang 2008: 137).

25) Mandarin Chinese
a. *Shu-**a**, baozhi-**a**, bai-man-le zhengge shujia.*
book-and newspaper-and put-full-PF whole bookshelf
'Books and newspapers, AMONG OTHER THINGS, occupied the whole bookshelf.'
b. *Tamen tiao-**a** chang-**a**, huanqing shengli.*
they dance-and sing-and celebrate victory
'They sang, danced, AMONG OTHER ACTIVITIES, to celebrate the victory.'
c. *Yin-(*a) yang-(*a) duili.*
yin-and yang-and opposite
'Yin and yang are opposites.'

Japanese has an extremely rich system of non-exhaustive connectives (Chino 2001, Tanimori 1994: 121-122, 265). *Ya* in 26) implies that the linked items are examples taken from a larger group of items. By contrast, *to* implies that the items stated are the only ones under consideration. *Ya* is often combined with *nado* ("and such", see below) reinforcing its basic meaning, and can only be used at the nominal level. With verbs the suffix *-tari* is attested, as exemplified in 27). Again, as we observed for Koasati, *-tari* may also be attested after a single exemplar, as in 27b).

26) Japanese (Chino 2001: 41)
Watashi no heya ni wa,[*konpyūtā **ya** sutereo ga*] *oite arimasu.*
I DET room in TOP computer and stereo SBJ place-SUSP be-POL.NPST
'In my room there is a computer, a stereo AND OTHER SIMILAR THINGS.'

27) Japanese (Chino 2001: 108-109)
a. *Nichiyōbi wa taitei tomodachi to tenisu* [*o shi-**tari** eiga o mi ni it-**tari***]
Sunday TOP usually friend with tennis OBJ do-and film OBJ see to go-and
shimasu
do.POL.NPST
'On Sundays I usually do SUCH THINGS AS play tennis with my friends or go to see movies.'

b. *Tenki no warui hi ni wa, ie de* [*ongaku o kii-**tari***]
weather DET bad-NPST day on TOP home at music OBJ listen-and
shimasu.
do.POL.NPAST
'On days when the weather is bad I listen to music and do OTHER SUCH THINGS at home**.'**

Recently, also Italian developed a non-exhaustive connective, which is still restricted to the colloquial variety, namely *piuttosto che* (cf. Bazzanella and Cristofoli 1998, Brucale 2012, Mauri and Giacalone 2015). This originally preferential construction (meaning 'rather than') is nowadays attested both with its source value and with a disjunctive 'or' meaning. However, with the latter it can only be used when the speaker's aim is to name some potential exemplars of a non-exhaustive list, as in 28) and 29), in order to construct (or imply) an ad hoc category ('customizations you may wish to have' in 28), 'pills to lose weight' in 29)). Disjunctive *piuttosto che* cannot occur in alternative questions aimed at a choice, which by definition imply an exhaustive list of alternatives.

28) [parlando di desktop] *c'e' il vantaggio che ti puoi customizzare la*
there.is DEF advantage that CLIT can.2SG customize DEF
*macchina come vuoi, in relazione alle tue esigenze (grafica,**piuttosto che***
machine as want.2SG in relation to.DEF your.PL need.PL graphics *piuttosto che*
*sviluppo, **piuttosto che** giochi…)*
development *piuttosto che* games
'[talking about desktop] there is the advantage that you may customize the machine (pc) as you prefer, depending on your needs (graphics, development, videogames OR SIMILAR THINGS…'

29) [talking about diets] *ti dico la verità io sono contraria a pastiglie in generale*
CLIT tell.1SG DEF truth I am against to pills in general
*es.: kalo **piuttosto che**... però forse perchè non le ho mai provate*
ex kalo *piuttosto che* but maybe because NEG CLIT have.1SG never tried
'I'll tell you the truth I am against pills in general, es.: kalo OR STUFF LIKE THAT… but maybe it's because I've never tried them' (discussion in a forum: http://forum.alfemminile.com/forum/fitness1/__f1443_fitness1-Messaggio-a-bionda-73.html)

Non-exhaustive connectives appear as a frequent, though little studied, strategy to construct ad hoc categories in discourse. The notion of non-exhaustivity indeed implies reference to potential further members of a given set, in addition to the ones explicitly mentioned. It thus implies abstraction over the explicit members through a context-determined similarity reasoning, leading to the construction of an ad hoc category.

3.1.5 *General extenders*

When speakers compose non-exhaustive lists, they frequently recur to so-called general extenders, especially when no dedicated connective (of the type described in the previous section) is attested in their language. There is great terminological variation when referring to this construction type and true cross-linguistic perspective is lacking (with English and

French holding a monopoly in the literature): Dubois (1993) calls them *extension particles*, Dines (1980) *set marking tags,* Aijmer (1985) *utterance-final tags*, Channel (1994) *vague category identifiers,* Overstreet (1999) *general extenders*. The latter is the most widespread label.

