UNIVERSITY OF BERGAMO

School of Doctoral Studies

Doctoral Degree in Linguistic Sciences

XXIX Cycle

SSD: L-LIN/01

EXEMPLIFICATION AND CATEGORIZATION:
THE CASE OF JAPANESE

Advisor

Chiar.ma Prof.ssa Caterina Mauri

Doctoral Thesis

Alessandra BAROTTO

Student ID 1031611

Academic year 2015/16

Acknowledgements

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

The following abbreviations are used in the interlinear glosses of language examples. In examples taken from a descriptive grammar, the glosses are generally the same as those used in the grammar. in examples taken from the corpus, the glosses have been made either by the author.

1SG	= first person singular	GRD	= gerundive
2SG	= second person singular	HON	= honorific
3SG	= third person singular	IMP	= imperative
1PL	= first person plural	LIK	= linker
2PL	= second person plural	LOC	= locative
3PL	= third person plural	MOD	= modal
ACC	= accusative	NEG	= negative
AOR	= aorist	NML	= nominalizer
ASP	= aspectual marker	NOM	= nominative
ADJ	= adjective	NPST	= non-past
ADV	= adverb	PASS	= passive
AUX	= auxiliary verb	PAST	= past
CAUS	= causative	PL	= plural
CLIT	= clitic	POL	= politeness
CLF	= classifier	POT	= potential
COM	= comitative	PP	= pragmatic particle
COND	= conditional	PRS	= present
COP	= copula	Q	= question marker
DAT	= dative	QT	= quotative marker
DEF	= definite	SG	= singular
DEM	=demonstrative	STA	= stative
DES	= desiderative	STR	= instrumental
EVID	= evidential	TE	= *te*-form (gerundive)
GEN	= genitive	TOP	= topic

1 THEORETICAL FOUNDATIONS: THE NOTION OF EXEMPLIFICATION

1.1 INTRODUCTION AND AIMS
1.1.1 WHY EXEMPLIFICATION?

> *The domestic life of a concept is a series of examples.*
>
> (Birk 2007: 5)

This research investigates the linguistic coding and functions of exemplifying constructions (i.e., linguistic constructions that signify exemplification), with a special focus on their role in constructing on-line contextually relevant categories at the cognitive level (cf. Barsalou 1983, 2010, Mauri 2016). More specifically, we argue that exemplifying constructions are used as overt strategies to make explicit the online construction of conceptual categories, allowing the hearer to identify relevant exemplars as starting points for inferential and abstraction processes (see section 1.3).

Exemplification has been variably regarded (or disregarded) in different research fields, prompting Lyons to affirm that example is "metaphor's forgotten sibling" (1989: 4). In fact, examples and metaphors are both rhetorical figures with a long theoretical tradition, yet they have undergone different fates. While metaphors have been recognized as instruments of cognition, examples have been left in a background of indifference, often dismissed as an inferior form of reasoning shaped in a too obvious a form (cf. Aristotle 1984, Lyons 1989). For instance, metaphors have been thoughtfully examined in cognitive linguistics for their important contribution in the elaboration of complex ideas by using other more straightforward ideas (e.g., the relation between time and space, cf. Lakoff 1987). On the contrary, examples have not received the same attention, despite the apparent similarities between these two practises: both are used to understand complex (and often abstract) information starting from more concrete and straightforward material.

Indeed, exemplification is first and foremost a widespread process to elaborate and communicate complex information. Training manuals, scientific papers, journalistic articles move systematically from general statements to specific exemplifying instances, or vice versa, from examples to the generalizations extracted from them (Manzotti 1998: 99). In this regard, we can say that examples function as cognitive reference points, triggering inferential processes and thus facilitating the categorization of information (Zillmann 2002).

Additionally, the episodic narrative structure associated with examples provides a cognitive interface which is intrinsically more accessible than the one provided by abstract information (Bruner 2002, Lischinsky 2008).

On this basis, recent studies on epistemology (cf. Kuhn 1970) and psychology (cf. Rosch 1973, Medin and Schaffer 1978) have restored the view on exemplification, showing that any form of abstract thought comprises the traces of the specific exemplars that have been experienced. Of the wide range of devices introduced by rhetoric, few can exhibit both the sheer pervasiveness and the cognitive depth of examples (Lischinsky 2008: 243). However, despite providing a solid basis for future investigation, these studies do not deal systematically with the process of exemplification.

In linguistics, exemplifying constructions have been usually discussed (cf. Halliday and Hasan 1976, Longacre 1983, Hobbs 1985) with respect to the structure of discourse, focusing (often solely) on those analytic constructions that overtly signal the status of example of the following or preceding element(s), such as *for example* in English. In this work however, the notion of exemplifying construction will not be defined in structural terms, but it will be identified on the basis of functional criteria, since we are mainly interested in the cognitive mechanisms underlying the notion of exemplification. Specifically, a given construction will be defined as exemplifying construction when it is used to provide examples in discourse by suggesting the presence of further elements beyond those explicitly mentioned (see section 1.3 for detailed definitions of exemplification and exemplifying constructions in functional terms).

Recent studies on vagueness have investigated exemplifying constructions also as discourse-pragmatic strategies used by language users to indicate a vague categorization, that is, to assign entities and events to conceptual categories (Jucker et ali. 2003, Mihatsch 2010a, 2010b, Ghezzi 2013). While these studies mention the cognitive role of exemplification in the creation and communication of ad hoc categories (Barsalou 1983, see further in section 1.2.3), they tend to assume sociolinguistic or pragmatic perspectives, leaving the cognitive aspect a bit in the background, thus aiming at a different purpose than the one pursued here. Beyond that, none of these studies addresses the phenomenon of exemplification in a systematic way.

All in all, within the field of linguistics there is no dedicated and systematic survey on exemplification as a linguistic practice for cognitive purposes. Furthermore, except for few cases (Mihatsch 2010b, Ghezzi 2013), exemplification has been mainly analyzed without a

solid empirical evidence, which on the contrary, is crucial to demonstrate the cognitive relevance of examples and their actual role in communicating the construction of categories.

We aim to move a step forward towards filling such a theoretical and empirical gap, in the belief that, just like metaphors, also examples behave as cognitive devices that facilitate the elaboration of complex information, and for this reason they should be the object of systematic linguistic analyses. Metaphors and examples indeed provide a breeding ground for Croft and Cruse's (2004: 2) claim that "the representation of linguistic knowledge is essentially the same as the representation of other conceptual structures, and that the processes in which that knowledge is used are not fundamentally different from cognitive abilities that human beings use outside the domain of language". More specifically, we believe that the cognitive contribution of exemplification in verbal communication deserves more attention and should be investigated on the basis of actual corpus-data.

In this regard, our study constitutes the first attempt to provide a comprehensive study on exemplifying constructions starting from empirical data (i.e., corpus-based approach, see section 2.2 for complete information on the methodology). We will direct our attention both to the discourse and to the cognitive level, 1) by studying the linguistic coding of exemplification when it is used as a cognitive tool to create and communicate conceptual categories, and 2) by identifying the patterns enabling the development of further more pragmatic and discourse-oriented functions of exemplification (e.g., hedging functions, cf. Mihatsch 2010b, Ghezzi 2013).

This research may possibly lead to upgrade exemplification from forgotten sibling to a restored (to value) sibling.

1.1.2 WHY JAPANESE?

To answer our questions about the role of exemplification at the linguistic and cognitive level, we focus on a case study, in order to have a corpus-based empirical basis and investigate exemplifying constructions in their context. Furthermore, a case-study oriented methodology allows us to perform different types of analysis (both qualitative and quantitative), in order to obtain in-depth insights regarding the phenomenon under examination.

More precisely, this work is a case study based on Japanese. However, the specific language is not the final scope, but just a tool to observe the phenomenon under examination. Our aim is not to provide an exhaustive account of how exemplification is linguistically codified in Japanese, but to investigate what we believe to be a universal cognitive mechanism, namely exemplification, starting from a selected sample of its many

linguistic interfaces, in order to determine cognitively-motivated tendencies in its linguistic coding.

Japanese is not a random choice, but is motivated by several reasons.

First of all, most linguistic studies on exemplification and exemplifying constructions have been conducted on European languages, such as English (e.g., Longacre 1983, Hobbs 1985, Mihatsch 2010b), Italian (e.g., Manzotti 1998, Ghezzi 2013, Voghera 2013), French (e.g., Rossari and Jayez 1999) and other Romance languages (e.g., Mihatsch 2010b). Moreover, much attention has been devoted to exemplifying constructions that signal the status of example(s) through analytical expressions, such as *for example*, *ad esempio*, *par example*.

In our study, we would like to provide a different perspective on exemplification. For this reason, we have chosen a language without any connection to those mentioned above. Not only Japanese is completely unrelated to European language, but it also shows substantially different structural patterns (e.g., words order, cf. Iwasaki 2013). In addition, as will be explained in section 1.3, we will focus solely on strategies that do not rely on analytical expressions to explicitly signal the status of example. Since we start from the assumption that exemplification is a universal cognitive mechanism, we expect to find linguistic patterns and tendencies that are potentially valid for any language, despite language-specific structural preferences.

Another reason for choosing Japanese concerns the acknowledgement of exemplifying constructions by grammars. It is mostly unusual for grammars to interrogate and give feedback on the strategies used to provide examples, especially because exemplification is often achieved by lexical transparent means (e.g., constructions like *for example*). From this point of view, the case of Japanese is rather peculiar. As it will be made clear in section 1.3.2, in Japanese, dedicated exemplifying strategies are described by grammars (even those targeting L2 learners, cf. Chino 2001). Mostly they are synthetic strategies (e.g., dedicated connectives like *ya*), which belong to the grammar and follow precise morphosyntactic rules.

For instance, let us consider the case of the non-exhaustive connective *ya* (i.e., connectives that indicate non-exhaustive lists of items). Kuno (1973: 115) explicitly describes the main function of the non-exhaustive connective *ya* as exemplification. Similarly, Chino (2001: 41) notes that "*ya* implies that the items stated are taken as examples from a larger group of items". This is not an isolated case: Japanese exhibits a large group of analogous constructions (e.g., the connectives *tari, dano, yara, toka*) to provide examples

in discourse. Because of these peculiarities, Japanese seems to be an interesting point to start our analysis on the linguistic coding of exemplification.

Moreover, the fact that these strategies are part of the grammar and follow strict morphosyntactic rules makes them even more interesting to investigate. For instance, as will be explained in section 1.3.2, some markers can be used only with noun phrases, while other can be used only with verbal phrases. It follows that speakers must make precise choices regarding how to formulate the message to select the most suitable maker. In this regard, we will try to investigate these choices with an eye to identify patterns and tendencies regarding exemplification. In fact, as will be pointed out in sections 2.3.2.1 and 2.3.2.2 with respect to the semantic and syntactic coding of examples, the way speakers formulate their messages can have a strong impact on the cognitive elaboration thereof.

In addition to this, some of the strategies exhibited by Japanese to specifically provide examples are cross-linguistically rare. We have seen above the case of the non-exhaustive connective *ya*, representing a type of connectives used solely to indicate open-ended lists and in which "the conjuncts are taken as representative examples of a potentially larger class" (Haspelmath 2007: 24). Dedicated non-exhaustive connectives (or representative conjunction, cf. Haspelmath 2007) are less widespread than connectives which are simply compatible with non-exhaustivity (e.g., the English connectives *and* and *or*). Despite this, as pointed out above, Japanese exhibits an extremely rich system of non-exhaustive connectives which can be used to express exemplification.

Finally, another reason underlying the choice of Japanese is that it is a well-documented language both at a synchronic and diachronic level: we are provided with descriptive grammars covering a period of hundreds of years, written and spoken language corpora, and a variety of specific diachronic and synchronic studies. This ultimately allows us to pursue the type of investigation described at the beginning of this section, that is, an in-depth analysis (mainly at the synchronic level), where it is possible to examine many occurrences of exemplifying constructions in their actual context.

1.1.3 OVERVIEW

The work is organized as follows. Section 1.2 outlines a a state-of-the-art survey of the subject of exemplification considering different fields and focusing specifically on its role in the reference of context-relevant categories. In section 1.3, the notion of exemplification is defined in functional terms and a structural core is identified, then we point out which Japanese exemplifying constructions will be examined throughout the work. In chapter 2 the

object of this analysis will be delimited and the methodology will be described. First, the corpus and the questionnaire will be examined, then the parameters of analysis will be explained and discussed.

Chapter 3 and 4 focus on the cognitive functions of exemplification, providing data on the linguistic coding of lexicalized and non-lexicalized categories respectively. Each chapter starts with a definition of the type of construction examined, describing it on the basis of the presence or absence of a category label. Then, the analysis of the data will highlight tendencies and recurring patterns in the linguistic coding of these constructions. Finally, we will discuss the division of labour between category labels and examples and the role of the context in directing the underlying inferential processes.

Chapter 5 focuses on the pragmatic functions of exemplification. Specifically, we will investigate the usage of exemplifying constructions to perform semantic approximation and pragmatic hedging. We will argue that these functions are closely connected to the fact that exemplification is systematically associated to lack of referentiality.

To say that some given entity is an example implies reference to further potential alternatives, thus to non-exhaustive sets. However, Japanese exemplifying constructions are also attested in exhaustive contexts. Chapter 6 focuses precisely on these cases, highlighting those functions that do not seem to be related to the notion of non-exhaustivity. More specifically, we will examine the usage of the so-called non-exhaustive connectives to encode exhaustive lists, arguing that this function depends on the type of readings these connectives can assume. Furthermore, we will describe the usage of some exemplifying constructions to intensify the negative polarity of the utterance, linking this function to the lack of referentiality at the basis of the categorization and hedging functions.

In chapter 7 the coding patterns attested for non-exhaustive connectives and general extenders will be compared, highlighting further tendencies in the functional extension of exemplifying constructions. In particular, on the basis of the attested tendencies, a functional space of exemplification will be proposed. Finally, we will take a diachronic glance on the developments of the markers examined, showing a strong and cognitive-motivated relation between the notion of irreality and that of exemplification.

Chapter 8 summarizes the major results of this research and briefly indicates some possible developments which may further enrich future studies on exemplification and exemplifying constructions. Finally, the appendix presents an excerpt of the table used to classify corpus-data, in order to provide an instance of how the classification of occurrences

has been brought about, and the questionnaire used to verify the specific readings identified, both in the Japanese and in the English version.

1.2 THE NOTION OF EXEMPLIFICATION

For a long time, exemplification has been a minor object of study in classical logic and rhetoric, often dismissed or ignored because considered an inferior form of reasoning shaped in a too obvious a form. Lyons has called the example "metaphor's forgotten sibling" (1989: 4), comparing the different destinies of the two rhetoric figures. He points out that even though they both can trace their status in rhetorical theory at least back to Aristotle's Rhetoric, metaphor has received a greater amount of attention.

Even recently, while the value of metaphor has been recognized not only as a rhetorical figure, but also as an instrument of cognition, exemplification has been left in the background, still considered "an addition to what has already been said and proven independently" (Birk 2007: 3).

However, the fact that exemplification exhibits a greater depth is evident right from the description of the word *exemplum* in medieval Latin: "a clearing in the woods". Lyons elaborates the concept noting that:

> Only the clearing gives form or boundary to the woods. Only the woods permit the existence of a clearing. Likewise, example depends on the larger mass of history and experience, yet without the "clearings" provided by example that mass would be formless and difficult to integrate into any controlling systematic discourse (Lyons 1989: 3).

The main point of the clearing (the example) is its discernible structure, which contrasts with the unclear surrounding wood. Moreover, the clearing suggests a precise duality between the "inside" and the "outside", consequently shaping the boundaries of the wood. At the same time, it gestures toward the "outside", that is "some commonly recognized basis in a reality shared by speaker and listener, reader and writer" (1989: 4). It appears clear that behind the notion of example there is much more than just an obvious ornament.

Exemplification has been examined mainly as a communication process "through which meaning is clarified or supported by a second unit which illustrates the first by citing an example" (Hyland 2007: 270). These two functional cores make exemplification a widespread device whenever people need to build and communicate a new form of thought: examples can clarify it, illustrate it, justify it, discover new areas of application, propose or

suggest something new (Manzotti 1998: 107). In fact, exemplification is so pervasive that its strategic role in argumentative or persuasive discourses tends to go unnoticed. Nevertheless, an example is first of all a chosen representation of a concept: this means that people do not choose randomly one concrete instance among many others, but they deliberately choose that particular example that frames the general concept in a specific light or that highlights certain concrete features (Lischinsky 2008: 244).

Now the question arises of whether examples exhibit only a discursive value, or also a cognitive depth. In this section, we explore, through a survey of the state-of-the-art, the role that exemplification plays in different fields, analysing in particular its cognitive and communicative functions. This survey does not aim to be neither exhaustive nor complete, but our goal is to provide a concrete explanation as to why exemplification may prove to be an important communicative and cognitive strategy and why cognitive linguistics should continue to analyse linguistic strategies that codify exemplification.

1.2.1 EXEMPLARS IN PURE COGNITION

Although exemplification has often been perceived merely as fact of communication, studies on knowledge acquisition have shown that exemplification constitutes an essential element in cognitive processes, well before the mediation of the language, pointing out how the abstract thought can never be completely detached from the concrete experiences. In these studies, the example is considered of particular interest because of its duality nature: examples can be used to depict an average tendency, but also as an ideal prototype. The former aspect is what allows examples to enter into relations of similarity with perceived similar events, to the point of becoming representative of a larger category. The latter regards the cognitive guidance of the example, that is, the ability to behave like a model to shape and direct human behaviour. These two important cores have been analysed especially by research on knowledge accumulation and transmission and by categorization studies.

1.2.1.1 EXEMPLARS AS A SOURCE OF KNOWLEDGE

The role of examples has been recently explored in correlation with the notion of "tacit knowledge" (Polanyi 1967). The categorization of knowledge into tacit and explicit knowledge is just one of many possible models, although it has proved to be quite valuable in current research on knowledge accumulation and transmission. Explicit knowledge can be codified, transferred, shared and managed (Nonaka and Konno 1998). It is expressed in

words and numbers and can be shared by means of books, manuals, etc. An example of explicit knowledge is the notion that Rome is the capital of Italy: this is a piece of knowledge that we can write down, transmit and share through material means. In contrast, loosely speaking, tacit knowledge collects all those things that people know how to do but perhaps do not know how to verbalize. Stenmark (2001: 10) defines it as "a cultural, emotional and cognitive background, of which we are only marginally aware".

The term "tacit knowledge" has been introduced by Michael Polanyi, a Hungarian chemical engineer turned philosopher of science, who has addressed the lack of recognition of practical experience in laboratories during the elaboration of scientific theories. The idea behind the concept of tacit knowledge can be summed up by his assertion "we can know more than we can tell" (1967: 4). Polanyi explains his theory by stating that tacit knowledge comprises a range of concepts and sensory images that can be brought together to make sense of new situations. One of the most convincing examples of tacit knowledge is the process of language learning: the ability to speak a language is not just a matter of learning a list of grammatical rules, but it requires something entirely different, first of all a great deal of direct personal experiences. That is, it requires something that is not always known explicitly. Although it is used by all people even in everyday life situations, the collection of highly subjective insights and intuitions that is subsumed in the concept of "tacit knowledge" is difficult to capture and articulate because people carry it in their minds and they are not often aware of it.

Because tacit knowledge can only be gained through personal practical experience, examples seem to play an important part in this process, acting as focal models to shape and direct personal experience. In his book *The Structure of Scientific Revolutions* (1970), Kuhn acknowledges the work of Polanyi, who claims that the success of scientists depends directly on tacit knowledge acquired through practice and especially through the collection and the analysis of particular cases (e.g., laboratory experiments). This point has been expanded further by Kuhn who highlights the importance of laboratory examples, which he calls *exemplars*. According to Kuhn, exemplars are the fourth element in the disciplinary matrix (along with symbolic generalizations, metaphysical presumptions and values), and by them he means "the concrete problem-solutions that students encounter from the start of their scientific education, whether in laboratories, in examinations, or at the ends of chapters in science texts. ...All physicists, for example, begin by learning the same exemplars: problems such as the inclined plane, the conical pendulum, and Keplerian orbits" (1970: 187). Exemplars are crucial in the formation and evolution of scientific developments,

because "in the absence of such exemplars, the laws and theories he has previously learned would have little empirical content." (1970: 187-188). For this very reason, texts typically present not only abstract rules and theories, but also instances of scientific research, that is, the applications of those theories in the solution of important problems, along with the new experimental techniques employed in those applications.

These exemplars are regarded and used as models of exemplary science and consequently fulfil some important functions: they suggest new puzzles and new solutions to solve those puzzles. From this perspective, science proceeds on the basis of perceived similarity to exemplars: this is the principle that guides scientists through scientific research and what enables them to grasp new puzzle-situations and hence new potential solutions.

1.2.1.2 THE ROLE OF EXAMPLES IN CATEGORIZING INFORMATION

Even if Kuhn's approach on exemplars concerns a very specific field, still it has the great merit of emphasizing the two important cognitive cores of exemplification: 1) the relations of similarity and 2) the cognitive guidance of specific instances raised as exemplars to shape and direct human behaviour.

Zillmann (2002) expands further the notion that all forms of thought retain traces of the specific exemplars upon which they are shaped, explaining how exemplification can work at the cognitive level in extrapolating information to gain tacit knowledge about the world. Constantly surrounded by a thick continual flow of information, the human brain has been able from the dawn of time to extract experiential chunks through a deep-rooted mechanism.

The extrapolation is not made randomly, but with a precise focus on those events that are deemed vital for the welfare of individuals and could not be ignored without a loss. However, the simple extrapolation of event is not enough to acquire real knowledge. If these events are processed individually and in isolation, they may provide only a small amount of knowledge, since, as postulated by Heraclitus, there are no two events that are ever exactly the same. Therefore, to maximize the positive cognitive effect, the brain processes them through a systematic comparison with those previously collected, applying automatically relations of similarity to organize them into manageable chunks (see Bruner et ali. 1956, Chater and Hahn 1997). Whenever two or more events are deemed alike for sharing of a number of defining features "to a degree that makes them classifiable as members of the same population of events" (Zillmann 2002), they are grouped together in one category.

This is a crucial step because such categorization of vital events allows them to stop being only representative of themselves and start being representative of other (potential)

events. Therefore, they become exemplars of a larger population. Consequently, the brain is able to extrapolate information about other events within the same category and to potentially predict information about all other similar events. Thus, through a spontaneous inductive inference, this small collection of experiences serves as a basis for acquiring knowledge about past occurrences that could guide future behaviour. Furthermore, another important achievement of this categorization of information is to make easier and faster the identification of new events simply by placing them in an already known class.

The crucial role of categorization processes is immediately evident in the acquisition of new knowledge. Actually, there is no real novelty in this statement, since categorization has been long recognized especially by cognitive psychology (e.g., Bruner et ali. 1956) as an important cognitive instrument to store and sort information. However, Zillmann's insightful, albeit brief, discussion has the merit of focusing also on the cognitive value of a specific event whenever it is processed as representative of a larger population, that is, when it functions as a cognitive model. Indeed, the relationship between categories and exemplars seems to be a central notion in knowledge acquisition processes, nevertheless it has remained in the background until the work of Eleanor Rosch in the seventies.

1.2.1.3 PROTOTYPE THEORIES AND EXEMPLAR-BASED THEORIES

Discussions related to categorization have always been pervasive in psychology because of people's natural tendency to "see something as X" rather than simply seeing it (Wittgenstein 1953/1978): an interpretation of the world that is fundamentally an act of categorization.

Until the later part of the twentieth century, rule-based accounts of concept representation have prevailed (see Bruner et ali. 1956). According to this so-called "classical view", categories are considered to be well-defined, context-independent and pre-stored in memory. Category membership is determined by a series of rules (that is, criteria properties) that underlie the representation of the concept and that allow to determine whether an entity belongs within a specific category.

In the last decades, a growing dissatisfaction towards the assumptions of these rule-based models has been brewing. For instance, Wittgenstein (1953/1978) challenges the classical view anticipating many of its inadequacies addressed later by the so-called "contemporary view". He notes that it is not always possible to find a property (or even a set of properties) shared by all members by virtue of which they are grouped together. Wittgenstein addresses the question using *Spiel* "game" as an example of category in which

the members do not seem to share a common set of defining proprieties. According to him, the category is structured by "a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail" (Wittgenstein 1978: 33). In other words, some members of the category share a set of attributes and other members share another set. Therefore, there may be the case in which some members do not share any common attribute.

The discussion on the nature and structure of categories has been further triggered by empirical evidences collected by cognitive psychologists. In particular, the research of Rosch (1973, 1975) on natural categories deeply influenced the contemporary view on categories, changing the focus of categorization theories from abstract representation to concrete occurrences, that is, the exemplars. More precisely, we can define an exemplar as a "specific remembered instance" of members belonging to, or being representative of, a specific category (Reisberg 1997).

Through a series of empirical experiments, Rosch confirms that the categorization process is not a simple matter of sharing a specific set of common features. She investigates the structure of natural categories observing that some concrete instances of a category act as cognitive reference points, that is, people consider them to be more representative of a category than other members. Two main types of theories were elaborated in order to explain how categorization can derive from concrete instances: *Prototype* theories and *Exemplar-based* theories.

According to the *Prototype theory* (cf. Rosch 1973, 1975), categories show an internal graded structure. At the centre of the category there is the prototype which exhibits the highest concentration of defining attributes, while at the boundaries of the category there are those members which exhibit fewer characteristic features. The concept of prototype can be understood as a schematic representation of the conceptual core of a category (Taylor 1995: 60). Following this approach, members are not prototypical but are judged as good examples of the category in virtue of their similarity to the prototype. This graded structure of categories might suggest that the prototype has a major role with respect to the members of the category. Nonetheless, it is noteworthy that the prototype is in turn constantly re-shaped based on concrete experience, since it should be regarded mainly as an abstract summary representation.

Rosch (1978) describes categories as organized in hierarchies horizontally and vertically. She states that there is a horizontal dimension that "concerns the segmentation of categories at the same level of inclusiveness – the dimension on which dog, cat, bus, chair

and sofa vary" (1978: 30), but also a vertical dimension, that is, "the level of inclusiveness of the category – the dimension along which the terms collie, dog, mammal, animal and living thing vary" (1978: 30). The vertical dimension is based on the notion that not all levels of inclusiveness of categories are equally cognitively useful. Specifically, Rosch identifies the "basic level" as the most culturally salient and as the one that includes the optimal amount of stored information. Indeed, when people are asked to name presented objects as quickly as possible, they tend to use basic-level names. For instance, a picture of a Labrador would be named "dog" rather than "animal" (superordinate level) or "Labrador" (subordinate level).

The role of the exemplar is even more central in exemplar-based models, an alternative type of theories on categorization first proposed by Medin and Schaffer (1978). They claim that conceptual representation consists only of the actual individual instances that one has observed. Thus, while the prototype representation for the category *dog* consists of a collection of the most common features across all dogs, in an exemplar-based model the category *dog* is actually "the set of dogs that the person remembers" (Murphy 2002), that is, the category is represented by all the instances (exemplars) that belong to it. Thus, in contrast to the prototype models, there is no abstract summary representation that stands for all dogs. In a practical sense, it means that a new stimulus is acknowledged and identified based on personal experiences with that concept and compared to multiple well-known exemplars in a category in order to see which things it is most similar to. Not all memorized exemplars share the same status: "some memories might be more salient than others, and some might be incomplete and fuzzy due to forgetting" (Murphy 2002: 49). Nevertheless, one consults them all to decide about a specific concept.

Although the exemplar-based and the prototype models have often been presented in opposition to each other, actually both rely heavily on the concept of similarity, differing only in the type of process leading to their inference: only by resembling a prototype or a stock of exemplars a new stimulus can be located properly into a category.

1.2.2 THE COGNITIVE IMPORT OF EXAMPLES IN COMMUNICATION

At the discursive level, examples have been examined mainly for their argumentative and persuasive value (i.e., proving a point) and for their ability to organize the discourse, thus helping the comprehension (i.e. clarifying a point).

The cognitive import of examples provided in communication has never been explicitly discussed, yet it still emerges in different fields, with different degrees of recognition. In fact,

even if exemplification has been acknowledged as one of the most widespread and effective device both in classical and contemporary rhetoric, for a long time it has been also considered merely an inferior form of reasoning (Lischinsky 2008: 244), deficient in logic validity. This view has been attributed mainly to the fact that exemplification is based on an inductive form of inference (Aristotle 1984). However, Lyons (1989) provides another explanation for the devaluation of examples, noting that unlike metaphors that preserve a hidden core, examples are too open and "too obvious to attract attention" (1989: 5). Nevertheless, as Lyons himself notes, this very feature is the biggest strength of exemplification, that is "to shore up the 'inside' of the discourse by gesturing toward its 'outside', toward some commonly recognized basis shared by speaker and listener, reader and writer." (1989: 3-4). Moreover, being narrative anecdotes, examples can function as ornaments to embellish the discourse, since they are narratively more vibrant and richer in features that the abstract rule they are assumed to demonstrate.

1.2.2.1 EXAMPLES IN RHETORIC TRADITION

The first acknowledgement of exemplification can be traced back to classical rhetoric, where examples have been studied as forms of argumentation. Aristotle (1984) examines the forms of rhetorical argument distinguishing between those that are part of the inductive method and those that are part of the deductive one. Following this model, he identifies two means of argumentation: the enthymeme, as part of the deductive reasoning, and the example (*paradeigma*) as part of the inductive reasoning.

Even if enthymemes and examples seem to complement each other in the demonstration process, it is evident that for Aristotle there is a strong hierarchy, both from the point of view of the success of the speech ("speeches that rely on examples are as persuasive as the other kind, but those which rely on enthymemes excite the louder applause", 1984: 10), but also of the accuracy of the argumentative method ("where we are unable to argue by Enthymeme, we must try to demonstrate our point by this method of Example, and to convince our hearers thereby. If we can argue by Enthymeme, we should use our Examples as subsequent supplementary evidence. They should not precede the Enthymemes: that will give the argument an inductive air, which only rarely suits the conditions of speech-making", 1984: 111).

While Aristotle recognises the role of examples as argumentative devices, he also believes that they are fit for the kind of "persons who cannot take in at a glance a complicated argument, or follow a long chain of reasoning" (Aristotle 1984: 10). In other words, examples

are not appropriate for what he considers to be serious thought: they just serve to make up for the lack of stronger evidences, mainly because in many cases induction reasoning is not a suitable demonstrative argument.

More recently, Perelman and Olbrechts-Tyteca (1969) have addressed exemplification shifting the focus on inferential validity of examples. In their taxonomy on argumentative techniques, an entire section is devoted to the description of arguments from examples, which are described as devices that establish the structure of reality.

Following a distinction already highlighted by Aristotle (who separates examples used as inductive devices from those used as testimony), they distinguish between argument from examples proper, that is, the presentation of particular instances to establish a rule, and illustration, where examples are exhibited to support an already established regularity. Perelman and Olbrechts-Tyteca admit that this distinction can be quite subtle (1969: 358), nevertheless it enables us to see that sometimes rules and examples are used with different purposes than those traditionally ascribed to them. For example, "sometimes the rule is stated in order to lend support to the particular cases that appear to corroborate it" (1969: 359).

Perelman and Olbrechts-Tyteca seem to follow a very dynamic approach on how to use exemplification, both at the discursive and cognitive level. For instance, the number of instances and the similarity among them are essential features in order to regard particular cases as examples and not as mere information ("a single attorney, appearing on the stage, may be taken as a particular, rather than a representative, character. But if two attorneys are introduced in the same play, their behaviour will seem to exemplify the whole profession" 1969: 351). Therefore, in order to establish a rule, it takes something more than one particular case to activate the generalization inference in the hearer's mind. Specifically, we need at least a second example, through which to compare the first one, in order to find those similarities that lead to the abstract generalization.

They also observe that examples interact with each other, playing an active role in making a more accurate reference to an abstract generalization: "the mention of a further example modifies the meaning of the examples previously given, making it possible to define accurately the point of view from which the facts given earlier should be regarded" (1969: 354). This intuition about the semantics of the mentioned examples provides also interesting insights at the cognitive level. Let us imagine a list of three examples. Once the speaker mentions the second example, the hearer will immediately and spontaneously compare it with the previous one in order to identify potential similarities. The same happens after the

third example, which will be in turn compared to the second and the first ones. Only through the comparison it does become clear what features the hearer should consider to grasp the generalization and what features can be ignored without loss. Thus, we can say that examples play an active role in making a more accurate reference to an abstract generalization, to the point that they can increasingly adjust the reference generated by those mentioned before.

As for the illustration, Perelman and Olbrechts-Tyteca note that its role is quite different from that of analogy, since it does not replace the abstract with concrete, but it just represents a particular case that validates the rule (1969: 360). Furthermore, often the real purpose of illustration is to help the comprehension of the rule, by providing an unquestionable instance of its application.

Perelman and Olbrechts-Tyteca recognize that different examples might exhibit different discursive purposes: "while the first ones must be beyond question in order to count as heavily as possible in the discussion, the following profit ones by the credit attaching to the first, and the last ones may only serve as illustrations." (1969: 359). Therefore, the passage from example to illustration might even occur in the same utterance. Just through the comparison between the first mentioned examples, the rule may be established to the point of becoming well-attested, so that the following examples relate directly with the rule, functioning effectively as illustrations.

Finally, they argue that illustration runs less the risk of being misinterpreted, since people are guided in their inferential process by the well-known rule (1969: 358). This suggests that there is an active mutual relationship between examples and abstract rules: while the example provides a concrete support to the rule, the rule helps to grasp the reference of the example, indicating in a more direct way the features that are useful for the generalization.

1.2.2.2 THE INFLUENCE OF EXAMPLES ON JUDGEMENTS MAKING

While Perelman and Olbrechts-Tyteca dethatch the inferential validity of examples from their persuasive value focusing solely on the former, the latter has been explored thoughtfully by current research in communication, especially regarding informative discourse. The so-called *exemplification theory* (cf. Zillmann 1999) explores how concrete examples shape and influence people's opinions about the likelihood of facts. Zillmann (2002) notes that with the emergence of linguistic competencies and communication skills, relevant experiences have become communicable: people can rely also on communicated information in addition to those learned from first-hand experience. The ability to communicate events and

information is obviously advantageous in enabling also the acquisition of knowledge on issues that lay beyond the personal limited experience. However, at the same time, it has the disadvantage that experiences related by others can prove to be not completely reliable. This is particularly true since the dawn of mass media communication, due to the fact that "the capacity to reach large audiences carries with it the risk of misleading the public in case the disseminated information proves to be distorted and inaccurate or simply in error." (Zillmann 2002: 21).

The core foundation of the exemplification theory is built upon the idea that information sources can be categorized as base-rate or exemplar (Gibson and Zillmann 1994). Base-rate information are numbers, facts and figures, while exemplars are episodic illustrations that describe "causes, importance, and consequences of the problem under consideration form the unique perspective of an individual" (Brosius and Bathelt 1994: 48). Base-rates are considered more veridical and less partial. On the other hands, exemplars are generally perceived as more concrete, emotionally interesting and – being episodic narrative – more entraining. Usually in news reports, exemplars are used in combination with base-rates in order to provide not only a description, but also a demonstration of the event being discussed. Unsurprisingly, almost no journalistic text fails to include examples (Zillmann and Brosius 2000).

Different studies have questioned the relationship between these two sources of information by conducting experiments to assess their influence on the human mind in the process of creating judgements. More specifically, several investigations from social psychology (e.g., Baesler and Burgoon 1994, Bar-Hillel 1980) show that people tend to overlook explicit statistical evidence. On the contrary, anecdotal evidence proves to be more persuasive and seems to capture the attention of the hearer since it deals with the abstract in the more direct and intense way.

Furthermore, research on communication studies (e.g., Gibson and Zillmann 1994, Brosius and Bathelt 1994) demonstrates that the distribution of exemplars affected the way recipients perceive the event population: the estimated frequency of a certain event tends to be linked to how often that specific event is illustrated by mean of exemplars. In these investigations, the persuasive role of exemplars proves to be so intense that even when the ratios apparent from exemplar distributions are contradictory to those explicitly stated, recipients form their judgments on the exemplars, ignoring base-rates information (see Brosius and Bathelt 1994). Thus, once the representations of these events are biased, people's opinion could be easily misled. Empirical evidence shows that a collection of one-

sided exemplars can lead to one-sided judgements and the more extreme the selectiveness, the more extreme will be the corresponding judgements (for an overview, see Brosius 2003). Moreover, exemplification effects are stable in time even if recipients seem to be oblivious to them (Zillmann and Brosius 2000).

To be fair, it is necessary to point out that most of these experimental studies prove to be quite biased in their methodology, due to their limited and homogeneous range of topics and respondent subjects (mostly college students, see Brosius 2003: 183). Nevertheless, there is still enough evidence to reasonably accept the basic claim about the strong persuasiveness of examples, which consequently brings up the question: why are examples so influential in the process of forming judgements?

The fact that exemplars overweight base-rate information has often been explained in terms of the theory of heuristics (Tversky and Kahneman 1974). In psychology, heuristics are mental shortcuts that allow people to solve problems and form judgments quickly and efficiently, focusing on one aspect of a complex problem without using all potentially available information. Exemplification theory argues that media messages are processed through two cognitive heuristic mechanisms: representativeness and availability heuristics.

The representativeness heuristic indicates that judgements about event populations are based on the scrutiny of a collection of representative exemplars without any considerations on the size of the sample or sampling methods. The phenomenon called "base-rate fallacy" (Bar-Hillel 1980) is related to this heuristic and it illustrates the devaluation of abstract information and the dominant attention to concrete events.

The availability heuristic refers to the fact that judgements about event population are dependent on exemplars available for cognitive manifestation at the time judgements are made. This retrieval mechanism is considered a "function of ease" (Zillmann 2002) by which exemplars in memory are retrieved and consequently impose themselves, exerting a disproportionate influence on the evaluation of the exemplified event population and on the construction of mental judgements.

The mechanism of the availability heuristic is often expressed in term of vividness and salience of information (Taylor and Thomas 1982). Examples are perceived as more concrete and more easily retained. Moreover, examples are emotionally interesting, that is, they contribute to make the story more intense and dramatic, increasing the emotional involvement of the hearer (Brosius 2003). For example, Baesler and Burgoon (1994) claim that anecdotal evidence is usually more concrete and vivid than statistical evidence. In order to prove this, they manipulate a report providing both vivid statistical and anecdotal evidence

and non-vivid statistical and anecdotal evidence. The results show that when the evidence's vividness is balanced, the statistical evidence proves to be more convincing than the anecdotal evidence. This suggests that the persuasiveness of examples is to be ascribed to their usual greater vividness.

Finally, exemplars prove to be meaningful not only as persuasive devices, but also in the comprehension process. Hendriks Vettehen and van Snippenburg (2004) demonstrates that exemplification stimulates complexity of thought about (the various aspects of) an issue, helping to process the information. This is due to the fact that, as already investigated in section 1.2.1, exemplars have the ability to make the abstract comprehensible. Brosius (2003) explains the phenomenon claiming that since the dawn of time the human mind has been accustomed to rely only on information based on exemplars in order to come to a judgment (e.g., first-hand experiences, opinions of neighbours, experiences of friends and family members, reports). Base-rate data about reality are a relatively new source of information. Therefore, we can assume that the human mind may be less accustomed to them.

1.2.2.3 EXAMPLE AS A DISCURSIVE DEVICE

The elaborative value of exemplification has been examined, albeit briefly, also by different studies on coherence in discourse. Focusing on all those relations that "bind contiguous segments of text into a global structure for the text as a whole" (Hobbs, 1985: 1), these studies recognize exemplification as an important functional relation, ascribing to it slightly different functions. For example, according to Longacre (1983), exemplification is an elaborative device: an integral part of the structure of language, which further facilitates the discourse by providing concrete information to an abstract formulation. Hobbs (1985) defines exemplification as a positive expansion coherence relation whose purpose is to ease the listener's difficulties in comprehension.

In addition, existing literature on specific languages has shown that exemplifying strategies can function as discourse-pragmatic strategies, i.e., linguistic strategies that operate on the discourse-pragmatic level. Speakers may use them to signal their attitude towards the content of the utterance they are producing or towards the interlocutor.

For example, different language-specific studies suggest that exemplifying strategies can be used to perform hedging functions. This fact is confirmed by Taylor (2010) and Ohori (2004) in their studies of some Japanese exemplifying constructions (i.e., *nado, toka, tari*). They note that these constructions may be used as vagueness markers to make the

utterance fuzzier, and thus to attenuate statements. Speakers may use them to limit their own commitment to the truth of the propositional content conveyed by the utterance.

Similarly, Ghezzi (2013) analyses the pragmatic value of some Italian exemplifying constructions. In Italian, exemplification is signalled by a formally heterogeneous group of discourse markers. However, all of them exhibit some sort of hedging functions. For instance, she notes that the analytic marker *per esempio* (for example) "is employed to hedge strong assertions, which represent the speaker's point of view and are relativized by being presented as an arbitrary choice among many potential others" (2013: 164). Similar hedging functions are performed by epistemic markers such as *magari,* which indicate one of a set of possible options and therefore the arbitrariness of the example (Manzotti 1998).

(1.1) A: *Cosa potremmo fare domani pomeriggio?*
           What can- PRS.1PL do tomorrow afternoon?
           'What can we do tomorrow afternoon?'