Overstreet (1999: 3) calls these expressions "[…] 'general' because they are non-specific, and 'extenders' because they extend otherwise grammatically complete utterances". According to her, "the general extender has been treated as a form that indicates additional members of a list, set, or category. The general assumption has been that these expressions combine with a named exemplar (or exemplars), […], some non-specific form of reference" (1999: 11; cf. also Channel 1994, who calls such constructions *vague category identifiers*, distinguishing them from approximators 'kind of, sort of…etc.').

Most constructions analyzed in the literature are analytic, with a structure like the following: [connective + proform (+ similarity)], as shown in Table 1. Analytic general extenders are transparent with respect to the operations underlying their function, which consists of linking exemplars of a category (through connectives, e.g. *and, or*) and referring to further potential members (through preforms, e.g. *something, stuff, what…*), which are associated to the one(s) mentioned explicitly by virtue of some context-dependent similarity (frequently denoted by similarity markers, e.g. *of the like, similar, so…*).

Conjunctive general extenders		**Disjunctive general extenders**
and stuff (like that)	and the rest	or *something* (like that)
and all (that)	and this and that	or *anything* (like that)
and everything (like that)	and whatever	or *what*
and blah blah blah	[…]	or *whatever*
and that		or *what* have you
and the like		or *anyone* (like that)
and such		or *anybody* (like that)
and so on		or *someone* (like that)
and so forth		or *someplace* (like that)
and whatnot		or *somewhere* (like that)

Table 1: *Conjunctive and disjunctive general extenders in English (Overstreet 1999: 4, adapted).*

A cross-linguistic perspective as the one adopted here, however, shows that general extenders may also be synthetic strategies. Example 30) from Hausa shows the morpheme *kàzā,* that Jaggar (2001: 356) analyses as a "non-specific pro-form used to express similarity, substituting for an implied noun or hypernym with a similar meaning to preceding nouns", attested only for inanimate nouns. *Kàzā* is analyzed as a post-head modifier with a non-specific 'such-and-such' value 30a). The construction *dà* + *kàzā* in 30b) thus consists of the additive connective *dà* 'and' followed by the non-specific pro-form *kàzā*:[5]

30) Hausa (Jaggar 2001: 356)
a. *zā tà tàfī* ***jāmi'ắ kàzā***
'she's going to SUCH AND SUCH a university'
b. *yā jē Kanò, dà Zāriyà, dà Kādūna,* ***dà kàzā dà kàzā***
'he went to Kano, Zaria, Kaduna, AND SO ON AND SO FORTH.'

Synthetic general extenders may derive from originally analytic strategies: *etcetera* is the

[5] For example (23) glosses are not provided in the source.

result of a univerbation process from Latin *et cetera* 'and the remaining (similar) things'. A similar expression is found in Dutch, where *enzovoorts* derives from *en zo voorts* (cf. Eng. *and so forth*).

Crucially, general extenders need not occur in a true list, but may follow a single exemplar, just like one-slot connectives in 27b) and 29), or the already discussed cases of special plurals and derivational strategies, which by definition do not involve any list. From a purely structural point of view, in all these cases we are faced with one exemplar and some dedicated marker encoding reference to further, similar elements, generating a process of abstraction leading to the relevant ad hoc category. As already mentioned, in some cases it is possible to identify some diachronic link between the strategies described.

First, non-exhaustive connectives that allow for a one-slot construction (such as *piuttosto che* in Italian or *–tari* in Japanese) may be analyzed as connectives that reached a more advanced stage of pragmaticalization, developing the function of general extenders. Another frequent diachronic source for synthetic general extenders are interrogative or other indefinite pronouns, meaning 'what, whatever'. In 31) two examples from Galo are provided, where *joò* 'what' (31a) first developed a 'whatever' value and then what Post (2007: 344-346) analyzes as "universal pro-form 'etcetera; and all that sort of thing; and so on'" (31b):

31) Galo (Tibeto-Burman, Western Tani, Post 2007: 344-346)

a. *əráp=əəm agùm akə=əə jə́ə bəre*
door=ACC exterior DST.ABL.SLEV=TOP who CJEC
ɲíi=əə com ***jòo=əə*** *com cɨ́ɨ-nə́*
person=COP.IPFV GUES **what**=COP.IPFV GUES slap-MOVE.1
cɨ́ɨ-bó-káa
slap-MOVE.2-PF
'Someone...who could it be? Is it a person OR WHAT?...knocked on the door.' (Post 2007: 45)

b. *hottúm-horə́ rɨ-kú-nam rɨ-nam=əəm dó-pàk-là(a)*
bear-boar do-CMPL-NZR:RLS do-NZR:OBJ=ACC eat-RID-NF
jòo-là(a)
and.so.on-NF
'All that we in the end produced was eaten up AND ALL by wild animals.' (Post 2007: 345)