       B: *Mah... potremmo fare un giro in bicicletta, magari.*
           Mah… can-PRS.1PL do a tour in bicycle, MAGARI
           'Mah, maybe we can go for a bike ride.' (Manzotti 1998: 109)

Interestingly, beyond hedging (see chapter 5 for further theoretical clarification), exemplification exhibits other functions at the discursive level. For instance, recent studies have demonstrated that some exemplifying strategies may also be used as focus markers to highlight a specific item. In Italian*, tipo* functions as an approximator (1.2), but also as a focus marker (1.3). Moreover, the focusing function proves to be very common in requests for information or clarification (Voghera 2013: 21).

(1.2) *Alla fine gli lascio un messaggio proprio tipo a*
        at.ART end him/her leave-PRS.1SG a message just type at
       mezzanotte.
        midnight
       'In the end I leave him/her a message at around like midnight.' (Voghera 2013: 20)

(1.3) *Quanti piercing hai? Tipo che ho*
        How many piercings have- PRS.2SG? Type that have-PRS.1SG
       già risposto 20 volte a questa domanda.
        already answered 20 times to this question.
       'How many piercings do you have? I've replied to that question, like 20 times.' (Voghera 2013:

21)

Similarly, *per esempio* (for example) can be used as a focus marker before a list (Bonvino et ali. 2009).

Therefore, apart from a pragmatic dimension of vagueness and hedging, we can also recognize a 'purely pragmatic' exemplification, in which the exemplifying construction works as a focus marker or a discourse relational device.

1.2.3 LANGUAGE AND COGNITION: THE ROLE OF EXAMPLES IN THE COMMUNICATION OF CATEGORIES

Up to this point, we have analysed how, over time, exemplification has been addressed by different fields for its cognitive value. While the contribution from linguistics has been relatively low so far, this trend has been recently reversed by some studies that investigate the role of exemplification in the online construction of context-relevant categories (cf. Barsalou 1983, Mauri 2016).

Concepts and categories should not be considered just as matter of perception and pure cognition, but also as sharing a strong connection with language, which allows them to transcend direct experience and to be easily communicated. This connection can be seen as a bidirectional relationship, that is, "one's repertoire of concepts may influence the types of word meanings that one learns, whereas the language that one speaks may influence the types of concepts that one forms." (Goldstone et ali. 2013: 622). While the research around the first issue seems to be more consolidated (e.g., Rosch et ali. 1976, Kersten and Billman 1997), the second proposal is still controversial. Nevertheless, some evidences have been provided indicating that whether a concept is learned in the presence or absence of language[1] may influence the way in which it is learned (see Gluck and Bower 1988).

Along these lines, the so-called discursive approach to cognition has emphasized the reciprocity between discourse and cognition, especially in categorization processes. Scholars who adopt this approach argue that "categorization is something we do, in talk, in order to accomplish social actions (persuasion, blamings, denials, regulations, accusations, etc.)" (Edwards 1991: 517). Categorization processes are always encountered as a part of a discourse, and therefore, they should be investigated in terms of the kind of discursive work they are designed for within a specific context of interaction. This fact is crucial,

---

[1] For instance, as it will be explained further in chapter 3, regarding categorization processes, the presence or the absence of a category label (or category name) may influence the way a category is inferred.

because if we assume that categorization is a "situated discursive practice", we also recognize that it is subject to the working of indexicality (that is, how situational and linguistic context are invoked), rhetoric and, more generally, pragmatics. Thus, the centrality of membership is itself something that language users focus on in their talk and it may be understood by means of indexicality or as a matter of disputation (Edwards 1997: 236). For instance, as for indexicality, "if, in a pencil and paper test situation, I am asked to 'name some typical birds', I may very likely mention robins and sparrows. […] But neither is at all likely to come to mind when I am greeted at the door with: 'I've just put the bird in the oven'" (Heritage 1984: 149). Or again, as for rhetoric, "with imagination one could envisage a situation where the choice between terms such as 'armchair' and 'chair' might be keenly contested: this could be a situation involving law courts and allegations of broken contractual agreements between furniture manufacturers and wholesalers" (Billig 1987: 136).

The connection between language and cognition becomes even more central in the light of recent evidence that category structure is not stable but dynamic, context-dependent and computable within a given situation. This new approach argues that concepts and categories are part of our general knowledge about the highly variable world. People do not learn concepts in isolation (as happens during many psychology experiments), on the contrary, they learn them as part of their understanding of the world. Barsalou (1982) describes this notion using as an example the category "frog". He notes that when people process the word frog in isolation, they do not think of "edible" as an attribute of the category. However, "edible" becomes highly accessible when the word is read in the context of "French restaurant". In other words, depending to the context, the prototypical features of the category can change and, accordingly, the location of category boundaries changes as well (see also Smith 2005).

Following this intuition, Barsalou (1983, 1991, 2003, 2010) has conducted several experiments culminating in the identification and conceptualization of *ad hoc categories*, i.e., categories that are spontaneously constructed to achieve a goal relevant in the current situation (e.g., "things to do in Beijing" while planning a vacation). These categories do not reside in long-term memory, that is, they are not based on a previously represented concept. However, once ad hoc categories are created and frequently used, they become well established in memory. Moreover, they show a graded internal structure, as postulated by Rosch regarding natural categories (cf. Rosch 1973, 1975).

In recent developments, the concept of ad hoc category has taken on a more radical form. Smith and Samuelson (1997) refuse completely the notion of fixed categories with

permanent representations, suggesting that "individual acts of categorization do not require an already represented concept." (Smith and Samuelson 1997: 167). All categories are created in a specific context and intended for a specific discursive or practical goal. Moreover, they have to be constantly re-interpreted according to the context. This idea seems to be supported by psychological experiments, which demonstrated that ad hoc categories are ubiquitous in everyday cognition (e.g., Medin et ali. 2006) and used rapidly and without troubles by different types of subjects in different types of situations (e.g., Lucariello and Nelson 1985). From the perspective of cognitive linguistics, Croft and Cruse (2004: 92) label this new view on categories "dynamic construal approach", referring to the fact that according to it "all aspects of conceptual categories are subject to construal". In other words, they depend on how the speaker conceptualizes the experience to be communicated, for the understanding of the hearer.

Following this new approach, some issues arise regarding what sort of strategies people might use to refer to these dynamic and context-dependent categories. Barsalou himself recognizes the crucial role of the language in the forming of ad hoc categories ("Both conceptual and linguistic mechanisms appear central to forming ad hoc categories. [...] Linguistically, people combine words in novel ways to index these concepts.", 2010: 86), calling for further study ("Much further study is needed to understand their structure and role in cognition. Important issues include the following: How do productive conceptual and linguistic mechanisms produce ad hoc categories?", 2010: 87).

For many decades, labels have been considered the primary strategy to communicate categories. This relates directly to the fact that for a long time categorization theories have mainly focused on what have been called "common" or "natural" categories, that is, "concepts designatable by words in natural languages" (Rosch 1975: 193). Common categories have been considered well established in memory with stable associations with linguistic expressions – that is, they are usually expressed by short and stable linguistic means (e.g., "fruits", "gifts"). However, different studies on semantics (e.g., Cruse 1986, Wilson and Carston 2007) have highlighted some flaws in this approach. For instance, Cruse (1986: 151) demonstrates that in everyday life people deal with categories lacking a specific label but having a conceptual reality (i.e., "covert categories", such as "among the verbs of locomotion for living creatures"), laying the foundations for a criticism to the concept of stable association between categories and labels.

The issue becomes even more complex when we recognize the essential role of context: if categories are not stable entities, but they are constructed dynamically any time they are

needed according to the specific context, the role of labels becomes problematic. While, labels can still be used to make reference to more specific context-dependent categories (e.g., Wilson and Carston 2007), there may be cases in which labels are perceived as inadequate, because linked to more general and abstract concepts. In these cases, alternative linguistic strategies are needed. In this regard, some scholars suggest exemplification as a potential strategy to bridge this linguistic gap.

For instance, Wittgenstein (1953/1978) addresses the role of exemplification in the communication of concepts without a precise set of attributes shared by all members. In these cases, in order to facilitate the comprehension, he proposes that the speaker should list some exemplars of the category: "How should we explain to someone what a game is? I imagine that we should describe *games* to him, and we might add: 'This *and similar things* are called 'games'" (1953/1978: 33).

More recently, some language-specific studies (e.g., Mihatsch 2010a, 2010b, Channel 1994, Overstreet 1999, Ghezzi 2013) have referred to the cognitive role of providing concrete exemplars (therefore, of exemplifying) in the communication of ad hoc categories. For example, Mihatsch recognizes the case of categorization through illustration or exemplification (2010a: 52).

Mauri (2016) moves further by proposing that "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." (2016: 3). She recognizes three inferential processes that enable the construction of abstract categories starting from the mention of examples: 1) saturation, 2) associative reasoning, 3) abstraction. Let us see in detail these processes.

First, the language user signals the status of example(s) of the mentioned item(s) by means of some overt, dedicated strategies to guide the hearer to recognize the existence of additional members beside the mentioned ones. These strategies have a precise referential function, that is "to mentally open an empty folder, where such further items can be 'saved'" (Mauri 2016: 4), thus indicating the presence of a variable X, whose identity needs to be saturated based on the specific context. This step is vital because it ensures that the hearer does not process the mentioned elements solely on the basis of their referential meanings, but as representative of a larger set that should be inferred.

The saturation process is fulfilled through associative reasoning (cf. Recanati, 2004), in the sense that the additional members must be associated, or associable, to the mentioned exemplars by virtue of a defining shared property P. In order to identify this property P, the

hearer compares the mentioned exemplars looking for their minimum common denominator that is relevant in the specific context.

Finally, on the basis of the recognized property P, the hearer is able to determine the inclusion or exclusion of other potential members. This ultimately leads to the construction of the superordinate category which includes explicit exemplars and implicit members through an abstractive process.

This process of interpretation by abstraction outlined by Mauri is exactly the type of cognitive work that we would like to investigate in the present study, i.e., how examples can work as inferential triggers to imagine other contextually related examples, ultimately leading to the construction of contextually relevant categories. This cognitive core would be the keystone of our working definition of 'exemplification' in section 1.3.

In this regard, as noted in the introductory section, we believe that exemplification, like its more acknowledged sibling the metaphor, can provide an important contribution to cognitive linguistics. More specifically, the study of how language users create and communicate categories in discourse starting from selected exemplar(s), may provide interesting insights concerning the storage and organization of knowledge in the human mind.

1.2.4 IDENTIFYING THE P: SIMILARITY-BASED AND FRAME-BASED CATEGORIES

Before moving forward with our analysis, it is essential to clarify some issues regarding the property P common to all the examples. In the previous section, we have seen that the members of the category are grouped together on the basis of a shared property P (see Mauri 2016). This means that, from the perspective of exemplification process, it represents the potential of being associated with other similar elements.

However, this notion of 'associability' needs further clarifications. More specifically, it is essential to understand on what basis items can be associated to build categories. Therefore, we need to answer two questions: 1) what type of property is the property P? and 2) how can people correctly identify the shared property P?

Different categorization studies have use the notion of similarity to explain the relationship established among category members, starting from the very intuitive idea that we classify together those things that we find similar. For instance, some categorization models in the '70s (Medin and Schaffer 1978, Rosch 1975) have strongly relied on the notion of similarity, seen as the main criterion according to which items can be grouped together in one category: items must be deemed alike for sharing of a certain number of essential features to a degree that makes them classifiable as members of the same category. According to

the Prototype model (see section 1.2.1.3), for example, a category is composed by all items that are sufficiently similar to the prototype. Additionally, these models trace back the ability to note similarity (and thus also differences) to our biological inheritance and social and cultural environment.

Later investigations (cf. Murphy and Medin 1985) have criticised the notion of similarity deemed as insufficient to account for the structure of categories, arguing in favour of an explanation-based model. They note that we use concepts and categories as ways of explaining the world to ourselves and others. It follows that categories are determined by our selecting the concept that better explains the instance to be categorized. Murphy and Medin (1985) provide the example of someone jumping into a swimming pool fully clothed. We may categorize this person as 'drunk', but similarity-based model cannot explain this process, because it is unlikely that the category "drunk" includes such a specific feature like 'jumping into a swimming pool fully clothed'. They argue that this categorization is possible because we have a theory about inebriation that explains this behaviour. Therefore, the observation of this behaviour in a specific situation might lead us to classify this person as "drunk".

Another element that surfaces from these investigations is the importance of the context. Consider again the example of the man in the swimming pool. Changing the context, the categorization changes as well: if someone is drowning, the same behaviour may be classified as "heroic". This means that there is a dependence of categorization (but also similarity) judgments on context. In this regard, Tversky (1977) notes that the relevant feature of a category is often "inferred from the general context". So, when people are asked about the similarity between USA and Russia, the answer strongly relies upon the feature that has more weight in that specific context. At the linguistic level, the same importance of the context was recognized by Lang (1984), who notes that the deduction of the common integrator (or, in our words, the Property P) does not involve only the semantic meaning of the conjuncts (or, in our words, the explicit members of the category), but also "involves various other factors obtainable only from either the situational context or the interactional setting of the given utterance or from extra-linguistic systems of knowledge, belief-systems etc." (1984: 27).

Similarity-based and theory-based models are not mutually exclusive, and similarity is still considered an important criterion in categorization, while acknowledging that there are other. Nevertheless, these developments have contributed to change the view on categories,

to the point that categorization is currently considered more like problem-solving than attribute-matching.

Less straightforward cases of categorization fit into the discussion as well, for example ad hoc categories. Let us consider the case of a slipper, a newspaper, a foot, and a can of bug spray, which all fit within the category 'things to use to kill a roach' (Overstreet 1999: 42). Thus, the members of this category are considered similar in one important respect: they are all things that can be used in an everyday situation to successfully kill a roach. However, this seems a tautological similarity, in the sense that the degree to which something is a good member of the category (and therefore similar to the other exemplars) is related solely to its property of being in the category (cf. Hampton 2001). A similar issue is raised when we consider a cricket bat, a cricket ball, a cricket umpire. While the connection seems intuitive, the modifier cricket does not pick out similarity shared by bats, balls and umpires (Lakoff 1987: 21). Again, it appears that to understand this category we cannot rely solely on similarity. To understand this type of categories we need to have access to the narrative frame ("specific unified framework of knowledge, or coherent schematizations of experience", Fillmore 1985: 223). Regarding the cricket examples, Austin says:

> The reason that all are called by the same name is perhaps that each has its part – its own special part – to play in the activity of cricketing: it is no good to say that cricket simply means "used in cricket": for we cannot explain what we mean by "cricket" except by explaining the special parts played in cricketing by the bat, ball, etc. (Austin 1940: 73)

In other words, Austin says that these categories are based on the "holistic structure governing our understanding of activities like cricket" (Lakoff 1987: 21) and thus on those things that are part of the activity, rather than based on shared properties or similarity. In other words, in this case, the elements are associated with each other by virtue of their contiguity within a narrative frame, i.e., the activity of cricket. It is only through having access to the narrative frame that we can identify what defines the category and its members.

The same holds for the 'things to use to kill a roach' example. Again, the members are perceived as associable only when they are processed in a specific narrative frame, that is, the action of killing a roach with a tool. The only difference between these two situations is that while in the cricket case, the members of the category represent different parts of the same narrative frame (i.e., contiguity), in the killing a roach case, the members represent all

good alternatives of a single specific part of the frames (i.e., paradigmatic choices), that is, the tool used to kill the roach.

In the light of the above, to better explain the notion of 'associability', we can borrow the binary model proposed by Joosten (2010). In his study on collective nouns, Joosten (2010) recognizes the fundamental distinction between and *kind-of* relations (i.e., hyponymy, e.g. bird – animal) and *part-of* relations (i.e., meronymy, e.g. wheel – bike). While in *kind-of* relations, the cohesion is provided by (constitutional, functional) similarity, that is the resemblance between the members, in *part-of* relations, it is provided by contiguity by an external bond (cf. Joosten 2010: 38).

Starting from Joosten's insights, Mauri (2016) notes that the members of a category are associated on the basis on 1) an intrinsic property, which may define the nature or function of the members; or 2) an extrinsic property, based on the contiguity of members recurring within a specific schema. In the former case, the associative reasoning leads to the identification of a relationship of similarity (cf. Joosten 2010: 32). Under this respect, it is noteworthy that natural categories (cf. Rosch 1973, 1975) are prototypical instances of similarity-based categories, as noted by Prototype-based and exemplars-based models (cf. section 1.2.1.3). On the contrary, in the latter case, elements are associated with each other by their contiguity within a narrative frame (Lakoff 1987).

Consider the following examples (Mauri 2016: 3-4):

(1.4)   Similarity-based [healthy drinks]:
        It is necessary to drink [water, herbal teas, smoothies, and the like] to be healthy.

(1.5)   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.

In (1.4), the category members are associated on the basis of an intrinsic similarity, that is, being all healthy drinks that do not contain any harmful substance. On the contrary, in (1.5), the category members are grouped together based on their contiguity in the schema 'eating at the restaurant', exactly like, in Austin's example, the cricket bat and the cricket ball are associated by their contiguity within the schema 'playing cricket'.

To sum up, we can say that depending on the relationship between the exemplars, the inferential processes examined in the previous section lead to the construction of different types of categories. In this regard, Mauri (2016) argues that the inferential processes are always the same, regardless of the type of categories. However, in the associative reasoning

(cf. Recanati 2004), people are called upon to correctly identify the specific property P shared by the exemplars.

With this further important step, we can determine in its entirety the cognitive process that we are going to investigate, as summarized in Figure 1.1.

```
┌──────────────┐   ┌──────────────┐   ╭──────────────╮        ┌──────────────┐   ┌──────────────┐
│ Exemplifying │   │  variable X  │   │  Associative │     ┌──│ P depends on │───│  Frame-based │
│ construction │───│to be saturated│──│   resoning   │─────┤  │ some intrinsic│   │  categories  │
└──────────────┘   └──────────────┘   ╰──────────────╯     │  │  similarity  │   └──────────────┘
                                                           │  └──────────────┘
                                                           │  ┌──────────────┐   ┌──────────────┐
                                                           └──│ P depends on │───│Similarity-based│
                                                              │some contiguity│   │  categories  │
                                                              │within a frame│   └──────────────┘
                                                              └──────────────┘
```
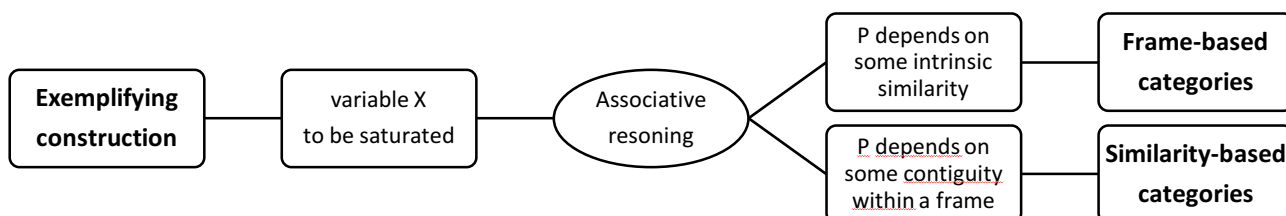
Figure 1.1: The creation of ad hoc categorization through exemplification

1.3 EXEMPLIFICATION IN JAPANESE

Now that we have analysed the theoretical picture currently available regarding the notion of 'exemplification', we need to shift our focus on the actual linguistic research. In other words, how can we apply what has been said so far to our analysis of Japanese exemplifying constructions?

At this stage of the investigation, we are adopting a top-down approach: we have started from a broad domain (in our case, exemplification), we have mapped sub-domains, acknowledging the one that will be crucial for our research (i.e., the usage of examples to refer to conceptual categories), and now we move forward to break down this sub-domain into a group of linguistic constructions. As noted in section 1.1.1, we are not looking for constructions exhibiting a specific structural form (e.g., analytic expressions like *for example*), but for any linguistic construction that can convey the postulated function we are analysing. This means that we expect to investigate a highly heterogeneous group of constructions.

To achieve our aim, it is necessary to elaborate a working definition of 'exemplification', which may provide an anchor for the investigation by defining the angle from which to approach the analysis. In this regard, the notion of 'exemplification' that has emerged from the previous sections is too broad to be used in an empirical linguistic analysis. Therefore, we need to formulate a definition that can specifically tell us 1) what is the main function performed by exemplifying constructions, and 2) what is the basic structure of exemplifying constructions.

To formulate this definition, we will start from the current conceptualization provided in the previous sections, trying to recognise the functional and semantic core of exemplification.

1.3 TOWARDS A DEFINITION OF EXEMPLIFICATION

As noted at the beginning of our theoretical survey, exemplification has been examined mainly as a communication process to clarify or to support abstract formulations by pointing out concrete instances of the phenomenon (Hyland 2007). In this regard, we have examined how in journalistic texts the usage of examples helps to make the abstraction more comprehensible (cf. section 1.2.2.2) and how examples can increase discourse coherence and improve the reader's understanding (cf. section 1.2.2.3). Moreover, exemplification is often used in argumentative and demonstrative processes to support or even formulate abstractions. As pointed out by rhetoric (cf. section 1.2.2.1), not only a concrete instance can be used as a guarantee of the applicability of a general concept (cf. illustration pattern in Perelman and Olbrechts-Tyteca), but the identification of similarities among examples can be used as a basis to formulate a general rule (cf. example pattern). The same process is frequently used (and sometimes even abused) also in persuasive or informative discourse (cf. section 1.2.2.2).

Nevertheless, these two functional cores (i.e., clarifying and supporting generalizations) appear to be strongly related to the cognitive dimension, specifically to the ability of exemplification to organize and elaborate complex information starting from more manageable material, namely, the concrete experience. It is no accident that examples play such an important role in heuristic theory (cf. section 1.2.2.2), functioning as mental shortcuts that ease the cognitive load of learning complex information. Indeed, one of the most basic heuristics identified by Pólya (1945) in the resolution of problems is "if the problem is abstract, try examining a concrete example".

We thus argue that exemplification is firstly a cognitive process that facilitates the organization of information, with the purpose of gaining knowledge about the world (cf. section 1.2.1). More specifically, it is a cognitive process in which an instance is profiled and construed as representative of an abstract formulation. From this basic core, all other manifestations of exemplification are generated.

Following this approach, as Mauri (2016) noted, the ultimate purpose of exemplification is to guide the interlocutor through a process of interpretation by abstraction, where the

mentioned examples[2] work as inferential triggers to imagine other contextually related examples (cf. section 1.2.3), and eventually pointing to the abstract formulation (i.e., the category). To achieve this, it is pivotal to mentally open a metaphorical empty folder that suggests the presence of further elements.

Therefore, we propose the following functional definition of exemplifying construction:

(1.6)   The label 'exemplifying construction' encompasses all those linguistic constructions whose ultimate function is to provide one or more exemplars in discourse. The status of example (i.e., representative item of a wider category) is signalled by suggesting the presence of further related elements.

Following this definition, even linguistic constructions that have not been directly identified as means to give examples fall within our scope of analysis. For example, the so-called "general extenders" (e.g., *and something like that*, *or something similar*) act as cues for the hearer to consider the preceding mentioned element "as an illustrative example of some more general case" (Dines 1980: 22). Interestingly, general extenders have been identified as strategies to encode ad hoc categories or, more generally, categories lacking a specific label (Channel 1994, Overstreet 1999).

In her cross-linguistic investigation, Mauri (2016) recognizes a wide range of strategies that can be used to provide concrete exemplars: from more transparent discourse-level strategies such as general extenders in English, to synthetic, less transparent means such as non-exhaustive connectives (i.e., connectives that explicitly convey the non-exhaustivity of the list), dedicated plurals (i.e., associative and similative plurals, see Daniel 2000 and Moravcsik 2003), derivational affixes or special types of reduplication (e.g. Turkish *m*-reduplication).

All these strategies share a basic structure that can be summarized as follows:

(1.7)   *i*) Exemplar(s) + *ii*) Non-exhaustive tag[3].

---

[2] In our analysis, with the label 'examples', we designate the concrete exemplars (or members, or items) mentioned in discourse to refer to the category that encompasses them. This means that in our analysis, the word *example* and *exemplar* are not equivalent. Following the usage of cognitive and psychology fields that employ the label *exemplar* to designate the items grouped together in a conceptual category at the cognitive level (cf. Rosch 1973, Taylor 1995), we employ the term *example* whenever discourse is the subject, reserving the term *exemplar* for strictly cognitive or psychological issues.

[3] Referring to general extenders, the term "vague tag" is commonly used (cf. Dubois, 1992). Nevertheless, we believe that the basic feature that general extenders share with all the others types of constructions here

The exemplar(s) serves to allow the hearer to identify the defining property of the set through associative reasoning. The non-exhaustive tag is used "to evoke some larger set" (Dubois, 1992:198). In other words, non-exhaustive tags allow to mentally open the empty folder by suggesting the presence of further elements and thus signalling the status of example(s) of the mentioned item(s), as described by Mauri (2006, cf. section 1.2.3). In this sense, we do not need explicit makers such as *for example*[4] to signal the process of providing examples, because the non-exhaustivity profiled by these tags can perform the same cognitive task.

Let us consider some of the strategies recognized by Mauri. While the order of the two elements (i.e., exemplars and non-exhaustive tag) may change depending on type of construction, the basic semantic core persists.

Similative plurals

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

[exemplar] [non-exhaustive tag]

| *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.'

Here, the similative plural marker *mbe* functions as a non-exhaustive tag which is added to the explicit exemplars *isu* "fish" and *nie* "oil" to open the reference. Specifically, similative plural (Corbett 2000: 101-111, Daniel 2000) differs from associative plural in that it denotes a set of objects sharing similar features rather than a group of closely related associates (Daniel and Moravcsik 2005). In this regard, it is noteworthy that while we can consider similative plurals as means to provide examples, associative plurals (which often apply to proper names to create set of people related to the exemplar) exhibit a more ambiguous status (see Mauri 2016).

---

considered is the ability to configure a non-exhaustive set of items. In some cases, and in some contexts, a vague reading can be added, mainly to elicit more pragmatic functions (cf. chapter 5).

[4] Indeed, canonical exemplifying constructions such as *for examples* can be considered non-exhaustive tags as well. In fact, their very semantic core is to signal the attached item as an example of a larger set of other potentially available options (cf. Manzotti 1998).

Derivational strategies

(1.9)   Kuuk Thaayorre (Australian, Pama-Nyungan, Gaby, 2006:642)

[exemplar =non-exhaustive tag]

pormpr=yuk   *koop*   *thiik-nhan*

house(ACC)=STUFF all   break-GO&:NPST

'All the houses and things will be broken [in a cyclone]'

Here, the suffix =*yuk* 'STUFF' acts as a non-exhaustive tag, since it can be added to nouns to designate a larger set. More specifically, it is used "to generalise their reference to include things normally associated with the denotatum of the noun in question" (Gaby, 2006: 642).

Reduplication

(1.10) Turkish (Göksel and Kerslake 2005: 91-92)

[exemplar] [non-exhaustive tag - exemplar]

Eve çat kapı bir alıcı geldi, **odalarí modalarí** dolaştı.

'Today a potential buyer came without notification, and looked at the rooms, etc.'

In Turkish, the *m*-reduplication is commonly employed to generalize the reference denoted by a specific word or phrase to include similar entities or events. It is attested especially in colloquial Turkish. It can be applied to animate and inanimate entities, proper nouns, and even to phrases with the same function.

Similar constructions are attested in South East Asian languages, such as Tamil and Lao, and are often described as *echo word formation* (cf. Keane 2005, Enfield 2007).

Non-exhaustive connectives

(1.11) Japanese (Kuno 1973: 115)

[exemplar] [non-exhaustive tag] [exemplar]

Biiru-ya   *sake-o*   *takusan*   *nomimashita.*

beer-and   sake-ACC   lots of   drink.POL.PAST

'[I] drank lots of beer and sake (and stuff like that)'

Here, the Japanese connective *ya* joins nouns to indicate a non-exhaustive list of items, making clear that the set is actually open and thus there are other instances to consider beyond those mentioned.

Indeed, the peculiarity of these connectives is that they occur only in open-ended lists.

For this reason, the label 'non-exhaustive connective' is well established in the literature on East Asian languages (Chino 2001, Zhang 2008) to describe the phenomenon. Haspelmath (2007: 24) briefly mentions this type of connectives using the label 'representative conjunction'. According to him, "in this construction, the conjuncts are taken as representative examples of a potentially larger class". Similar constructions are attested also in Kosati (Haspelmath 2007) and Hakha Lai (Peterson and VanBik 2004).

General extenders

(1.12) [exemplar] [non-exhaustive tag]

English (Dines 1980: 28)

she's sort of a child who swings and does somersaults and stuff like that

(1.13) Japanese (Chino 2001: 43)

sono	*hen*	*ni*	*wa*	*resutoran*	*ya*	*disuko*	*ya*	*eiga-kan*	*nado*	*ga*	*arimasu*
that	area	LOC	TOP	restaurant	YA	disco	YA	movie theatre	NADO	NOM	exist:POL

'There are restaurants, discos, movie theaters and so forth in that area.'[5]

The general extenders *and stuff like that* in (1.12) and *nado* in (1.13) act as non-exhaustive tags to widen the reference. Overstreet (1999: 3) labels these expressions as "'general' because they are nonspecific, 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" (Overstreet 1999: 11). It is noteworthy that general extenders include both analytic transparent constructions as shown in the example from English (1.12), but also synthetic expressions as shown in the example from Japanese (1.13).

After this brief cross-linguistic survey, we would like to point out that we are not excluding canonical exemplifying constructions (e.g., *for example* in English, *per esempio* in Italian). On the contrary, our aim is to demonstrate that once we apply a function-to-form approach, the actual linguistic constructions that can be used to provide examples form a much more heterogeneous group, where canonical exemplifying constructions are just one option among many others. Moreover, they also represent the more obvious and straightforward

---

[5] Here the general extender *nado* is used in combination with the non-exhaustive connective *ya*. This pattern is very common in Japanese (cf. Kuno 1973, Chino 2001).

option, where the examples are explicitly marked as such. However, this is not necessary: by means of non-exhaustivity, language users can provide examples in a much subtler way. In some cases, it is even possible that speakers are not aware themselves of the exemplifying process they are performing.

As previously mentioned in section 1.2.3, exemplifying constructions can be used to trigger the construction of contextually relevant categories. In this regard, while the mentioned exemplars indicate some concrete members of the category, the non-exhaustive tag suggests that the following or preceding element is to be construed as an exemplification of a wider category. In other words, the cognitive function of these strategies is to trigger and direct the inferential processes indicated by Mauri (2016, cf. section 1.2.3): saturation, associative reasoning and abstraction.

Let us examine in detail the linguistic interface of this cognitive process considering an example from English:

(1.14) *I grudgingly admit that it's a handy thing to have where I can check [email, Twitter, and etcetera]. In fact, these are the apps I'm using that are very essential to my daily life!* (enTenTen13)

Here, the exemplifying construction consists of two examples and a general extender. More specifically, with regards to the categorization process, the speaker refers to:

1. two explicit examples: *email, Twitter;*
2. additional implicit members by means of the non-exhaustive tag *etcetera*, which share with the examples the property 'accessible via an app', identifiable by drawing on the context;
3. a superordinate category [ONLINE SERVICES AND WEBSITES ACCESSIBLE VIA APPS], which includes explicit and implicit members.

This above is the cognitive mechanism that characterized exemplification and that we wish to investigate at the linguistic level. While this functional core is at the basis of our analysis, we will not consider it as merely a pre-determined postulation. On the contrary, during our investigation, the cognitive and linguistic processes underlying exemplification will be further investigated and particularly the role of non-exhaustivity will be discussed and examined (cf. chapter 4).

1.3.2 JAPANESE EXEMPLIFYING CONSTRUCTIONS

Using the working definition formulated in the previous section, we have selected four Japanese exemplifying constructions to be investigated: the non-exhaustive connectives *ya, tari* and *toka* and the synthetic general extender *nado.* Before moving to describe these strategies in details (see sections from 1.3.2.1 to 1.3.2.4), we would like to briefly motivate our decision to restrict the study to this specific set of constructions.

For instance, the reader may consider controversial our decision not to consider *tatoeba* "for example" in our analysis. We have made this decision for several reasons. At the structural level, *tatoeba* (like many similar constructions in other languages) tends to show a homogenous patter of usage (see Kaiser et ali. 2001: 80), where in one turn the speaker provides the abstract generalization and in subsequent turn he or she provides some concrete examples marked by *tatoeba* at the initial position of the utterance[6]. Since our aim is to monitor also quantitatively different structural patterns (e.g., the presence or the absence of a category label, cf. sections 1.2.3 and 3.1), the homogeneous behaviour of *tatoeba* may create a bias in the data.

Moreover, we would like to monitor strategies which do not require a conscious effort from the speaker to connect abstract generalization to concrete instance. In this sense, *tatoeba* suffers of being "too obvious" (to quote Lyons 1989)[7]. The process of providing examples is often a subtle one, where both language users are called upon to activate inferential processes even without explicit analytic expressions acting as triggers. This process is the type of exemplification process we would like to address, since it is also the less studied (cf. section 1.2.2.3 and section 1.2.3).

Another reason for choosing this specific set regards the peculiarities of the strategies attested in Japanese. As already mentioned in section 1.1.2, Japanese exhibits exemplifying constructions that are cross-linguistically rare (e.g., non-exhaustive connectives) and that follow more heterogeneous patterns of usage (e.g., do not require the explicit mention of the abstract generalization, unlike *tatoeba*).

Furthermore, we have chosen to investigate a small set of widespread constructions, rather than taking into consideration a larger group, in order 1) to have comparable amounts of data, and 2) to perform an in-depth analysis monitoring a good number of occurrences

---

[6] In grammars, *tatoeba* is often described as a means of introducing a paraphrase (cf. Kaiser et ali. 2001: 80).
[7] The kanji of *tatoeba* is that of "example" (*rei*). Therefore, not only *tatoeba* is glossed as "for example", but it literally refers to notion of example.

(200 occurrences for each construction), 3) to avoid constructions that overlap each other in usage and functions (thus creating potential biases in our quantitative study).

To achieve this, we have excluded constructions attested in grammar but with a low frequency in actual language[8] (e.g., the non-exhaustive connective *yara*, cf. Kuno 1973). Since our aim is a type of analysis that tries as much as possible to study the language as it is (as it will be explained further in chapter 2), we need also to consider the actual frequency of use.

In addition, we have also excluded register and style variants. For instance, *nanka* and *nante* are often considered register variants of *nado* used in less formal context and in spoken language (cf. Martin 1975: 160). As for their meaning, they are almost synonyms, that is, synthetic general extenders used to codify non-exhaustivity or lack of referentiality (cf. chapter 5). To avoid any potential bias, we have investigated only *nado*[9] because our corpus is mainly based on written texts (see section 2.2.1).

Taking all these issues into account, we have chosen four strategies that are widespread, but different to each other at several levels: 1) type of construction (i.e., non-exhaustive connective, general extender, connective that can also be used as general extender), 2) type of examples they can be attached to (i.e., noun phrases or verbal phrases), 3) register of usage (i.e., formal register vs. colloquial register), 4) channel of communication (i.e., written language vs. spoken language). In this way, we hope to monitor as many linguistic variants as possible. In Table 1.1, the major features of each strategy are schematized.

Table 1.1: Major features of Japanese exemplifying constructions. C = connective; GE = general extender

	ya	*nado*	*tari*	*toka*
Type of construction	C	GE	C/GE	C/GE
Syntactic level	NP	NP/VP	VP	NP
Register	formal	formal	formal/informal	informal
Channel of communication	written	written	written/spoken	spoken

---

[8] Information about the actual frequency has been gathered through a preliminary corpus-based investigation (for example using the Japanese section of the Leipzig Corpora Collection) and available quantitative studies on these constructions (cf. Taylor 2010, who confirms the high frequency of *toka* in spoken Contemporary Japanese).

[9] In Japanese section of the Leipzig Corpora Collection (LCC), the overall number of occurrences of *nado* is 232,986. The numbers of occurrences of *nanka* e *nante* are 1,022 and 3,847 respectively.

In the following sections, the modalities of usage of each strategy will be described in depth.

1.3.2.1 ya

The Japanese connective *ya* links nouns and noun phrases to indicate specifically non-exhaustive lists of items. In other words, it "implies that the items stated are taken as examples from a larger group of items" (Chino 2001: 41), as shown in (1.15).

(1.15) *Watashi-no   heya-ni   wa,   konpyūtā-ya sutereo-ga   oitearimasu.*
       I-GEN   room-LOC   TOP   computer-YA stereo-NOM   place:STA:POL
       'In my room there is a computer, a stereo, and such.' (Chino 2001: 41)

Ya is attested only at the nominal level. In other words, it cannot be used to connect verbs and verbal phrases. Moreover, it cannot be used as a general extender with only one example: it must be used to connect at least two examples. For this reason, it is incompatible with situations where only one item is mentioned.

When it joins three or more items, *ya* tends often to occur only between the first two items, while the others are connected by means of a comma, as shown in (1.16).

(1.16) *Gasorin-ya   kenzai,*   *denka*   *seihin-nado-ga*   *fushin*   *da.*
       Gasoline-YA   building.material   electricity   good-NADO-NOM   stagnation   COP
       'Things like electrical products, building materials and gasoline aren't doing well.'

As for its distribution, *ya* can be used only between two items. In other words, typically *ya* does not follow the last item of the list[10]:

(1.17) (a)   *biiru ya sake*
       (b)   **biiru ya sake ya*

The connective *ya* is often combined with the general extender *nado* (see below) to reinforce its basic meaning, as shown in (1.16).

---

[10] The usage of *ya* alter the final element of the list (i.e., A *ya* B *ya*) is not completely forbidden, but extremely rare. As Martin (1975) noted, "written Japanese also overwhelmingly prefers A *ya* B over A *ya* B *ya*, which is largely limited to set phrases" (1975 :156). In our corpus, we did not find any instances of A *ya* B *ya*. It is likely that in Contemporary Japanese it has completely fallen in disuse.

1.3.2.2 nado

Nado functions as a synthetic general extender, that is, an independent morpheme occurring at the end of non-exhaustive lists. More specifically, it is attached immediately after the last example to indicate that "the item(s) mentioned is/are representative samples" (Kaiser et ali. 2001: 258), as shown in (1.18).

(1.18)   *Watashi*   *wa*   *Suisu-nado-ni*   *sumitai*   *desu.*

         I   TOP   Switzerland-NADO-LOC   live.DES   POL

         'I would like to live in a place such as Switzerland.' (Lee 2004:317)

     Nado cannot be used as a connective. Therefore, it cannot be repeated after each item of the list. Consequently, to provide a list of two or more items, *nado* must be used combined with *ya* as shown in (1.16), or – less frequently – with other non-exhaustive connectives (e.g., *tari* and *toka*). Otherwise, it can be used at the end of a list of juxtaposed items.

     Contrary to the connective *ya*, *nado* can be attached to noun phrases and verbal phrases, with the same meaning and functions.

     While in Classical Japanese *nado* was used mainly in quoted speech during narrative tales (cf. Yu 1995), it is widely accepted that in Contemporary Japanese *nado* is a feature of the written formal language (i.e., Chen 2005), and it is not widely used in spoken language.

1.3.2.3 tari

Tari is a converb used to join verbs or verbal phrases, as shown in the following example.

(1.19)   *Osaka-de*   *kaimono-o*   *shitari*   *kankoku-ryoori-o*   *tabetari*

         Osaka-LOC   shopping-ACC   do:TARI   Korean-meal-ACC   eat:TARI

         shimasu.

         do:POL

         'In Osaka, I will do such things as shopping and eating Korean food.' (Banno 2000:215)

     Typically, it indicates a range of actions performed by the same agent, but in some cases, it can also be used with different participants (cf. Alpatov 1997).

     At the structural level, *tari* is a converb attached to the *ta*-base of the verb, i.e. the form of verbs to which the past suffix *ta* is attached. It follows that the predicates suffixed by *tari*

are in a non-finite form. Consequently, the last *tari* should be followed by a conclusive *suru* "to do". This rule tends to be less restrictive when *tari* is used in spoken language.

Contrary to *ya*, *tari* can be used also with only one item, as a general extender. More specifically, Ohori (2004: 54) notes that in colloquial speech *tari* can be attached to a single clause, functioning as an utterance-final marker indicating that the event described by the verb is representative of several other events. For instance:

(1.20) *Tenki-no   warui hi   ni   wa,   ie   de   ongaku-o kiitari   shimasu.*
      weather-DET bad   day   on   TOP   home LOC   music-ACC listen:TARI do:POL
      'On days when the weather is bad I listen to music and do other such things at home.' (Chino 2001: 108)

In Contemporary Japanese, the main function of *tari* is non-exhaustive representation (cf. Chino 2001, Suzuki 1998). Nevertheless, some other pragmatic functions have been identified especially in spoken language (e.g., hedging function, see Taylor 2010, Suzuki 1998).

Contrary to *ya* and *nado* which are mainly used in written language, *tari* is widespread both in written and spoken language. Moreover, it also occurs both in formal and informal registers (cf. Taylor 2010).

1.3.2.4 *toka*

Toka can be used as a connective to joint two or more items as shown in example (1.21), but also attached to just one item as a general extender (like *tari*), as shown in example (1.22).