A similar path is attested in Mandarin Chinese, where the interrogative pronoun *shenme* 'what' "is grammaticalized as an *indefinite pronoun* and indefinite adjective meaning 'and so on, etcetera, or something'" Hsieh (1997: 108):

32) Mandarin Chinese (Sino-Tibetan, Chinese)
ranhou bir u shuo wo gei ta de you yixie huikuei huo-shi-shuo ah ta bushi buhui sajiau ah ***shenme shenme*** *de*
'Then, for instance, I gave her some feedback that she can't... she doesn't know how to show femininity, WHAT AND WHAT (= AND OTHER SUCH THINGS).'[6]

An interesting diachronic path can be observed in the development of Modern Japanese general extender *nado* 'etcetera, and so on'. Synchronically, it is an independent morpheme occurring at the end of non-exhaustive lists. This form, however, was attested in Classical

[6] For example 32) glosses are missing in the source.

Japanese as *–nado* (see example 33) below) and had the function of a similative plural (representative plural, Vovin 2003:40), opposed to the additive plural form *–domo.* As can be observed in 33b), the locative suffix has scope over the similative plural, and thus over the potential referents of the ad hoc category. In the history of the Japanese system, we are then faced with a path going from morphology to lexicon, from grammar to pragmatics, along what has been called *degrammaticalization* (Ramat 1992): a bound morpheme that was part of the number paradigm (C. Japanese *-nado*) has become an independent, non obligatory morpheme characterized by syntactic flexibility and pragmatic function (M. Japanese *nado*).

33) Classical Japanese (Vovin 2003: 40)
   a. *wabi-uta-**nado*** *kak-ite*
      grieve-song-**REPR** write-SUB
      "He wrote grieving songs AMONG OTHER THINGS"
   b. *tani-no* *soko-**nado**-ni* *fa*
      valley-GEN bottom-**REPR**-LOC TOP
      "at the bottom of valleys AND OTHER PLACES LIKE THAT"

3.2 Tendencies and correlations

The cross-linguistic survey described in the previous sections reveals on the one hand some degree of variation, but on the other hand it can be described on the basis of a restricted number of recurring features, both at the formal and at the semantic level.

At the structural level, the strategies attested to convey the process of ad hoc category construction can be located along a *continuum*, based on the degree of morphological integration with the mentioned exemplar(s).

High morphological integration	< ---------- >		**Low morphological integration**
NOMINAL INFLECTION (E.G. SPECIAL PLURALS)	DERIVATIONAL STRATEGIES	ECHO-WORD FORMATION / REDUPLICATION	CONNECTIVES, GENERAL EXTENDERS (LISTS)
Inflectional	< ---------- >		**Syntactic**

Table 2. *Inflectional-to-syntactic continuum of the strategies expressing ad hoc categories.*

The continuum in Table 2 is not meant to be a semantic map (cf. Haspalmath 2003), although it is likely that, once more data are available, a representation in terms of functional contiguity may fruitfully describe the synchronic and diachronic patterns of multifunctionality of the strategies at issue. For example, there is some evidence for recurrent diachronic patterns deriving general extenders from dedicated plurals and non-exhaustive connectives. Yet, more research is needed to draw a semantic map and, as the language sample grows, it is likely that more strategy types will be included in the study. For now, it suffices to point out the fact that the cognitive and communicative process under exam is expressed through the mobilization of strategies at different levels (inflectional paradigms, word formation and syntax) and with different degrees of obligatoriness and morphological integration with the mentioned exemplar.

As can be observed in Table 2, at the left-hand side of the continuum we find inflectional morphemes, such as associative, similative and collective plurals, characterized by a high

degree of morphological integration with the root to which they apply. Moving towards the right-hand side of the continuum, we encounter derivational strategies (mainly deriving aggregate nouns, cf. Joosten 2010) and echo-word formation, which share with inflectional markers a certain degree of morphological integration with the exemplar, but crucially pertain to the processes of word formation, and as such do not belong to obligatory paradigms (see sections 3.1.2 and 3.1.3 for examples).

Typically, inflectional and derivational constructions take only one exemplar as the starting point for the construction of an ad hoc category. Furthermore, inflectional and derivational strategies appear to correlate with ad hoc categories denoting sets of entities (be they animate or inanimate), rather than classes. This is probably due to a structural reason and to a semantic one. The structural reason is that plurals and collectives are inherently nominal strategies, and therefore naturally lead to a nominal *denotatum*. The semantic reason has to do with the fact that plurality and collectives are inherently additional, in that they imply conceiving a whole as composed by a set of co-occurring items, linked by a contiguity relation, and this may explain why these strategies do not refer to classes, i.e. sets of similar alternatives. The motivation underlying the contiguity association triggered by inflectional and derivational strategies may also consist of a frame, as is the case for Italian *Berlusconame* in example 15b): in order to identify potential values of X, the hearer has to be familiar with a narrative frame activated by the exemplar Berlusconi (i.e. his way of acting, the kind of people he used to hang out with, etc.).

Frames of activities and classes of situations tend to be constructed with the strategies characterized by the lowest degree of morphological integration, at the right-hand side of the continuum. Here we find non-exhaustive connectives and general extenders (discussed in section 3.1.4 and 3.1.5), which are syntactic in nature and usually occur within lists, operating at the discourse level. These strategies may be employed to derive any type of abstraction, be it a set, a class or a frame, involving entities, properties or states of affairs. In other words, while at the left-hand side of the continuum we are faced with linguistic strategies showing a number of structural and semantic restrictions, at the left-hand side we encounter strategies that are applicable to a wider range of abstraction processes.

Furthermore, we may observe that the degree of context-dependence may vary in the construction of ad hoc categories. For instance, building the category "things I usually do on Sundays" through a non-exhaustive list 'jogging, going to museums and things like that' is highly context-dependent and could not be constructed without reference to shared knowledge regarding the speaker (it's hard to determine what other members belong to the set without knowing *me*). On the other hand, constructing the category 'people and situations revolving around Berlusconi' through a derivational process like *berlusconame* (see example 15b) above) does not require access to the specific speech situation, but more to a general cultural knowledge of Italian politics. Such a difference in the role played by context (and by the specific type of context necessary for the category to be constructed) tends to correlate with a difference in the strategy type. Especially in ad hoc categories conveyed through derivational strategies, the type of context that has to be accessed in order to construct the category tends to be broader and in some way less anchored to the specific speech situation. On the other hand, other strategies - lying more at the discourse level than at the lexical one - appear to be more flexible and allow for the construction of categories whose interpretation is heavily dependent on a specific knowledge of the interlocutors and of the speech situation.

Some variation is also attested in the role played by the mentioned exemplar(s) in the process of categorization. In derivational strategies such as *berlusconame* and *grillame*, but also in some associative plurals (see for instance example 11)), the exemplar has to be interpreted as both a member of the category and the property P that all the other potential

members must share. This feature appears to be present when the exemplar is human, and even more systematically when it consists of a proper noun. In other words, in such cases the exemplar is the pivot of the set and it is also what the other members have in common: in the associative plural described in example 11), in addition to *János*, the category includes other members that entertain the same relationship with the pivot (*János*' friends, relatives, colleagues, etc.). In *grillame* (example 16)) the category comprises Grillo (a political leader) and all the persons revolving around him: supporters and colleagues, basically, who share the same relationship with him. Crucially, in these cases there is just one named exemplar/pivot of the set. When the exemplar(s) denote non-human animate or inanimate referents, instead, they tend to be analyzed not as a property shared by the other members, but simply as pointers to the category, i.e. as salient exemplars that the speaker considers sufficiently relevant to allow for the ad hoc abstraction.

The strategies encoding the construction of ad hoc categories are synchronically and diachronically connected to a number of functions involved in the *creation of* (mainly heterogeneous) *sets*: plurality, collectives, connectives, lists. The notion of plurality is a prerequisite for the construction of any set, except for one-member sets, which are however marginal cases. Word formation strategies are typically linked to the necessity to create new labels for sets/types of entities, properties of activities, which are perceived as salient enough to require specific words to name them. Connectives encode and realize the connection between entities, and the consequence of such connections is precisely the construction of sets, whatever the specific relation between the members may be. In particular, non-exhaustive connectives, in addition to establishing relations, also encode a specific property of the set, namely its openness. Finally, general extenders typically occur at the end of lists to contribute to the (non-)delimitation of the set, encoding its open-ended nature and referring to further potential members beyond the ones explicitly mentioned.

The connection between the construction of ad hoc categories and notions involved in the creation of sets (such as plurality, connections between entities, (non-)exhaustivity) is not surprising, since a category IS a set. Only, ad hoc categories are sets whose members are not identified on the basis of natural or frequently occurring associations as in common categories, but rather on the basis of specific communicative goals. As a consequence, the ad hoc set frequently (though not necessarily) comprises heterogeneous members (cf. *berlusconame* in example (15b), which includes persons, attitudes, situations, etc.) and is typically open-ended, i.e. open to enrichments by the hearer.