(1.21) *Koohii toka   koocha   toka   iroirona mono-ga   arimashita.*
      coffee TOKA   tea   TOKA   various thing-NOM   exist:POL:PAST
      'There were various things such as coffee and tea.' (Maynard 1990:106)

(1.22) *Yasumi-ni   wa   jogingu-o   suru   toka   shiteimasu.*
      vacation-LOC TOP   jogging-ACC   do   TOKA   do:STA:POL
      'When I am on vacation, I do things like jogging.'

Among the strategies examined in the present study, *toka* is the one that exhibits less structural constrains. For example, it can be used to join both nouns and verbs. Moreover,

it can follow each item of the list as shown in (1.23a) or it can be used only between two items, like *ya,* as shown in (1.23b).

(1.23) (a)   *koohi toka koocha toka*
        (b)   *koothi toka koocha*

Toka is widely used in colloquial speech especially by the younger generation. For this reason, it has been thoroughly studied in the research area of *wakomono kotoba* "young people's language use" (see Yamamoto 2004). It is hardly used in written language (except for the language of Internet), and even more rare in formal (written) language.

As noted by Taylor (2015: 143), it is likely that the contemporary marker *toka* is derived from two distinct markers (i.e., exemplifying *toka* and quotative *toka*) which exhibit different diachronic pathways. Exemplifying *toka* (connective and general extender) is usually interpreted as the combination of the comitative and conjunctive marker *to* (that is, "with" but also "and") and the indefinite/interrogative marker *ka*. Both elements contribute to the overall meaning of the suffix: "since *ka* generally conveys uncertainty and hence the possibility of choice, the composite morpheme *toka* means more than 'and'." (Ohori 2004: 51). On the contrary, quotative *toka* is derived from the quotative marker *to* (as in *to iu* "to say that"), and it normally functions as a hedge to indicate that the speech may not be verbatim (Ohori 2004: 53, see chapter 5).

In Contemporary Japanese, the distinction between exemplifying *toka* and quotative *toka* does not seem to be that clear, to the point that some studies (cf. Ohori 2004, Suzuki 1998) discuss them as different functions of the same maker *toka*.

Since the connection between exemplifying constructions and hedging strategies is well attested (cf. Mihatsch 2010b, Ghezzi 2013) and worthy of being further investigated, in our analysis, we will consider *toka* as a single maker (at least, at the synchronic level) which exhibits different functions.

2. DATA COLLECTION AND PARAMETERS OF ANALYSIS

2.1 DELIMITING THE OBJECT OF ANALYSIS: TOP-DOWN AND BOTTOM-UP APPROACH

As noted in the previous chapter, so far, we have adopted a top-down approach (cf. section 1.3). We have started our analysis by sketching the domain that we wish to map based on a predetermined conceptualization (cf. section 1.2). Then, we have reduced the domain to a set of constructions used to specifically express exemplification, according to our working definition (cf. section 1.3.2).

Nevertheless, as noted at the beginning of this work (section 1.1), our goal is also to use this set of linguistic constructions to better understand the underlying cognitive and linguistic mechanisms of exemplification. To achieve this, a top-down approach is insufficient.

Therefore, we now reverse the perspective by adopting a bottom-up approach. We will start out investigation from the individual constructions selected in the previous chapter by determining their functional range. We will do so exhaustively, by identifying every possible emerging function, regardless of its apparent relation with the main domain formulated by the top-down approach (i.e., exemplification to create context-relevant categories). In other words, there are no primary and secondary functions: all attested functions exhibit equal status (see De Haan 2010). More importantly, we will perform our investigation without strongly relying on background information to predict the functional range of the constructions under study. Consequently, we will look at individual functions and then move from them to understand the entire functional domain of exemplification.

De Haan (2010) highlights the differences in these approaches by pointing out that while a top-down approach leads to questions such as "to what category [domain] does linguistic element X belong?", the bottom-up approach leads to questions like "what is semantic range of the linguistic element X?".

In our specific case, we will look for the functional ranges of *ya, tari, toka* and *nado*. Then, after their functional ranges are established, we will compare them, working our way up to the domain level.

Therefore, the object of this research is twofold. First, we aim to analyse the usage of exemplification to construct online context-relevant categories. In this sense, the bottom-up approach will help us to move from the linguistic evidences to the cognitive functions, without

imposing pre-existent theoretical frameworks. Secondly, this study is also meant to sketch the functional space of exemplification in its entirety, to better understand its linguistic domain and identify possible further connections with cognitive mechanisms. In this regard, the bottom-up approach requires to build and piece together several smaller functional spaces to give rise to a larger and more complex one, that represents the domain under study.

2.2 DATA SELECTION

Most studies on Japanese exemplifying strategies have relied solely on invented and de-contextualized sentences, without considering the influence of the surrounding context and the possible discursive goals that speakers can attempt to achieve by using them (cf. Taylor 2010).

Since the purpose of this study is to investigate the actual functions of these constructions in Contemporary Japanese with a special focus on their ability to create and communicate context-relevant categories, it is necessary to observe these constructions in large bodies of authentic discourse data. Moreover, it is also crucial to have access to the broader context to monitor the contextual environment of the utterance.

To address these methodological issues, a corpus-driven approach has been adopted and a corpus of Contemporary Japanese has been used. Corpus linguistics involves the study "of language based on examples of 'real-life' language use" (McEnery and Wilson 1996: 1). Therefore, it is a methodology in which the analysis of language is based on contextualized, naturally-occurring data, rather than on artificial or "made-up" data (Meyer 2002: xiii).

More specifically, the considerations we are going to discuss are the result of a corpus-driven methodology (Tognini-Bonelli, 2001), which differs from corpus-based approach since it is characterized by having no pre-defined domains to be searched for. Moreover, this methodology allows us to observe and investigate new categories emerging from the corpora, since it is also free of "pre-tagged texts" (Sinclair 2004: 191). In our specific case, it enables us to identify the varying functions of Japanese exemplifying constructions without imposing pre-empirical intuitions to the data.

Particularly, we have selected one corpus of Contemporary Japanese which provides us with the access to the broad context of the occurrences[11] to be analysed, namely the Japanese section of the Leipzig Corpora Collection (LCC). Further reasons for this choice will be clarified in section 2.2.1.

Finally, selected data from the corpus have also been further investigated and verified by means of questionnaires. As it will be explained in detail in section 2.2.2, the questionnaire consisted of 20 sentences containing at least one exemplifying construction at issue, whereby native speakers have been asked to evaluate the appropriateness of a certain number of paraphrases for each sentence (cf. section 2.2.2).

2.2.1 THE JAPANESE SECTION OF THE LEIPZIG CORPORA COLLECTION (LCC)

The Leipzig Corpora Collection (LCC) is a collection of corpora of comparable sources and equivalent processing for more than 250 languages (cf. Goldhahn *et ali*. 2012). The Leipzig Corpora Initiative was started during the 1990s at the University of Leipzig, using a variety of different sources to aggregate text material and supplementary data (cf. Eckart and Quasthoff 2013). The central resource used is the World Wide Web. More specifically, the major sources of texts are: 1) online newspapers, 2) randomly crawled texts from the WWW, 3) Wikipedia. If needed, supplementary data include also movie subtitles, chat room messages and Twitter messages.

The Japanese section of the LCC (jpn_news_2005-2008) consists mainly of newspaper texts and some randomly collected web pages from 2005 to 2008 and it contains 58.407.729 tokens. Due to the huge amount of data, we have analysed 200 occurrences for each Japanese exemplifying construction under analysis (for a total of 800 occurrences) taken from random samples.

This corpus has been selected for several reasons. First, for each sentence, it provides a clear and easy link to the main source, in order to monitor not only the immediate surrounding co-text, but also the entire contextual environment of the occurrence. Consider the amount of data and the main scope of our analysis, the ease to retrieve the broader context becomes a pivotal factor.

Secondly, it is based on online newspapers. As it was shown in the introduction (cf. section 1.2.2.2), recent studies on communicative theory have addressed the cognitive

---

[11] For this reason, we have been forced to exclude the use of corpora such as the Balanced Corpus of Contemporary Written Japanese (BCCWJ, at Kotonoha) which did not allow us to have easy and instantaneous access to the broader context of the sentences.

value of exemplification in informative discourse (cf. Zillmann and Brosius 2000). More generally, exemplification is an essential instrument whenever people need to explain, describe, and inform, such as in the context of journalistic articles. For this reason, we believe that a corpus based on newspaper articles can be a breeding ground for a research on the role of exemplification, and therefore, at the linguistic level, of exemplifying constructions.

Third, while most available texts are newspaper articles, the corpus still exhibits a significant amount of heterogeneity both in style and register. The variety of the texts it contains ranges from more traditional articles composed following the formal style of written language, to blog posts written in a borderline variety of language showing hybrid properties associable to both the speech discourse and to the written language, to even transcripts of interviews which can be considered as (naïve) transcriptions of spoken language. Moreover, it also includes 1) newspaper editorials, 2) some instances of online message boards, and 3) Twitter messages (cf. also chapter 5). This variation allows us to investigate exemplifying constructions in different communicative situations.

As for the topics covered by the texts, the corpus tends to suffer from a lack of variation. While it is possible to find articles facing all sorts of issues (from a banking expert reporting advice on mortgages, to a school girl describing her issues at dating), it is also true that a large group of texts focuses on computer science and technological devices. While this can constitute a bias in the internal variation of the corpus, we believe that it is (at least partially) minimized by the variation in styles and registers. For instance, in our corpus there are several reviews of technological devices. However, while some of them are standard newspaper articles, with a formal register and long dry lists of product details, others are much more informal, reflecting speech patterns as if the writer is having a direct conversation with the reader. Consequently, despite the similarity in topics, they can still be considered different types of texts.

While it is important to address the bias that may arise from similarity in topics, we argue that register and style are the major factors in determining different patterns of usage. This fact is backed up for instance by the quantitative study of Taylor (2010) on some Japanese exemplifying constructions (i.e., *nado, toka* and *tari)*. He notes that the higher variation is triggered by different styles and registers (e.g., spoken language vs. written language, formal language vs. informal language, private conversations vs. public conversations), rather than by topic variation.

Finally, we have selected this corpus also for future research purposes. We firmly believe that this case study on Japanese is just the first step towards a more comprehensive study on exemplification and exemplifying constructions. In particular, the investigation of the functional space of exemplification needs to be examined and tested considering other languages. For this reason, we hope that this type of research will be repeated (inasmuch as that is possible) for other languages, even adopting a cross-linguistic perspective. Therefore, the usage of a corpus which is part of a collection of comparable corpora may provide a good starting point for future (comparable) research.

2.1.2 THE QUESTIONNAIRE

The questionnaire used to investigate Japanese exemplifying constructions is a multiple-choice questionnaire, whereby native speakers have been asked to evaluate possible (or impossible) paraphrases of sentences from our corpus containing at least one exemplifying construction (see the questionnaire in Appendix B). The context has been made clear in all sentences by providing a direct link to the source on the Internet, so that the various contextual clues could be clearly identifiable.

Sentences have been chosen from our corpus on the ground of being interesting with regards of one or more parameters of analysis (with a special focus on exhaustivity, cf. section 2.3.2.4). For each sentence, several paraphrases have been provided. Paraphrases have been written to be similar to each other in any aspect except one involving the parameter to monitor (e.g., exhaustive and non-exhaustive versions of the same paraphrase).

For each sentence, the speaker has been asked to evaluate the appropriateness of the paraphrases on a scale from 1 (not appropriate) to 3 (very appropriate). Moreover, at the end of each sentence, a blank space has been left for further comments. In case of disagreement towards all the paraphrases proposed, the native speaker has been asked to provide the most appropriate paraphrase according to him or her.

The questionnaire has been filled out by speakers with a good education level (PhD students or Master's degree students), whose proficiency in the language of the questionnaire is very good. Two versions of the questionnaire have been prepared, one in English and one in Japanese.

As it often happens (cf. Bernini and Ramat 1996), the questionnaire has been adjusted as the various feedback has been received, in order to make it more coherent with the aims of this study. At the beginning, the questionnaire was only written in English to avoid any

possible influence from Japanese. The decision to translate the questionnaire in Japanese has been made to facilitate those speakers who were not proficient in English. The Japanese version of the questionnaire has been compiled trying to avoid any possible bias due to the language, and with the help of two native speakers.

The questionnaire has been completed by a total of 35 respondents. Out of this 35, 6 native speakers completed the English version of the questionnaire (mostly PhD students) and 29 native speakers completed the Japanese version (mostly Master's degree students).

The questionnaire proved to be particularly suitable for 1) monitoring some parameters which strongly depend on the sensitivity of native speakers, such as exhaustivity, and 2) distinguishing instances of exemplifying constructions used to create context-relevant categories and instances performing pragmatic functions. As for the former, the evaluation of native speakers was essential in all those sentences where both an exhaustive and a non-exhaustive interpretation were possible for lack of other linguistic evidences (cf. section 2.2.2.5). As for the latter, since politeness strategies are often cultural-dependent and the author (whose mother tongue is Italian) does not always have the cultural sensibility to distinguish between nuances of meaning, the sensitivity of native speakers has been an essential factor to correctly identify and interpret instances of exemplifying strategies performing pragmatic functions.

Despite the utility of having a tool to check the appropriateness of different interpretations, we are aware that the use of questionnaire entails some issues that may cause methodological problems (cf. Bernini and Ramat 1996). For instance, even though all informants have a good level of education, the evaluation of the paraphrases basically depends on their sensibility as native speakers. This is especially true with regards to those sentences where informants have been asked to evaluate subtle nuances of meaning. Moreover, the fact that data do not come from spontaneous interactions, but are based on written text may have influenced some of the evaluations. Informants may have deemed some paraphrases as not appropriate in the written register, excluding from the survey their appropriateness at the oral conversational level.

These difficulties may lead to consider the use of questionnaires as unreliable to some extent. Nevertheless, it is the only realistic method to monitor some linguistic parameters while avoiding direct questions which in turn may elicit unnatural and non-spontaneous answers. For this reason, while considering the potential biases underlying the evaluations, we still believe that the questionnaire is an important tool to integrate data coming from other sources.

Furthermore, in several occasions, the questionnaire has been the starting point for further discussions, in order to monitor - albeit as much indirectly as possible - the underlying reasoning of the native speaker.

The aim of questionnaire collection has thus been to verify information coming from grammars and from our corpus data. Therefore, the corpus-based method and the questionnaire have been jointly used to pursue the same goal, that is, to identify the different functions of Japanese exemplifying constructions in actual communication.

2.3 PARAMETER OF ANALYSIS

The bottom-up corpus-driven approach adopted in this part of the study strongly influences the choice of the parameters of analysis, leading to the exclusion of criteria that would pre-impose functional distinctions to the data. Therefore, each occurrence will be examined on the basis of parameters comprising morphosyntactic and distributional features of the exemplifying constructions (cf. sections 2.3.1, 2.3.2 and 2.3.3) and textual features of the linguistic environment (cf. sections 2.3.3 and 2.3.4).

Some parameters have been selected to identify different functions performed by exemplifying constructions, such as the exhaustivity feature (to distinguish between instances of categorization and instances of exemplifying strategies used in exhaustive contexts, see chapter 6) and the topic continuity in discourse (to establish whether the main purpose of the speaker is to communicate a category or not, see chapter 5). Other parameters have been selected to analyse the usage patterns of Japanese exemplifying constructions, such as the morphosyntactic, distributional and textual properties. In some cases, parameters may provide useful information on both cases. For instance, category labels (cf. section 2.3.1) not only are evidence of ongoing categorization (therefore excluding other functions), but also contribute to the analysis of exemplifying constructions, because they interact in the same process.

In the following sections the collection of the parameters will be further discussed and explained[12].

---

[12] For the sake of clarity, the parameters will be explained by using mainly examples from English (invented by the author or taken from the enTenTen13 corpus, Sketch Engine). Nevertheless, whenever the parameters are related to language-specific issues, examples form Japanese (from our corpus or other papers) will be used.

2.3.1 CATEGORY LABEL

Generally speaking, a category label (or category name) is a word or a linguistic expression through which people specifically designate a conceptual category. Unfortunately, much of the existing literature is vague in providing a concrete working definition of category label, merely describing it as a word or short expression associated with a stable cognitive representation, that is, the category (cf. Rosch 1973, Channel 1994, Taylor 1995, Overstreet 1999).

This characterization is not enough to distinguish instances of category labels from instances that should not be considered as such. Therefore, it is crucial to establish a much more precise working definition which can be used in our analysis.

Intuitively we can define a category label as a linguistic expression whose main purpose is to classify a set of items. A given linguistic expression will be considered as a category label if 1) it is a sufficiently generic term, to which the explicit exemplars can be traced back, 2) it provides some semantic clue towards identification of the defining property of the category. This definition excludes placeholders such as *things* because they do not provide any semantic specification regarding the property shared by the category members. Consider the following examples:

(2.1)   *My house is full of things like lipstick, mascara, eyeliners and so on…*

(2.2)   *I like to collect makeup products. That's why my house is full of lipstick, mascara, eyeliners and so on...*

In (2.1), the word *things* certainly indicates a set of elements, but it does not provide any semantic clue to identify the defining feature of the category. For instance, it can be used both with categories of entities and categories of events, since it does not specify any semantic property of the exemplars. According to our definition, terms like *things*[13] are not proper labels, but just placeholders. Consequently, the Japanese words *koto* "things" and *mono* "things"[14] have been excluded.

---

[13] It is noteworthy that in some languages other than Japanese, unspecific terms like *things* crystallized into analytical general extenders (e.g., "and things like that" in English, *e cose così* in Italian) to express merely non-exhaustivity.

[14] While it is true that *mono* "things" tends to designate only (concrete) objects, in our corpus (cf. section 3.6), it is always used with abstract entities, performing a function much similar to that of a placeholder. For this reason, it was not considered as a category label.

On the other hand, in (2.2) the expression *makeup products* not only implies the presence of a set to which the mentioned examples can be traced back, but, at the same time, it also highlights effectively the defining property that all members of the category share.

Following our working definition, we exclude also lexical items which do not provide information regarding the defining property, but indicate other types of specification, for example the semantic relation among the items. Consider the following example:

(2.3)   *For example, [the times] when you have a headache or you are stressed.*

(2.4)   *There were [cases] of people being unable to obtain their medications or people being refused services and treatment.*

Terms such as *times* and *cases* have been excluded as well since 1) they are too unspecific in their semantic reference, and 2) they simply signal that the members of the category are all "times" or "cases", that is, alternatives that may happen in different situations. In other words, they do not specify what kind of times/cases the members actually are, but just how they relate to each other. For this reason, they have been considered as placeholders and not proper labels. Consequently, Japanese nouns such as *toki* "times", *kēsu* "cases", *baai* "situations" have been excluded.

It is noteworthy that placeholders can be considered a proper labels as long as they are attached to other linguistic adjuncts that provide a higher semantic specification.

(2.5)   *Things like hiking and climbing.*

(2.6)   *Things to do during a trip on the mountain, like hiking and climbing*.

In (2.5), the word *things* is not a label, but a placeholder. However, in (2.6) *things to do during a trip on the mountain* is a proper category label.

2.3.1.1 Presence of the category label

This parameter distinguishes between *lexicalized* and *non-lexicalized* categories. The former is characterized by the presence of an explicit category label in addition to the examples as shown in (2.7). The latter is characterized by its absence as shown in (2.8).

(2.7)   *Small animals such as dogs, rabbits and so on.*

(2.8)   *Dogs, rabbits, and so on*.

In the case of non-lexicalized categories, the defining property of the category needs to be inferred from the situational context and from the semantic properties of the mentioned examples. On the contrary, when a category label is provided, the property is lexicalized (at partially[15]) through the label itself, as shown in (2.9).

(2.9)   *Relaxing drinks such as water, herbal teas, smoothies and such.*

Here, the label indicates that the speaker is referring only to drinks that help people to relax. Therefore, the hearer may exclude other types of drinks, such as coffee or soft drinks.

As for the role of placeholders, we will consider as non-lexicalized categories those occurrences in which the label consists only of a placeholder as in (2.10), and as lexicalized categories those occurrences in which the label consists of a placeholder attached to other linguistic adjuncts as in (2.11).

(2.10) *Things like hiking and climbing.* [non-lexicalized category]
(2.11) *Things to do during a trip on the mountain, like hiking and climbing.* [lexicalized category]

2.3.1.2 Syntactic types of labels

All the attested labels are examined on the basis of their syntactic structure.

As already noted, for many decades, category labels have been described as short conventional linguistic means, such as general simple nouns (e.g., "fruit", "furniture") or small phrases (e.g., "alcoholic drinks", "grocery stores"), without any other specification about their syntactic structure.