4. Categorization in discourse: from *ad hoc* categories to *online* categorization

The evaluation of the discourse relevance and discourse phenomenology of ad hoc categories allows to identify different ways in which ad hoc categories may be built in discourse, the function of the conversational move of creating an ad hoc category in various discourse situations, as well as the pragmatic mechanisms underlying such a move, both in terms of the speaker-hearer relationship and in terms of topic management. For the purpose of this paper, we will take into account the possible motivations underlying the choice of an ad hoc category strategy in discourse, focusing on the presence of a category label. The following discussion is based on a preliminary examination of corpus data from Italian and English (see section 2.3 for methodology).

First of all, ad hoc categories may be constructed with different functions in discourse, they may refer to a category lacking a label, or they may refer to a category having a label which is perceived as too generic for the specific context. In both cases, the process of categorization is anchored to the speech act through exemplification, that is, through a

bottom-up approach that starts from relevant exemplars to abstract higher-order entities. Let us consider example 34), where a ready-made linguistic label is available in the language, but the speaker chooses not to use it and rather to construct the category as ad hoc, in order to stress its context-dependence:

34) Japanese (Kuno 1973: 115)
*[Biiru-**ya** sake-o]* *takusan* *nomimashita.*
beer-and sake-ACC lots drank
'[I] drank lots of beer and sake (AND STUFF LIKE THAT)'

In 34), the speaker uses the non-exhaustive connective *-ya* to link 'beer' and 'sake'. By using this connective, the speaker implies that the list is not restricted to the mentioned exemplars, and makes reference to a more or less abstract category 'alcoholic drinks that I could order at the restaurant', that the addressee can construct based on the two mentioned exemplars. If the speaker wanted to refer only to 'beer and sake', she should have used the exhaustive connective *-to* instead of *–ya* (see section 3.1.4). Why does the speaker use a non-exhaustive list instead of using the label 'alcoholic drinks'? Because reference is made not to the common category, but to the particular set of alcoholic drinks that the speaker could have drank at the restaurant, thus requiring access to context and to the speaker's habits (e.g. never drinking vodka). The category is thus constructed as ad hoc in order to anchor its interpretation to the specific speech situation.

As argued by Ariel and Mauri (2016, on the creation of higher-level categories through *or*), the choice to construct an ad hoc category instead of using common categories (when they are available) may be due to a necessity to narrow down the abstract concept and tailor it to the specific context. For example, Ariel and Mauri provide the following explanation for example 35). *Position* is a rather general concept, which may imply reference to money, to stability, to power, etc. Montoyo's mentioning of two exemplars of the category in 35) helps us not just to arrive at the ad hoc construction of the correct set, but it also narrows it down to something like 'power and authority over people'.

35) MONTOYO: ... If I am,
... for example,
... the president (H) .. of .. a major labor union,
.. or a major corporation.
... the position,
.. (H) as president of that entity,
... gives me so much power. (SBC: 012)

In other words, ad hoc categories may not only satisfy a speaker or a hearer's need in abstracting over exemplars, but they may also help in the ever-necessary process of adjusting the linguistically expressed concept to the specific context (cf. lexical pragmatics, Wilson & Carston 2007, Carston 2010).

Barotto (2016) finds the distinction between lexicalized categories and non-lexicalized categories useful to account for the distribution of exemplifying constructions in Japanese, showing that speakers frequently recur to exemplars in order to refine and modulate the sets of referents denoted by a given label, resulting in a tailored category that avoids useless efforts for the hearer. Barotto identifies two distinct patterns in Japanese: on the one hand, speakers make reference to a category through two separate elements, i.e. a label and one or

more exemplars; on the other hand, speakers make reference to a category or a frame through one or more exemplars, without using any overt label.

Data from Italian confirm the fact that exemplifying constructions, i.e. linguistic strategies involving the identification of one or more exemplars, occur more or less in 50% of the cases together with a category lexicalized through some label (cf. Lo Baido 2016). In such cases, we are faced with an exemplar-driven abstraction process aimed at positioning the category borders closer to the relevant exemplars: the hearer processes the label meaning starting from the most accessible interpretations, and the choice of exemplars crucially determines what aspects of the label are processed first. Let us consider the following three utterances:

36) Italian

a. *Ho* *letto* *dei* *giornali* *oggi* ***tipo*** *il* *Corriere*
AUX.1SG read some newspapers today **such.as** DEF Correre
della *sera* *o* *Repubblica [...]*
della sera or Repubblica
'I have read some newspapers **such as** *Il Corriere della sera* or *Repubblica* [...]'
(Lip Corpus, M C 4 47 W)

b. *la* *XYZ* *gli* *ha* *pure* *detto* *quanto* *sei* *brutta* *vattene*
DEF XYZ CLIT.3SG AUX even said how be.2SG ugly.F go:IMP.2SG
via ***cose*** ***di questo*** ***genere*** *perche' lei* *ha* *vinto* *il* *dottorato* *a*
away **things** **of this** **type** because she AUX won DEF PhD in
francesistica
French studies
'XYZ even told her "How ugly you are!" "Go away!" and such things because she won a place in the PhD Program of French studies' (LIP Corpus R B 2 178 B)

c. *quando* *tu* *telefoni* *a* *una* *persona* *gli* *domandi* *come* *sta* ***eccetera***
when you call:2SG to INDEF person CLIT.3SG ask:2SG how stay:3SG **etcetera**
'When you call a person you ask him how he is doing etcetera.' (LIP Corpus F A 4 126 A)

In 36a) we have a label followed by two exemplars: the label *giornali* 'newspapers' is followed by *Corriere della Sera* and *Repubblica*, which help the hearer in restricting the borders of the category around the subclass of newspapers [MAJOR NATIONAL NEWSPAPERS], excluding smaller and local newspapers. In 36b), on the other hand, no label is produced by the speaker and the two explicit exemplars are the only clues available for the hearer to abstract the class [BAD THINGS AN ANGRY PERSON COULD SAY]. In 36c) *come sta* 'how he is doing' is followed by *eccetera*, and is thus the only exemplar that the speaker provides for abstracting the frame [BEGINNING OF A PHONE CALL], already activated by the premise, i.e. *quando tu telefoni a una persona* 'when you call a person'. In 36a) we are faced with the pattern [label + exemplars], while in the other two cases there is no overt label for the category. Of course, potential labels are available also for 36b) and 36c), but they would be rather complex and difficult to process, if compared to 'newspapers'.

As argued by Barotto (2016), "the possibility of lexicalizing a category is not an inherent feature which can draw the distinction between types of categories, but rather an arbitrary communicative strategy to create and express categories in specific contexts". The question then arises whether any category, be it common or *ad hoc*, may be constructed in an *ad hoc* way, that is, by using some relevant exemplars as anchors to the specific context.

In this perspective, the strategies under examination have to be considered as tools to convey a specific *way* of constructing categories, rather than tools to denote a specific

category type. [MAJOR NATIONAL NEWSPAPERS] in 36a) cannot be argued to be an ad hoc category like the ones constructed in 36b) and 36c), however the inferential process leading to construct the specific class of newspapers intended by the speaker is similar the one applied for the other two cases: saturation, associative reasoning leading to the identification of the Property P (relevant for the specific context) and abstraction (see section 2.1).

Discourse data thus point to the fact that exemplar-driven abstraction is attested both with and without a category label, both with common and ad hoc categories. *Ad hoc* categorization appears to be a context-dependent way of abstracting a category (*any* category) in discourse. Mauri and Sansò (in press) speak about *online* categorization processes, whereby the term *online* refers to the idea of constructing the category in real time, without directly importing a set from long-term memory.

In cognitive linguistics, Croft and Cruse (2004) extend the idea of context dependence to all types of categories, including those that Barsalou had called 'common' (see also Samuelson and Smith 1999). According to them, all categories are "inherently variable, and *on-line* created as and when needed." In this perspective, all categories are the result of a construal process, which is determined by the context. What linguistic data seem to provide evidence for is not the intrinsic nature of categories, but the fact that any category can be construed in a context-dependent way, and the strategies attested for this function show non-random patterns of variation. More corpus-based research will shed light on the actual discourse properties of these categorization processes.

5. CONCLUSIVE REMARKS AND PROSPECTS FOR FUTURE RESEARCH

The object of this paper is the linguistic reflections of a basic cognitive and communicative process, namely the construction of *ad hoc* categories. After a definition of ad hoc category in purely cognitive and semantic terms, the cross-linguistic variation observed in a typological survey based on 60 languages has been discussed and exemplified, revealing a great degree of formal and functional diversity. After discussing the observed regularities, the analysis of the construction of ad hoc categories in discourse has pointed to a different perspective, shifting the focus from a specific category type, namely ad hoc categories, to a specific type of categorization process, namely exemplar-driven, context-dependent abstraction, and ultimately to the idea of *online* construction of (any type of) categories.

This study constitutes the first step of a wider project in which the typological perspective is planned to be complemented by a diachronic and a discourse analysis (LEAdhoC project, see footnote 1). A comprehensive picture of how languages encode (and speakers use) ad hoc categories may have a strong impact on the disciplines involved in the modeling of human conceptual processing, providing them with a theory on the role played by verbal communication in the construction and use of non-stable categories. More generally, the analysis of how ad hoc categorization is expressed and used across languages may lead to important anthropological considerations: a typological perspective indeed naturally tackles the question whether there is something universal in categorization processes or whether, and to what degree, the construction and communication of categories is affected by specific and local cultural and linguistic factors.