The shortcomings of this model emerged even within the analysis of natural categories alone. Cruse (1986) highlights that not every level of the vertical dimension of categories can be expressed by means of a single word or a short expression. These lexical gaps can be covered only through complex expression (e.g., furniture on which people can sit). Additionally, also the identification of ad hoc categories emphasises the possibility of using complex expressions as labels to designate specific context-relevant sub-sets. In fact, despite being often studied as instances of non-lexicalized categories (cf. Channel 1994, Overstreet 1999), ad hoc categories can be designated by means of linguistic expressions

---

[15] As it will be made clear in section 3.2.1, the presence of a category label does not exclude the necessity of inferential enrichment, because, in some cases, the property may be lexicalized only partially (e.g., simple labels). Therefore, while the category label represents an important clue to correctly identify the property shared by category members, the key role of the context cannot be excluded completely.

that consist of a nonspecific superordinate noun (generally, *things*) and an infinitival purpose clause (*to do X*), such as in "things to take on a camping trip" (Overstreet 1999: 42).

Starting from the definition formulated in the previous sections (which is deliberately very unspecific regarding the syntactic parameter), any instance of proper label is considered and examined on the basis of its syntactic structure. First, we will distinguish category labels expressed by simple nouns (2.12) and category labels expressed by noun phrases. Then, the latter is further subdivided into: 1) noun phrases containing adjectives (2.13), 2) compounds (usually N + N) (2.14), 3) noun phrases containing genitive clauses (2.15), 4) noun phrases containing relative clauses (2.16), 5) noun phrases consisting in two or more adjuncts (2.17).

(2.12) *Animals, such as rabbits*.
(2.13) *Small animals, such as rabbits.*
(2.14) *Country animals, such as rabbits*.
(2.15) *Animals of the Chinese zodiac, such as the rabbit*.
(2.16) *Animals that lives in countryside, such as rabbits*.
(2.17) *Small animals that lives in countryside, such as rabbits*.

2.3.1.3 Semantic properties of the label

Beyond their syntactic structures, all the attested labels will be examined also on the basis of their semantic properties, specifically general-specific relations. This strongly relies on the semantic notion of hyponymy (cf. Lyons 1968, Cruse 1986). Specifically, in semantics, hyponymy is a relation of inclusion where X is a kind/sort/type of Y. In the words of Lyons (1968: 453), "...the meaning of *tulip* [or *rose* or *geranium*] is said to be 'included' in the meaning of *flower*.". Therefore, the term 'tulip' is a hyponym with respect to 'flower'. From the point of view of categorization, it means that the word 'flower' can be seen as a label for the category that encompasses members such as 'tulip' or 'rose' or 'geranium'.

These semantic relations have been organized and schematized by means of semantic taxonomy hierarchies (cf. Cruse 1986). Consider the following schema taken from Cruse (1986: 136):
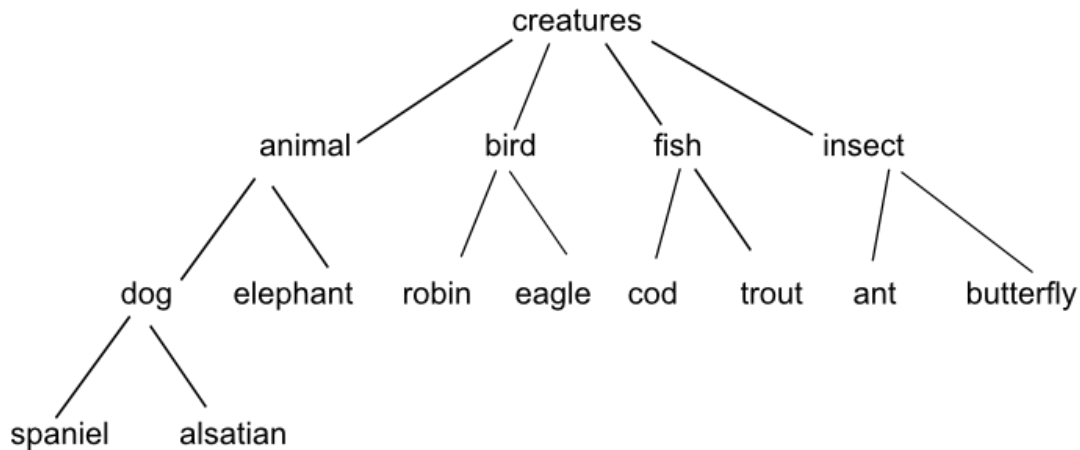
Figure 2.1: Semantic taxonomy hierarchy

If we want to refer to the category whose members are 'spaniel' and 'alsatian', the more specific label (as the more characterizing) would be 'dog', while 'animal' would be considered general, and 'creatures' even more general. Therefore, the more specific label is the one on the immediately higher level (e.g., 'dog'). On the contrary, as we move higher up the hierarchy, labels become more general and inclusive (i.e., 'animal' is more general then 'dog' but more specific than 'creature'), encompassing more heterogeneous groups of items. This specificity of the daughter-nodes with respect to mother-nodes is given by the relation of inclusion: "we may say that the meaning (sense) of *apple* is richer than that of *fruit* and includes, or contains within it, the meaning of *fruit*" (Cruse 2000:150-151).

Ideally, people would use labels that make precise reference to the category they have in mind, therefore, the one in the immediately higher level than the category members. However, this immediately higher level is not always available, since a specific superordinate noun may do not exist in a particular language. Consider the schema below:
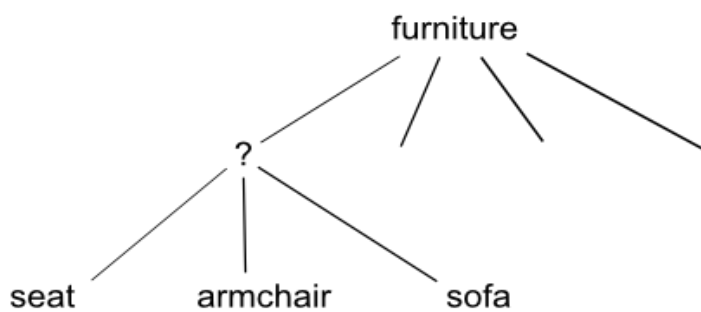
Figure 2.2: Semantic taxonomy hierarchy of 'furniture'

In English, there is no superordinate noun that functions as a label for the category of pieces of furniture on which people can sit, like 'seat', 'armchair' or 'sofa'. Therefore, in order to cover this gap, the speaker can decide to 1) create a syntactically complex label, starting from the general label available and adding linguistic adjuncts (e.g., 'furniture on which people can sit') or 2) use the closer general label available and then specify it through examples and/or contextual elements. In the first case, the label is specific and syntactically complex, in the second, the label is general but syntactically simple.

For our analyses, we will distinguish between general labels and specific labels following the methodology provided by semantic taxonomy hierarchies. In this regard, it is noteworthy that the general-specific semantic relation is a relative and not an absolute concept, because it strongly depends on the members of the category. Consider the following examples:

(2.18) *Animals such as alsatians and spaniels*.
(2.19) *Animals such as dogs and elephants*.

The noun 'animal' is a generic label when it is used to designate a category that encompasses 'spaniel' and 'alsatian'. Yet, it is a specific label when it refers to a category that covers 'dogs" and 'elephants'. Therefore, unlike the syntactic parameter, the analysis of the semantic properties of the category label will also consider the mentioned examples.

The investigation of the syntactic and semantic properties of the labels would allow us 1) to better understand how speakers lexicalize (i.e., create and use category labels) categories in real speech situations, and 2) to identify possible tendencies in the usage of category labels depending on the type of categories (i.e., categories of entities, categories of events, categories of properties, etc., cf. section 2.3.2.1).

2.3.1.4 Position of the category label

Category labels will be examined also on the basis of their location inside the utterance, with respect to the mentioned examples. This parameter is analysed along two dimensions: 1) the label can precede or follow the example(s), 2) the label can be part of the exemplifying construction, that is, directly linked to the example(s) by means of linguistic connectors (e.g., *such as, like*) or it can be external to it, that is, it can occur before or after the examples without being connected directly to them. The intersection of these two dimensions allows us to identify four overall values for this parameter: 1) inside precedent (2.20); 2) inside posterior (2.21); 3) outside precedent (2.22); 4) outside posterior (2.23).

(2.20) *Relaxing drinks such as water, herbal teas, smoothies, and such*.

(2.21) *Water, herbal teas, smoothies, that is, relaxing drinks*.

(2.21) *When I am very nervous I drink a lot of relaxing drinks. You know, water, herbal teas, smoothies, and such.*

(2.23) *Water, herbal teas, smoothies, and such are useless in the morning. I don't need relaxing drinks, I need just coffee to stay awake*.

In Japanese, the inside precedent pattern is structurally impossible because of the order of constituents. In fact, to link directly the label to the example(s), the label must be added at the end of the construction as shown in (2.24). The label can be placed before the examples only if it is not directly attached[16] to them as shown in (2.25), that is, following the outside precedent pattern.

(2.24) *Kiniro-ya   giniro **nado** toitta   torofī-no   iro*
       Gold-YA   silver **NADO** such as   trophy-GEN   colour
       "Colours of the trophies such as gold, silver and so on"

---

[16] Examples and labels can still be in the same sentence or utterance. The only constrain is not to be directly connected.

(2.25) *Kigyō*   *wa*   *seitōna*   *riyū-ga*   *areba,*   *naitei-o*   *torikesu*

company TOP legitimate   reason-NOM   exist:COND   unofficial offer-ACC   revoke

koto-ga   *dekiru.*

thing-NOM   do:pot

Tatoeba,   *naiteisha-ga*   *naitei-go-ni*   *hanzai-o*   *okashi**tari**,*

For example, nomenee-NOM   offer-after-LOC crime-ACC   perpetrate:**TARI**

gakureki-o   *sashō*   *shite*   ***itari***   *shita*

academic-background-ACC   false.statement   do:GRD   exist:**TARI**   do:PAST

baai   *da.*

case   COP

"If there are any legitimate reasons, companies can cancel the job offer. For example, in case the nominee commits a crime after the nomination or in case he makes a false statement about his educational background"

It is important to note that the attribution to the values inside/outside is strictly related to the presence of a direct linguistic connector between the label and the example(s). For this reason, we decided to consider instances like (2.26) as "outside posterior".

(2.26) *Blockbuster-ya*   *eBay*   **nado,** *iroirona*   *kyōryoku*   *saito*

Blockbuster-YA   eBay   **NADO,** various   collaboration   website

"Various partner websites, such as Blockbuster and eBay"

Even if *nado* sometimes can used as a connector (see section 3.4), in (2.26) there is a pause, that is, a comma between *nado* and the label. The presence of the comma places us in a difficult position because we cannot automatically assume that the label and the examples are part of the same construction and that *nado* is indeed the linguistic connector. Therefore, we decide to monitor this type of construction by assigning the value "outside posterior" and annotating the presence of the comma. The reason for this choice is to monitor potential differences in the linguistic coding between instances without the comma and those with a pause like (2.26).

2.3.1.5 Linguistic links between label and example(s)

Finally, we will investigate linguistic links between labels and examples. In other words, whenever the label is directly linked to the example(s), there must be some sort of linguistic marker that highlights the relationship between the two components of the construction. In other words, we will investigate linguistic constructions that express the relation "X is an

example of Y", or more generally, "X is included in Y". For instance, consider the following example from Japanese:

(2.27) *Saru-ya*   *tanuki-ya*   *inu-ya*   *kitsune toitta*   *dōbutsu.*

Monkeys-YA   raccoons-YA dogs-YA   foxes   such as   animals

'Animals such as monkeys, raccoons, dogs and foxes.'

Here, the linguistic construction that connects the category label (i.e., *dōbutsu "animals"*) with the examples (i.e., *saru* "monkeys", *tanuki* "racoons", *inu* "dogs", *kitsune* "foxes"), is *toitta* which can be glossed as *such as*.

Linguistic constructions that express hyponymy have been thoughtfully studied by semanticists. For instance, Cruse claims that "A useful diagnostic frame for taxonomy is: An X is a kind/type of Y" (1986: 137). Here, X is a hyponym of Y, like in "a rose is a type of flower". From the perspective of categorization processes, X is an exemplar of Y, and Y is the label of the category that encompasses exemplar like X. The same linguistic structure has been studied by Lyons (1977: 292-293).

Ahmad and Fulford (1992: 13) investigate a list of English constructions to indicate hyponymy and tested some of them for effectiveness on a corpus. Pearson (1996: 820) labels these structures "hinges" and highlights different hinges depending on the position of X and Y respectively. Some examples of hinges are: X *is/are*, X *consist(s) of*, X *is/are defined as*, and *is called a* X.

Borillo (1996: 113) proposes a study on French to "voir comment la relation d'hyperonymie se manifeste linguistiquement dans les textes, i.e. sous quelles formes lexico-syntaxiques elle trouve à s'exprimer" (1996: 114). The analysis reveales a wide variety of French constructions, including (but not limited to) the following:

(2.28) (a)   Na *est un* NX [Na is the hyponym and NX is the hypernym]

      (b)   Na *est un sorte/espèce/type/variété de* NX

      (c)   Na *est un* NX *très* Adj

      (d)   Na *et autres* NX, Na *et tout autre* NX

      (e)   NX *et plus particulièrement* Na

      (f)   Na, *le* NX *le plus* Adj

      (g)   NX *comme (par exemple)* Na [similative and exemplifying constructions]

Hearst (1992) conducted a similar study for Englísh starting with the following list of constructions: *such as*, *or other*, *and other*, *including*, and *especially*.

From the perspective of categorization processes, Italian similative constructions (e.g., Y *come* X, where Y is the label, X is the example) have been briefly studied by Barotto and Mauri (2016) as linguistic structures to directly connect a category label to the example(s), like in *cani come i bassotti*.

To identify these constructions, we will follow a bottom-up corpus-driven approach: we will check and investigate the linguistic constructions attested in our corpus. Moreover, we will also signal possible systematic patterns in the selection of the linguistic material to form linguistic connectors in Japanese. The aim is to sketch a preliminary typological survey of these linguistic constructions while analysing the intra-linguistic variation as it is attested in our corpus data.

2.3.2 ITEM(S)
2.3.2.1 Semantic properties of the example(s)
The mentioned examples will be examined on the basis of their semantic properties. The aim of this parameter is to depict tendencies in the selection of exemplars to create categories.

Most cognitive studies (cf. Rosch 1973, 1975, Barsalou 1983) tend to focus solely on categories of (concrete) objects, while ignoring the possibility of creating categories of actions or properties. For example, Rosch herself describes a category as "a number of objects which are considered equivalent" (Rosch et al. 1976: 383). This is particularly true for natural category studies whose main concern is the way in which we categorize concrete objects that surround us (from animals to inanimate objects). Not even the introduction of ad hoc categories by Barsalou (1983) has changed this approach: while a more abstract level is not excluded *a priori*, the focus remains on concrete objects. Therefore, our aim is to investigate this cognitive fact through a linguistic parameter, specifically the semantic properties of the mentioned examples.

Lexical items can be categorised into three basic semantic classes (Croft 1991, 2001): 1) objects, 2) actions and 3) properties[17]. Beyond them, since in cognitive psychology special attention has been paid to concrete objects, the semantic class of objects (entities in our

---

[17] To avoid potential misunderstanding with further sub-divisions (e.g., regarding the concreteness parameter), we use the more neutral label 'entities' instead of 'objects' (cf. Givón 2001: 56), and the label 'events' instead of 'actions'.

analysis) will be further subdivided on the basis of the parameters of concreteness and animacy. Our analysis follows the markedness hierarchy proposed by Givón (2001: 56):
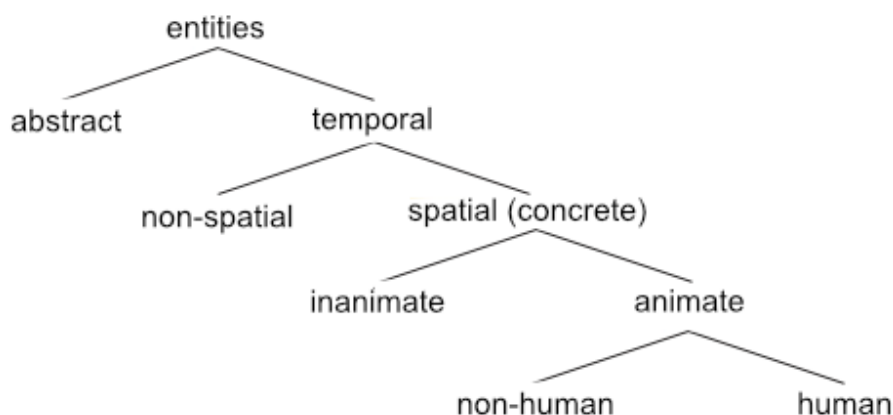
Figure 2.3: Markedness hierarchy of major semantic features of events (Givón 2001: 56)

As can be seen in the figure above, concreteness can be investigated with more precision through two hierarchic features: temporality and spatiality. Abstract entities such as 'peace', 'freedom', 'whiteness', 'refusal', etc. exist neither in time not in space. Temporal entities such as 'May', 'day', 'Monday', 'anniversary', etc. exist in time but not in space. On the contrary, fully concrete entities (or objects) such as 'bottle', 'tree', 'woman', 'sun', 'earth', etc. exist in both space and time[18]. Furthermore, concreteness can be further specified with regards to animacy and/or humanity. That is, concrete entities may be also animate; otherwise, they are considered by default as inanimate. Finally, concrete animate entities can be human or non-human (e.g., animals).

In our analysis, we are mainly concerned with three sub-divisions: 1) abstract entities which do not exist in time and space, 2) concrete objects which do exist in time and space (i.e., the focus of cognitive psychology), 3) animate entities, as a sub-class of concrete objects. This further subdivision will allow us to have a clearer picture on the items used by speakers to build categories and potential related tendencies. For instance, consider these few examples:

(2.29) Category of events:
       You can [read a book, make a drawing or something].

---

[18] In our analysis, the class of concrete objects encompasses also entities that exist in the virtual space. This means that entities that refer to the Internet and/or computer science such as 'videos', 'software', 'comments', 'forum', etc. are considered as concrete objects, since they exist in time and in (virtual) space.

(2.30) Category of properties:

　　　　He is very [shy and modest and such].

(2.31) Category of concrete objects:

　　　　[Water, herbal teas, smoothies, and such] are useless in the morning.

(2.32) Category of animate entities:

　　　　I would love to have a [small rabbit or a hamster or something similar].

(2.33) Category of abstract entities:

　　　　He doesn't show [love, hate, anger and so on].

2.3.2.2 Syntactic properties of the example(s)

The mentioned examples are also examined on the basis of their syntactic properties. The aim of this parameter is to depict tendencies in the selection of exemplars (at the cognitive level), but also in the linguistic coding of the examples (at the linguistic level). Since syntax is concerned with the rules and the processes that govern the structure of sentences, it may reveal correlations between the way we process examples at the cognitive level and the way we encode them at the linguistic level.

Croft's (1991) cross-linguistic study on grammatical categories establishes three typological prototypes which can be called 'noun', 'verb' and 'adjective', by pointing out at the relationships between the three major semantic classes (i.e., objects, actions and properties) and the propositional act functions with which they are used in particular constructions. Specifically, the three major propositional acts are: 1. reference, 2. modification (of a referent), 3. predication (i.e., ascribing something to a referent). The hypothesis of Croft is that "the semantic classes of OBJECTS, PROPERTIES, and ACTIONS are the TYPOLOGICAL PROTOTYPES of referring, attributive, and predicating constructions respectively" (Croft 2001: 87). Therefore:

(2.34) (a)　　noun = reference to an object

　　　(b)　　adjective = modification by a property

　　　(c)　　verb = predication of an action　　　　　　　　　　　(Croft 2001: 89)

These are the typologically unmarked combinations, but other combinations are still possible, albeit marked (e.g., action words to reference, object words to predicate, etc.).

In our analysis, we are interested in both marked and unmarked situations, with a slight preference towards the former. As it was pointed out by different scholars (cf. Croft 1991, Langacker 1987a, 1987b, 1991b, Givón 2001), the choice of encoding concepts in a marked

way influences how language users process those concepts. This may be explained in terms of a typological universal identified by Croft (1991) called *behaviour potential*, which refers to the range of the grammatical power of different combinations. This universal states that the range of grammatical behaviour of unmarked combinations will be at least as wide as, if not wider than, that of marked combinations. This excludes the case of marked members having more inflectional possibilities than unmarked members. The consequence is that the behaviour of marked combinations is often impoverished compared to that of unmarked combinations. For instance, when property words and action words are encoded by nouns rather than by adjectives or verbs, they lose their reference to their gradable nature (in the case of properties) or to the distribution through time (e.g., tense, in case of actions). This intrinsically affects the way people elaborate these words (cf. Langacker 1991b). Because the speaker deliberately decides how to encode a specific concept (and therefore also the examples), we argue that instances of marked encoding may be related to the cognitive process of building conceptual categories. Therefore, we will consider separately the semantic and the syntactic parameters, to detect any possible discrepancy and monitor tendencies regarding the encoding of different types of examples in categorization processes. For instance, consider these few examples:

(2.35) Categories of events expressed by verbs
If you need to relax, you can [sleep, drink a tea, read a book or something similar].
(2.36) Categories of entities expressed by nouns
I need [flower, eggs, butter and so on.]
(2.37) Categories of events expressed by nouns
[Smoking, drinking and so on] are bad habits.
(2.38) Categories of properties expressed by nouns
I was scared it would make the [redness etc.] worse.