ABBREVIATIONS

ABL = ablative
ACC = accusative
AOR = aorist
ASSOC = associative plural marker
ERG = ergative
EX = exemplary conjunction
EXCL = exclusive
F = feminine
PF = perfect
PL = plural
POL = polite
PROG = progressive

AUX = auxiliary
CJEC = conjectural
CLIT = clitic
CMPL = complement clause
COLL = collective
CONN = connective element
COP = copula
DAT = dative
DEF = definite
DET = determiner
DEM = demonstrative
DIR = directional
DM = discourse marker
DST = distal
EMPH = emphatic
GUES = guess
HON = honorific
IMP = imperative
IMPERS = impersonal
INAN = inanimate
INDEF = indefinite
INF = infinitive
IPFV = imperfective
LOC = locative
M = masculine
NEG = negative
NF = non finite
NPST = non-past
NZR = nominalizer
OBJ= object
PRX = proximal
PST = past
PURP = purposive
Q = question marker
QUOT = quotative
REPR=representative plural
RLS = realis
SG = singular
SIML= similative
SLEV = same topographic level
SUB = subject
SUBJ = subjunctive
SUPERL = superlative
SUSP = suspensive form
TOP = topic

REFERENCES

Aijmer, K. (1985) "What Happens at the End of Our Utterances?— The Use of Utterance-Final Tags Introduced by 'and' and 'or.'" *Papers from the 8th Scandinavian Conference of Linguistics,* Copenhagen, Institut for Philologie, Kopenhaven University, 366-389.

Ariel, M. & C. Mauri (2016) "Why use *or*?" *Submitted.*

Barotto, A. (2016). *Exemplification and categorization: the case of Japanese.* PhD Thesis, University of Bergamo.

Barsalou, L W. (1983) "Ad hoc categories" *Memory and Cognition* 11/3, 211-227.

Barsalou, L. W. (1991) "Deriving categories to achieve goals". In G.H. Bower (ed.), *The psychology of learning and motivation: Advances in research and theory*, San Diego, CA: Academic Press, 1-64. [Reprinted in A. Ram & D. Leake (Eds.), Goal-driven learning (1995, pp. 121-176). Cambridge, MA: MIT Press/Bradford Books]

Barsalou, L.W. (2003) "Situated simulation in the human conceptual system" *Language and Cognitive Processes*, 18, 513-562.[Reprinted in H. Moss & J. Hampton, Conceptual representation (pp. 513-566). East Sussex, UK: Psychology Press.]

Barsalou, L.W. (2010) "Ad hoc categories". In P.C. Hogan (ed.), *The Cambridge encyclopedia of the language sciences*, New York: Cambridge University Press, 87-88.

Bazzanella, C. & M. Cristofoli (1998) "Piuttosto che e le alternative non preferenziali: un mutamento in atto?" *Cuadernos de filologia italiana* 5, 267-278.

Brucale, L. (2012) "L'uso non canonico di piuttosto che coordinativo in italiano contemporaneo". In P. Bianchi, N. De Blasi, C. De Caprio, F. Montuori (eds.), *La variazione nell'italiano e nella sua storia. Varietà e varianti linguistiche e testuali,* Firenze, Cesati, 483-493.

Cardoso, H.C. (2009) The Indo-Portuguese language of Diu. PhD dissertation.

Carston, R. (2010) "Metaphor: Ad Hoc Concepts, Literal Meaning and Mental Images", *Proceedings of the Aristotelian Society* 110 (3pt3), 295-321.

Channell, J. (1994) *Vague Language*, Oxford University Press, Oxford

Chino, N. (2001) *All about Particles: A Handbook of Japanese Function Words*. Tokyo: Kodansha.

Corbett, G. (2000) *Number*, Cambridge University Press, Cambridge.

Daniel, M. (2000) *Tipologija associativnoj množestvennosti*. Russian State University for Humanities. (The typology of associative plurals).

Daniel, M. & E. Moravcsik (2005) "Associative plurals". In M. Dryer, M. Haspelmath, D. Gil and B. Comrie (eds.), *World Atlas of Language Structures*, Oxford, Oxford University Press, , 150–153

Delbrück, B. (1893) *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Strassburg: Trübner.

Dines, E. (1980) "Variation in Discourse—and stuff like that." *Language in Society* 1, 13-31.

DuBois, S. (1993) "Extension particles, etc." *Language Variation and Change* 4, 179-203.