To sum up, through the investigation of the semantic and syntactic parameters, we hope to address not only cognitive issues such as what types of examples are more easily selected by language users (and consequently what types of categories are more frequently built), but also linguistic issues regarding the linguistic coding of exemplifying constructions. The latter will be considered also for its effects at the cognitive level, regarding how the examples are ultimately elaborated.

2.3.2.3 Number of examples

We will investigate the number of examples mentioned by the speaker. Specifically, whether the speaker provides one single example as shown in (2.39) or a list of examples as shown in (2.40).

(2.39) One example

    I need to buy [milk and stuff].

(2.40) List of examples (two or more)

    I need to buy [milk, flower, eggs and stuff].

While there are no actual structural constrains[19] on the number of examples to provide in discourse, there are several reasons for monitoring this parameter.

We have seen already that examples need to be compared to infer their common property (cf. associative reasoning, section 1.3.1). This has consequences on the number of examples provided, because if the speaker provides only one example, the hearer is forced to infer the property elsewhere (cf. Perelman and Olbrechts-Tyteca 1969). More specifically, he or she must deduce the most relevant feature of the single example looking at the specific situation, by comparing the example to the multi-dimensional context. This process may require a greater cognitive effort showing consequences also on the encoding of the single example.

Moreover, comparison is a dynamic process. As noted by Perelman and Olbrechts-Tyteca (1969: 354), whenever new examples are provided, not only should they be interpreted in the light of those previously mentioned, but they also adjust the reference generated by the other examples. It follows that the inference of the common property is a complex and dynamic process and it may be influenced by the number of examples provided by the speaker.

In addition, while there are no actual structural constrains on the number of examples, it is also true that communication is ordered by a principle of linguistic economy, which consists in tending towards the minimum amount of effort that is necessary to achieve the maximum result, so that nothing is wasted. Therefore, we may wonder whether there exists

---

[19] As we have seen in the introduction (cf. section 1.3.2), *ya* is the only strategy that exhibits a structural constrain on the number of items, which should be at least two. Nevertheless, this is not an universal constrain of exemplification: all other strategies can be used also with only one example, and therefore does not invalidate our statement.

a number of examples beyond which the further mention of category members is considered as a "cognitive waste".

Therefore, even without precise constrains, there may be still some tendencies. If so, it is crucial to monitor them, since they may be motivated by cognitive mechanisms.

2.3.2.4 Exhaustivity

The last parameter of this section is exhaustivity. The aim of this parameter is to detect possible usages of Japanese exemplifying constructions in exhaustive contexts.

At this point, the reader may wonder as to why monitor such parameter, when all the constructions to be investigated explicitly codify non-exhaustivity (cf. section 1.3.1). Nevertheless, different studies (e.g., Alpatov 1997, Narrog 2012) have suggested that at least some of these strategies can be used in exhaustive contexts to perform different functions. For instance, *tari* is typically used to create non-exhaustive lists of events, as shown in example (2.41), but it can be also used to indicate the discontinuous repetition of an action or several similar actions during one period of time (cf. Alpatov 1997), as shown in example (2.42).

(2.41) *tari* to exemplify

    Nichiyōbi wa taitei   tomodachi to   [tenisu-o   shitari eiga-o   mi   ni   ittari]
    Sunday   TOP usually friend   with [tennis-ACC do:TARI film-ACC see   to   go:TARI
    shimasu.
    do.POL
    'On Sundays I usually do such things as play tennis with my friends or go to see movies.'
    (Chino 2001: 108-109)

(2.42) *tari* to indicate the repetition of an action

    Nisando   wakamono   no   mae-o   [ittari   kitari]   shita.
    Two.three.times   young.man   LK   front-ACC   go:TARI come:TARI   do:PAST
    'Two or three times he came and went in front of the young man.' (Alpatov 1997: 393)

The usage of *tari* to express the repetition of opposite actions is attested also in different grammars (cf. Chino 2001, Kaiser *et ali*. 2001).

Since we aim to investigate the functional space of exemplification, it is essential to monitor other emerging functions (even those in exhaustive contexts), in order to ultimately identify 1) which exemplifying constructions can perform these functions and which

constructions cannot, and 2) potential similarities or links between non-exhaustive functions and exhaustive functions.

To distinguish exhaustive instances from non-exhaustive ones, we will take into consideration other clues, such as 1) the semantic relationship between examples (e.g., antonyms); 2) the presence of different participants (in case of distributive meaning, cf. Alpatov 1997); 3) aspectual nuances (e.g., distributive, iterative, cf. section 2.2.4.1); 4) encyclopaedic knowledge (in some instances, our knowledge of the world can suggest that there are no other options beyond those explicitly mentioned).

Furthermore, this parameter has been also investigated and verified by means of the questionnaire (cf. section 2.1.2).

2.3.3 UTTERANCE

2.3.3.1 Aspect

As previously noted regarding exhaustivity (cf. section 2.3.2.4), some instances of *tari* used in exhaustive contexts have been linked to aspectual meanings, like iterative and distributive (cf. Alpatov 1997).

To investigate further this correlation, we will monitor aspectual values, distinguishing between: 1) perfective aspect (Comrie 1976: 16); and 2) imperfective aspect (Comrie 1976: 16). The latter is further subdivided into: 1) habitual aspect, which expresses "any situations that can be protracted sufficiently in time, or that can be iterated a sufficient number of times over a long enough period" (Comrie 1976: 30); 2) iterative aspect, which expresses the repetition of an event; 3) distributive aspect, which expresses situations where an event is applied to members of a group one after another.

Since, Japanese does not exhibit overt aspectual markers[20] that may help us to operate these distinctions (see Iwasaki 2013: 133), information about aspect will be mainly deduced from other elements. For instance, the speaker can convey such distinctions using frequency adverbs. Consider the example (2.42), repeated here as (2.43).

---

[20] Like English, Japanese does not have a special marker for the perfective aspect (cf. Iwasaki 2013: 133, Alpatov 1997). As for the imperfective aspect, Iwasaki (2013) notes that the structure *–te iru* is used to signal progressive and continuous aspects with continuative verbs. Depending on the transitivity of the verb, they are interpreted as either progressive (for transitive verbs) or continuous (for intransitive verbs). However, the same structure can also create a perfect-resultative meaning with instantaneous verbs (i.e., verbs that code an event that starts and ends simultaneously, e.g., *to die*). The structure *–te iru* cannot be used with stative verbs, which refer to a static condition that holds over time with no inherent initial and final point (cf. Iwasaki 2013).

(2.43) *Nisando*　　　　　　*wakamono*　*no*　　*mae-o*　　　*[ittari*　*kitari]*　　　*shita.*
　　　Two.three.times　　　young.man　LK　　front-ACC　　go:TARI come:TARI　do:PAST
　　　'Two or three times he came and went in front of the young man.' (Alpatov 1997: 393)

The adverb *nisando* "two or three times" highlights the repetition of the actions linked by *tari*, suggesting an iterative interpretation.

Similarly, the presence of different participants is an important clue for distributive aspect.

(2.44) *Soko-ni*　　　*wa*　　*mata*　*danjo-no*　　　　　　　　*kodomo-tachi-ga*
　　　there-LOC　　　TOP　again　male.and.female-DET　　　child-PL-NOM
　　　nanninmo　　　　　*oyoidari*　　　*moguttari*　　*shiteita.*
　　　as.many.as.come　　swim:TARI　　dive:TARI　　do:STA:PAST
　　　'Again as many boys and girls were swimming and diving there.' (Alpatov 1997: 393)

Here, the distributive interpretation is suggested by the presence of many participants and by the expression *nanninmo* here glossed as "as many as comes".

In other cases, aspectual distinctions can be suggested by the semantics of the entire utterance, as shown in the following example.

(2.45) *Sonnani*　　　*mado*　　　*o*　　　*aketari*　　*shimetari*　　*shinaide*　　*kureru?*
　　　like that　　　window　　acc　　open:TARI　close:TARI　do:neg:grd　give
　　　Could you stop opening and closing the window like that?

In the sentence above, there are no overt clues regarding aspectual properties. Nevertheless, the overall meaning of the utterance suggests an iterative (and exhaustive) interpretation.

2.3.3.2 Mode and modality

Modal verb forms and modality devices will be considered in our analysis to detect: 1) pragmatic functions, and 2) semantic relations between the examples (e.g., alternative relations).

We will focus on the *reality value* of a given utterance. As noted by Elliot (2000: 66-67), it is possible to distinguish between realis and irrealis propositions:

- A proposition is said to be REALIS when it asserts that a state of affairs is an "actualized and certain fact of reality" (Elliot 2000: 66).
- A proposition is said to be IRREALIS when "it implies that a state of affairs belongs to the realm of the imagined or hypothetical, and as such it constitutes a potential or possible event but it is not an observable fact of reality" (Elliot 2000: 67).

In our analysis, we will mainly focus on the irrealis value, which is linked to "the domains of imagination, possibility, wish, interrogation, necessity, obligation and so on" (Mauri 2008: 171). In other words, it codifies domains in which a given state of affair is presented as not having taken place or where the speaker is not sure about its occurrence.

For this very reason (cf. Fraser 1975, 2010), epistemic irrealis contexts may provide an ideal ground for pragmatic functions, since speakers are not sure about the occurrence of the reported state of affair, as shown in example (2.46).

(2.46) *I guess I can [chop them up or something].*

Moreover, this is also true for deontic irrealis contexts, where speakers are compelled to attenuate the assertiveness of the utterance to sound less direct[21].

Therefore, we monitor occurrences of overt irrealis markers (mainly verbal forms), that is, expressions encoding possibility, future, uncertainty, question, or similar domains which may require a mitigation of the speaker's commitment, but also irrealis devices such as lexical adverbs referring to irrealis contexts.

It is noteworthy that the correlation between irrealis contexts and pragmatic functions is not always straightforward. In other words, we cannot assume that exemplifying markers used in irrealis contexts always perform pragmatic functions. This is just a probable correlation that should be confirmed by means of other parameters (e.g., topic continuity, cf. section 2.3.3.3).

Exemplifying constructions can be used in irrealis contexts also to create categories of possibilities, that is, of alternatives. The correlation between the irrealis value and the notion of disjunction (cf. Mauri 2008) relates to the fact that alternatives are conceptualized as equivalent, mutually replaceable possibilities. It follows that until a choice is made or the

---

[21] This may be especially valid in Japanese, where being too direct is often considered impolite (cf. Iwasaki 2013, Taylor 2010).

speaker comes to know which hypothesis is realized, either alternative could be the non-occurring one and therefore both can be conceptualized as irrealis[22].

In this regard, the aim of this parameter is also to depict semantic relations that can be codified by non-exhaustive connectives (i.e., *ya, tari* and *toka*).

2.3.3.3 Topic continuity in discourse: categories and examples

Another parameter to identify different functions is topic identification. Under this regard, we can assume that, if the examples do not bear any independent reference but are mainly arrows to a context-relevant category, they cannot be the main topic of the discourse (whereas the category can). On the contrary, if the example bears an independent and discourse relevant reference, they can be the main topic of the discourse.

Consequently, at the discourse level, we rely on the notion of 'reference-tracking' (Foley and Van Valin 1984, Comrie 1989), which relates to the speakers' ability to track entities from one clause to the following clauses in an on-going discourse.

Therefore, we will monitor the organization of texts and referential paths in conversation (Robert 2008), revealing the topic continuity (Givón 1983) of categories and examples. Under this respect, we distinguish between cases in which the category is or becomes the topic of discourse and stays active through the subsequent text, and cases in which it is the specific example that is selected as topic. To achieve this, we will monitor the occurrence of words and phrases which semantically correspond to the main thematic field of the co-text in which the exemplifying constructions occur. For instance, consider the following sentence:

(2.47) *She got pregnant, so she had to start eating for the baby. She ate very healthy, except on weekends she would sometimes indulge a little on [cookies or pizza or something]. Now that she's had the baby, she is a lot bigger than she wants to be and hardly eats again.* (enTenTen13)

Here, the author refers to two examples: 1) cookies, 2) pizza. Examining the co-text, it appears clear that they do not bear any independent reference, but are used only to make reference to the category 'non-healthy foods that the subject still eats from time to time'. For instance, we can identify many occurrences of the lexeme 'eat' (i.e., *eating for the*

---

[22] Unsurprisingly, in some languages, irrealis markers are used to codify alternative relations. For instance, in Japanese, the marker used to codify exhaustive disjunction, namely *ka*, is also an interrogative marker (cf. Chino 2001: 45).

baby, ate very healthy, hardly eats again), which is the actual main thematic field of the text, since the author is comparing the eating habits of the subject before and after her pregnancy. Moreover, she makes explicit reference to the opposite notion to the one expressed by the category (i.e., *she ate very healthy* vs. *she would sometimes indulge a little on cookies or pizza or something*)*.* This further highlights the fact that the mentioned examples are only arrows to the category.

In this case, the author indents to focus solely on the category and she further manifests her intention by coming back to the domain of the at the end of the paragraph (e.g., *and hardly eats again*). Therefore, we classify this occurrence as an instance of exemplifying construction used to categorize.

In other cases, topic continuity suggests a different picture. Consider the following example:

(2.48) Elise: *Oh, the usual. Um, so, do you want to go to [dinner or something] sometime?"*
Tom: *Um... sure.*
Elise: *Great! Where do you want to go?*
Tom: *I don't know. Where do you like to eat?* (enTenTen13)

Here, Elise provides an example, namely, *dinner* (*or something*). Nevertheless, it appears clear that *dinner* is the actual topic of the sub-sequent text, where the two speakers discuss where to eat. Therefore, even if Elise uses an exemplifying strategy, her purpose is not to communicate a category, but to perform some other communicative functions. This type of occurrences will be investigated further in chapter 5.

To sum up, topic continuity will be investigated to identify or verify the function performed by the exemplifying construction in a specific occurrence.

2.3.3.4 Position of the exemplifying construction

We will briefly consider the position of the exemplifying construction in the utterance to identify different types of functions. This parameter is mainly language-specific, although other studies on exemplifying constructions have pointed out that different positions in the utterance may correlate with different discursive functions (cf. Ghezzi 2013).

Specifically, regarding Japanese, Taylor (2010, 2015) notes that whenever the scope of Japanese exemplifying constructions (i.e., *nado*, *tari* and *toka*) is not limited to noun phrases or verbal phrases, but it is extended to the entire utterance, their function is not to categorize, but to attenuate the assertiveness of the utterance.

This conclusion is backed up by the fact that many Japanese morphemes (e.g., *kedo* "but", *kara* "because") used at utterance-final position can produce pragmatic effects, mitigating the speaker's assertion, especially in spoken language (cf. Maynard 1989, Iwasaki 1993, Iguchi 1998). For instance, consider the following utterance:

(2.49) *Kaichuu-dentoo*   *toka ne?*   *ato*   *rajio* *toka* *mottari*   *toka.*
       flashlight   TOKA PP   and   radio TOKA hold:TARI   TOKA
       'Also taking something like flashing or radio toka' (Taylor 2010: 136).

Here, Taylor (2010) notes that while the first two *toka* (i.e., those attached to *kaichuu-dentoo* "flashing" and *rajio* "radio") are used to exemplify and therefore to communicate a category of similar objects, the third token of *toka* is used to soften the assertiveness.

Since we are adopting an approach that avoids imposing pre-existent theoretical definitions (cf. section 2.1), we will consider and monitor this parameter without assuming *a priori* such a strong connection between the position in the utterance and the function performed.

2.3.3.5 Types of speech acts

In our analysis, we also consider types of speech acts and performative occurrences.

Generally speaking, a speech act is an utterance that has a performative function (cf. Austin 1962) in communication. In *How to Do Things With Words*, Austin outlines the theory of speech-acts and the concept of performative language, according to which "to say something is to do something" (1962: 94). Speech acts can be analysed at different levels[23], but in our analysis, we mainly focus on illocutionary acts, "such as informing, ordering, warning, etc., i.e. utterances which have a certain (conventional) force" (1962: 108).

Searle and Vanderveken (1985) note that an illocutionary act is characterized by a propositional content and particular illocutionary force, that is, the speaker's intention in producing that utterance (e.g., asserting, promising, inquiring, ordering, etc.). For example, if we compare the two utterances "You will leave the room" and "Leave the room!", they have the same propositional content, but the former has the illocutionary force of a prediction and the latter has the illocutionary force of an order (1985: 109).

---

[23] Beyond illocutionary acts, other levels are 1) locutionary act, which is "roughly equivalent to uttering a certain sentence with a certain 'meaning' in the traditional sense", and, in certain cases, 2) a further perlocutionary act, that is, "what we bring about or achieve by saying something, such as convincing, persuading, deterring or surprising" (1962: 108).

The notion of illocutionary force has often been linked to that of hedging, in the sense that illocutionary acts often need to be attenuate in order to avoid face threatening situations (cf. Brown and Levinson 1987, Fraser 2010). This connection will be further discussed in chapter 5, but for now we should consider two factors: 1) some makers can assume pragmatic functions whenever are used in face threatening situations, such as illocutionary acts (cf. Fraser 1975, Brown and Levinson 1987); and 2) as suggested by Erman (2001: 1341), strategies implying a vague categorisation (such as exemplifying constructions) can have a face-saving function. Therefore, we will monitor the potential correlation between types of speech acts and different functions of exemplifying strategies. For instance, consider again the example (2.48), repeated here as (2.50).

(2.50) Elise: *Oh, the usual. Um, so, do you want to go to [dinner or something] sometime?"*
      Tom: *Um... sure.*
     Elise: *Great! Where do you want to go?*
     Tom: *I don't know. Where do you like to eat?* (enTenTen13)

According to the parameter of topic continuity, we have already established that the topic of discourse is not a context-relevant category, but the mentioned example. At this point, we should also note that Elise is performing an illocutionary act (i.e., a request) and therefore she may feel compelled to attenuate the illocutionary force of her utterance. Thus, we classify this occurrence as an instance of exemplifying construction used as hedging strategy (cf. chapter 5).

It is noteworthy that, as for irrealis context (cf. section 2.3.3.2), the correlation between speech-acts and pragmatic functions is not always straightforward. We need to consider also the case of exemplifying constructions used to create and communicate categories in illocutionary acts. For instance, consider the sentence below:

(2.51) *The fridge is empty. Buy a pizza or something.*

Here, despite the illocutionary act (i.e., an order), the exemplifying construction is used to refer to a context-relevant category 'take away food', as confirmed by the topic continuity (i.e., *The fridge is empty*), and not to hedge the assertiveness of the utterance.

Therefore, while this parameter is useful in identifying different communicative functions, it is should not be considered a defining feature.

2.3.4 CONTEXT

The final parameter to be considered in our analysis is the context.

Since one of our goals is to examine the dynamic construal of context-relevant categories, the role of the context inevitably holds an important place in our analysis. For instance, we have already noted that the access to the situational context is essential to properly construe the inferential process (cf. sections 1.2.3 and 1.3.1).

In this regard, we need to investigate how the context can influence and direct these cognitive processes. Some studies in this regard have already been conducted, both at the pure cognitive level (cf. Barsalou 1983, 2010), but also at the linguistic level (cf. Croft and Cruse 2004, Wilson & Carston 2007, Carston 2010). Nevertheless, we believe that a more focused investigation on the role of context in exemplification may add further insights regarding the relation between cognition and language.

To achieve this, we need a working definition of *context*. Indeed, context should not be considered as an inert setting for cognitive and linguistic processes, but as a multi-dimensional element consisting of different components, such as shared knowledge and interpersonal relations. More precisely, following the taxonomy proposed by Croft and Cruse (2004: 102-103), we can distinguish four types of context: 1) *linguistic context*, which includes preceding discourse, immediately adjacent co-text, and the type of speech; 2) *physical context*, which includes elements selected on the basis of perception; 3) *social context*, including the relationships between the interlocutors; and 4) *stored knowledge*, which regards that "background of a vast store of remembered experiences and knowledge, which is capable of affecting the likelihood of particular construals" (2004: 103), and therefore includes information related to the speaker, the listener, to their background and habits. All these dimensions of the notion *context* can effectively influence the elaboration of the examples and thus the inferential processes.

While monitoring as much as possible all these dimensions, we will devote particular attention to the linguistic context and the role it may play in actively directing cognitive processes.