Enfield, N. (2007) *A grammar of Lao,* Berlin, New York, Mouton de Gruyter.

Gaby, A. (2006) *A Grammar of Kuuk Thaayorre.* Unpublished PhD thesis, University of Melbourne.

Göksel, A. & C. Kerslake (2005) [Third reprint 2010], *Turkish, A Comprehensive Grammar*, London, Routledge.

Haspelmath, M. (2003) "The geometry of grammatical meaning: semantic maps and cross-linguistic comparison". In M. Tomasello (ed.) *The New Psychology of Language,* New York: Lawrence Erlbaum Associates Publishers, 211-43.

Haspelmath M. (2007), "Coordination". In Shopen, T. (ed.) *Language typology and syntactic description*, vol. II: Complex constructions. 2nd ed., Cambridge, Cambridge University Press, 1-51.

Hsieh, F. (2003) "Mandarin Chinese Shenme in Interaction". *Language, Information and Computation*: Proceedings of the 17th Pacific Asia Conference, 1-3 October, 2003, Sentosa, Singapore: 106-114.

Jaggar, P. (2001) *Hausa,* London Oriental and African Languages Library, Amsterdam: John Benjamins.

Jespersen, O. (1965) *The Philosophy of Language.* New York: Bantam.

Joosten, F. (2010) Collective nouns, aggregate nouns, and superordinates: When 'part of' and 'kind of' meet. *Lingvisticæ Investigationes* 33(1): 25-49.

Keane, E.L. (2005). "Phrasal reduplication and dual description". In Bernhard Hurch (ed.), *Studies on reduplication.* Mouton de Gruyter, Berlin, 237–259

Kuno, S. (1973) *The Structure of the Japanese Language*, Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Lo Baido, C. (2016). *Roads to exemplification in Italian: cognition and discourse*. Submitted.

Mauri, C. & A. Giacalone (2015) "*Piuttosto che*: dalla preferenza all'esemplificazione di alternative". *Cuadernos de Filología Italiana* 20: 49-72.

Mauri, C. & Sansò A. (in press). Strategie linguistiche per la costruzione *online* di categorie: uno quadro tipologico. In S. Caruana (ed.) *Atti del XLIX Convegno della Società di Linguistica Italiana*.

Moravcsik, E. (2003) "A semantic analysis of associative plurals", *Studies in Language* 27, 469–503.

Overstreet, M. (1999) *Whales, Candlelight, and Stuff Like That: General Extenders in English Discourse*, New York: Oxford University Press.

Overstreet, M. (2014) "The role of pragmatic function in the grammaticalization of English general extenders" *Pragmatics* 24:1, 105-129

Overstreet, M. & G. Yule (1997) "Locally contingent categorization in discourse", *Discourse Processes* 23, 83–97.

Peterson, D.A. & Vanbik, K. (2004). "Coordination in Hakha Lai (Tibeto-Burman)". In M. Haspelmath (ed.), *Coordinating constructions*. Amsterdam; Philadelphia, Penn, John Benjamins. 333-356.

Post, M. (2007). *A Grammar of Galo*. PhD Thesis. RCLT, La Trobe University.

Poletto, C. & N. Penello (2006), "Propagazione di suffissi derivazionali nel linguaggio giovanile", in G. Marcato (ed.) *Giovani*, *lingue e dialetti*. Atti del Convegno (Sappada - Plodn, 29 giugno - 3 luglio 2005), Padova, Unipress.
Ramat, P. (1992) "Thoughts on degrammaticalization". *Linguistics* 30, 549-560.
Recanati, F. (2014). *Literal Meaning*. Cambridge: Cambridge University Press.
Samuelson L. & Smith L. (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition* 73: 1-33.
Stassen, L. (2000) "AND-languages and WITH-languages". *Linguistic Typology* 4, 1-54.
Taylor, J.R. (2003) *Linguistic categorization*, Oxford, OUP.
Tanimori, M. (1994) *Handbook of Japanese Grammar,* Vermont, Tokyo, Tuttle Publishing.
Voghera, M. (2012) Chitarre, violino, banjo e cose del genere. In A. M. Thornton and M. Voghera (eds.) *Per Tullio De Mauro.Studi offerti dalle allieve in occasione del suo 80° compleanno.* Roma: Aracne, pp. 341-364.
Vovin, A. (2003). *A Reference Grammar of Classical Japanese Prose*. Alexander Vovin. RoutledgeCurzon
Wilson, D. & Carston, R. (2007) "A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts". In N. Burton-Roberts (ed.), *Pragmatics*, Palgrave Macmillan, 230-259
Zhang, Niina Ning (2008) "Encoding exhaustivity" USTWPL 4, 133-143.