3. EXEMPLIFICATION OF LEXICALIZED CATEGORIES

3.1 THE NOTION OF LEXICALIZED CATEGORY

Traditional theories of categorization conceptualized a constant mental representation underlying each category in order to explain the stability of cognition, that is the ability of performing "the same cognitive act over and over despite varying local circumstances" (Smith and Samuelson 1997). This notion of stability seems to have an impact also on the modality in which these categories are linguistically codified: stable categories can typically be encoded by short conventional lexical items (e.g., "birds", "fruit"), which are known as the labels or names of the categories. To describe this linguistic feature of common stable categories, Overstreet (1999) introduced the notion of lexicalized category, in opposition to non-lexicalized category which does not come with ready-made linguistic labels (e.g., ad hoc and goal-derived categories). The assumption behind this distinction is that there is no need to use other linguistic strategies, such as exemplification strategies, to make reference to a stable lexicalized category, because the speaker can easily and more precisely refer to it by its name.

Therefore, it appears to be a binary model: on the one hand, there are stable lexicalized categories, and on the other, (other types of) categories built on the fly, which need different linguistic strategies to be encoded.

However, this model exhibits some shortcomings addressing the existence of categories lacking a specific label but having a conceptual reality. This case encompasses some grey areas between the notion of "common categories" and other types of unstable non-lexicalized categories.

In chapter 1, we introduced the concept of "covert categories" (Cruse 1986: 151). While analysing the vertical dimension of categories, Cruse noted that some slots come without any names. For instance, English has the superordinate category name "furniture" and the basic-level category name "chair", however it has no single-word name for the category in between, that is "pieces of furniture that you can sit in". The latter is intuitively recognized as a conceptual category and people can easily mention some examples of it. Still, it has no label. The existence of these "lexical gaps" (Cruse 1986) is problematic, because there is no real difference between the nature of the category "furniture" and the nature of the category "pieces of furniture that you can sit in". The only substantial difference seems to

be the way they are linguistically encoded.

Channel (1994: 123) tries to address this issue by recognizing the existence of named common categories and common categories which do not have a name as two separate phenomena. However, this tripartite model actually puts into question the only linguistic parameter which can help us distinguish between different types of categories, especially when we take a broader interlinguistic perspective. While addressing specifically the question of how novel categories are named, Glucksberg (2001) notes that, unlike English, American Sign Language (ASL) has no single-word name for the superordinate category "furniture". This is not an isolate case: many natural languages do not have names for superordinate categories and have to employ other linguistic strategies to refer to such categories. For example, the Kiowa (Kiowa-Tanoan) speakers in western Oklahoma use cottonwood for trees in general, and real-cottonwood for the specific tree (Trager 1938). This is just a small measure of how arbitrary, cultural-dependent and context-dependent the process of lexicalising a category may be.

While most of these studies do not question the notion of stable common categories, they still recognize a difference between lexicalized or non-lexicalized common categories and non-lexicalized functional categories built on the spot. However, as already noted in chapter 1, the issue becomes even more problematic when we consider that there is a strong bias behind the idea of cognitive stability of common categories. In fact, most of the investigations on categories and category labels have been undertaken through experimental tests with constructed examples in restricted context. Recent empirical evidence (e.g., Barsalou 1983, Rips 1989, Smith and Sloman 1994) suggests a different scenario in which the context seems to play an important part and in which "acts of categorization are not simply repeated; they vary. Different tasks and contexts seem to create different categories." (Smith and Samuelson 1997: 167). These findings ultimately suggest that the notion of stable categories with constant representations is unrealistic, and that categories are inherently variable, created on the spot whenever they are needed. This idea of inherently variable and context-dependent categories exhibits consequences also on the usage of category names or labels.

While, experiments specifically dedicated to the usage of category labels and the theorizing on the matter are unsatisfying scant, the concept of ad hoc categories has recently been employed in research on lexical pragmatics, within the Relevance Theory approach (Wilson and Carston 2007, Carston 2010). These studies propose a radical contextual approach in order to account for why "the meanings of words are frequently

pragmatically adjusted and fine-tuned in context, so that their contribution to the proposition expressed is different from their lexically encoded sense" (see Wilson and Carston 2007). According to this inferential approach, the meanings of individual words as well as a number of phenomena such as hyperbole or metaphor are analysable as varieties of a single pragmatic process of lexical adjustment which involves either lexical narrowing or broadening. In addition, "narrowing and broadening are flexible, highly context-dependent processes" (Wilson and Carston 2007: 234). Therefore, these studies ultimately show that the interpretation of virtually every word depends on the context in which it occurs.

Actually, Wilson and Carston recognize the effects of this approach on categorization processes. Based on the insights provided by Glucksberg (2001), they theorise (Wilson 2004, Wilson and Carston, 2007) a phenomenon called "category extension" which can be conceptualized as a variety of broadening. This phenomenon is typified by the use of salient brand names (e.g., Kleenex) to denote a broader category ('disposable tissue') which includes also items from less salient brands. However, also common names can undergo the same process. For instance, in "brown is the new black", black is not just a colour, but it evokes the category of staple colours in a fashion wardrobe (Wilson 2004: 345).

The identification of ad hoc concepts and categories, and thus the key-role of context, has changed many assumptions about the way categories are built and communicated. On the one hand, there are several cognitive psychology studies that show how categories are indeed inherently dynamic, context-dependent and computable within at given situation. On the other hand, research on cognitive linguistics and psycholinguistics shows that also the meanings of individual words or phrases strongly depend on the context.

In such an account, it seems anachronistic to still rely on the notion of stable associations between categories and linguistic expression. On the contrary, it seems more sensible to assume that every category conveyed by lexicon is necessarily translated into a more concrete category anchored in the specific situational context (Mauri 2014, 2016).

At this point, we can sum up what has been said above in three main points:

a) Not all categories recognized as "common categories" are lexicalized and often their lexicalization depends solely on the specific language (Cruse 1986, Channel 1994, Glucksberg 2001);

b) Even categories which have been recognized as "ad hoc categories" may be lexicalized (in the sense of being encoded by short conventional lexical items), as long as the interpretation of context and the shared knowledge work to avoid any

misunderstandings (Wilson and Carston 2007, Mauri 2014, 2016);

c) In real-life situations, all categories are created and used to make reference to a very highly variable world, thus they are always context dependent, they are all ad hoc categories (Smith and Samuelson 1997).

What we have said above provides us with some theoretical background on the study of the linguistic codification of categories. However, since our analysis focuses on the role of exemplifying strategies, we can say that we are facing the same issue from the opposite side. In other words, our main concern is not to understand what we can do with category labels, but what types of categories have to be encoded by means of exemplifying constructions.

Traditionally, exemplifying constructions have been considered to create mainly ad hoc non-lexicalized categories (Mihatsch 2010a: 52). However, Overstreet (1999: 44) attested also few instances of examples and general extenders used to make reference to what she recognized as common lexicalizable categories, instead of using the available labels. But why would a speaker provide exemplars to make reference to a lexicalized category, when she could refer to the category simply by its name?

Consider (3.1):

(3.1)   *Most of 'em are evergreens around there I guess. Pine trees an' stuff.* (Overstreet 1999: 44)

In addition to naming the lexicalized category, the speaker decides also to provide a concrete example (pine trees) in order to illustrate what kind of items are designated by this label.

Overstreet (1999: 44) ascribes this phenomenon mainly to pragmatics, arguing that there are at least four pragmatic reasons for using exemplifying strategies to make reference to lexicalized categories: 1) the speaker does not remember the label; 2) the speaker thinks that the hearer does not know the label; 3) the speaker wants to emphasize the number of members in the category; 4) the speaker wants to emphasize or highlight certain members of the category.

According to Overstreet, in (3.1) probably the speaker suspects that the hearer who grew up in Hawaii, may be unfamiliar with the category evergreens. This is a likely interpretation, yet it may not be the only one. For instance, can we be sure that the category in the speaker's mind is the common (and I will also add "taxonomical") category "evergreen"? We

may also consider the possibility that the actual category is a more specific sub-set, such as "type of evergreen that grows around there". It may seem a slight difference, but it can still influence the way we interpret the expression "pine trees an' stuff". The category "evergreen" is not as broad as the category "tree", still it encompasses a wide variety of trees, from pines to olives. It seems unlikely that in a specific context, the speaker really wants to address the (taxonomical) category in its entirety. Following this approach, the label "evergreen" sounds too vague and may potentially be misunderstood. Thus, by adding "pine trees an' stuff", the speaker is trying to specify the reference to the category: he is talking about pines and other similar evergreen trees.

This hypothesis still explains why the speaker feels the need to add a concrete example and why he may think that the hearer is not familiar with the category as well.

Consider (3.2):

(3.2)   *I'm going to get some milk, and stuff.* (Overstreet 1999: 45)

According to Overstreet, the entire construction exemplar + general extender could be summed up by the label "groceries", but by choosing to say "milk and stuff", the speaker highlights one member of the category while also referring to the category.

In an unspecific situation, Overstreet's propose may be actually correct. However, depending on the context of the utterance, the underlying category may change as well. For instance, a father of a newborn baby may say something similar before leaving to the supermarket. In such a context, the label "groceries" would be too generic.

The reason behind these different interpretations is that there is no stable association between expressions such as "milk, and stuff" and specific categories: each utterance should be analysed and interpreted as part of a larger context.

Taking all these issues into account, we come across to a very different and more complex picture than the binary model postulated at the beginning of the section. And at this point it seems natural even to question the notion of lexicalized category as it has been conceptualized by Overstreet, because, at first glance, the difference between lexicalized and non-lexicalized categories ceases to exist: no category can be represented and described by a short and stable linguistic mean in a better way than other categories.

However, we propose that the notion of lexicalized category (and consequently non-lexicalized category as well) is still a valuable linguistic parameter, once it is conceptualized according to this new dynamic approach on categorization. In order to do this, we should

consider only the structure of the linguistic expression that codifies the category.

There is a significant difference between example in (3.1) and example (3.2). In (3.1) the creation of the category consists of two separate elements: a category label and a concrete member taken as an example. On the contrary, in (3.2), there is no label, but just a concrete member taken as an example of the category. In other words, we can distinguish between two main patterns: in the first one exemplified by (3.2), the speaker only uses the example(s) to build and communicate the category while in the second exemplified by (3.1) the speaker communicates the category through a label and one or more examples.

These two patterns are equally well attested in my corpus. Compare the examples below:

(3.3) *Koe-ya* *ugoki* *wa,* *dokusha-no* *sōzō-no* *hanchū* *da.*
Voice-YA movement TOP reader-GEN imagination-GEN category COP
‘Voices and movements are categories of the readers' imagination.’

(3.4) *Yōryō* *wa* *50MB-ni* *seigensareteiru* *ga,* *dokyumento-ya*
Capacity TOP 50MB-DAT limit:PASS:STA but document-YA
seishiga *toitta* *fairu* *deareba* *jūbun* *darou.*
still image such as file COP.COND enough MOD
‘The capacity per file is limited to 50MB, but if they are files such as documents or still-images, it should be enough.’

In (3.3), the speaker is complaining that the transposition of novels into comics takes away the pleasure of imagining the characters' faces from the reader. Only few features remain prerogatives of the reader's imagination, such as their voices and the way they move. Thus, in order to create and communicate the category (i.e., features that a reader can still imagine while reading comics), she lists some actual members as representative examples of the category, signalling their status of example by means of the dedicated non-exhaustive connective *ya*. Therefore, according to a context-based similarity reasoning, the hearer is able to infer other potential instances, leading to the construction of the category.

On the contrary, in (3.4) the speaker wants to make reference to the category "files that do not take up a lot of hard drive space". Thus, she uses a more complex construction consisting of a general label (i.e., *files*) linked to a short list of examples (i.e., *documents and still images*) by means of an approximator (i.e., *toitta*). Again, the dedicated non-exhaustive connective *ya* signals the existence of other potential members of the category and thus, that still-images and documents are only representative of a larger set.

Comparing these two sentences, a question arises: is the possibility of lexicalizing a category an inherent feature of the category itself (cf. Rosch 1975, Overstreet 1999), or just a communicative strategy adopted by the speaker? To provide a reasonable answer, we may try to rewrite both sentences according to the opposite pattern.

While just the category label would have been too generic to make a correct and specific reference, in (3.5) we note that it is possible to create and communicate the same category mentioned in (3.4) just by means of a list of concrete examples[24], that is, using the same pattern shown in (3.3).

(3.5) *Yōryō   wa   50MB-ni   seigensareteiru   ga,   dokyumento-ya   seishiga*
       Capacity TOP 50MB-dat   limit:PASS:STAT   but   document-YA   still image
       (nado)   deareba   jūbun   darou.
       etcetera   COP.COND   enough   MOD
       'The capacity per file is limited to 50MB, but if they are documents, still-images and so on, it should be enough.'

In the same way, in (3.6) we note that also the sentence in (3.3) can be rewrite using the same pattern shown in (3.4).

(3.6) *Koe-ya   ugoki   toitta   tokuchō wa,   dokusha-no   sōzō-no*
       Voice-YA   movement   such as   feature TOP   reader-GEN   imagination-GEN
       hanchū   da.
       category   COP
       'Features such as voices and movements are categories of the readers' imagination.'

To rewrite the sentence, we simply choose a category label (i.e., *tokuchō*) that could well encompass the concrete mentioned members, and then add a linguistic connector (cf. section 2.3.1.5) to link the label to the examples. In both cases, there is no substantial difference in the final interpretation of the category.

Therefore, through this little experiment, we may assume that:

a) Both the underlying categories in examples (3.3) and (3.4) are indeed ad hoc categories;

---

[24] We decided to rewrite the sentence using the synthetic general extender *nado* (i.e., "etcetera", "and so on") at the end of the list just because it sounds more natural in Japanese. However, its presence is optional.

b) In (3.4) the speaker chooses to lexicalize (that is, to provide a category label for) the category, while in (3.3) the speaker chooses to use only a list of examples;

c) To some extent, it seems possible to lexicalize any category (even ad hoc categories);

d) Lexicalizing a category (that is, to provide a category label) seems to be just an arbitrary choice;

e) The latter point adds a further question to the analysis: when do speakers choose to lexicalized the category?

We propose that 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 communicate categories in specific contexts. Therefore, for our analysis on exemplification, we define 1) as lexicalized categories all those categories that are expressed by means of a label and a list of examples[25]; 2) as non-lexicalized categories all those categories that are encoded only by examples. Furthermore, we propose that this distinction is important not only at the formal level, but also at the functional level, that is, that the functional role of exemplification changes depending on the presence or the absence of the label.

3.2 LINGUISTIC CODING OF THE CATEGORY

Before dealing with the analysis of the different constructions attested in the corpus, it is reasonable to explain the choice to devote attention to category labels. Since labels are often perceived as strategies alternative to examples, our choice may seem a contradiction in a study that aims to analyse exemplification. Nevertheless, our data suggest a very different interpretation of the phenomenon at the linguistic level, especially when we observe the following figures which show the frequency of labels and the distribution of their position in the utterance. The result is a picture that justifies a thorough investigation of the role of the label.

In Table 3.1, we examine the overall frequency of distribution of the lexicalizing strategy. The total number (664 occurrences) has been calculated excluding instances of

---

[25] Lexicalized categories are categories designated by means of a category label. This means that examples are not necessary. Nevertheless, since our analysis concerns exemplification strategies, we will not consider instances of category designated solely by category labels.

exemplifying constructions used for other purposes than categorizing function (e.g., pragmatic functions, cf. chapter 5; exhaustive occurrences, cf. chapter 6).

Table 3.1: Distribution of lexicalized categories.

	Total Categorization Function	Lexicalized categories
ya	170	62 (37%)
nado	194	101 (52%)
tari	148	36 (24%)
toka	152	67 (44%)
Total	664	266 (40%)

Based on the presence of an explicit label, it is possible to identify 264 occurrences of lexicalized categories in my corpus data, which means that the usage of a category label is a well-attested strategy (40 percent of the total number of occurrences). This first result shows how labels and examples are not merely competing strategies to make reference to categories, but instead they frequently occur together, in the same utterance. Nevertheless, this is just a partial information, because the fact that they can actually occur together does not prove that there is any substantial link between them, nor elaborate what type of relationship can be established between labels and examples. Therefore, to investigate these issues, it is necessary to examine another important parameter, that is, the position of the label with regards to the example(s), as shown in Table 3.1.
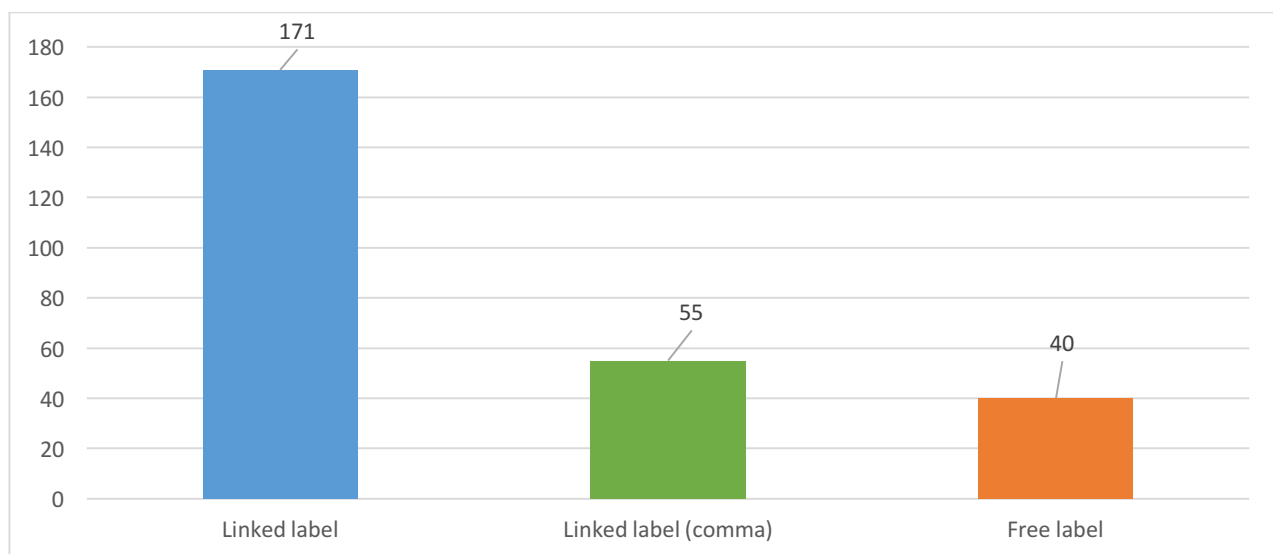
Figure 3.1: Positions of the category label in relation to examples

Interestingly, from the figure above, it emerges that, not only labels occur frequently with concrete examples, but also that, in most cases, they are directly linked to the example(s), forming a single combined linguistic construction: 226 of 266, or 85% of the total number of occurrences. Out of the 226 occurrences, in 171 occurrences (64% of the total number), the label is directly linked to the example(s) by means of a linguistic connector, that is, a linguistic construction that explicitly codify the relationship "X is an example of Y" (e.g., *toitta* "such as", cf. section 2.3.1.5). In 55 occurrences (21%), the label is still linked to the example(s), but without any connector. In such cases, a comma is used to signal where the list of examples ends, just before the label, and thus the different status of label and example(s). Finally, in 40 occurrences (15%) the label is completely detached from the exemplifying construction. These patterns are exemplified by the following instances:

Linked label: [example(s) + connector + label]

(3.7) *Orenji-ya*   *ierō*   *nado*   *no*   *bibiddo*   *karā*
Orange- YA   yellow **NADO** NML   bright   colour
"Bright colours such as orange, yellow and so on"

Linked label (comma): [example(s) , label]

(3.8) *Blockbuster-ya*   *eBay*   *nado, iroirona*   *kyōryoku*   *saito*
Blockbuster-YA   eBay   **NADO,** various   collaboration   website
"Various partner websites, such as Blockbuster and eBay"

Free label: [label] [example(s)]

(3.9) *Sakkon-no*   *rajio-kyoku-de*   *wa, chikahōsō-no*   *kazu*   *wa*
recently-GEN   radio-station-LOC TOP underground.broadcast-GEN number TOP
herimashita.
decrease:POL:PAST
'Kōgi   *no*   *chikahōsō'*   *to*   *mo*   *ieru*   'Radio Free Europe /*
broad   NML   underground broadcast   QT   also   say:POT   Radio Free Europe /
Radio Liberty',kyūba-muke-no   *'Radio Marti', Ajia-muke-no 'Radio Free Asia'*
Radio Liberty Cuba-for-NML   Radio Marti   Asia-for-NML   Radio Free Asia
nado   *wa*   *mada*   *nokotteimasu.*
NADO   TOP   still   remain:STA:POL
'As for radio stations in recent years, the number of underground broadcasts has decreased. "Radio free Europe/Radio Liberty" which can also be called "Broad underground

broadcasting", "Radio Marti" intended for Cuba, and "Radio Asia" intended for Asia and so on still remain.

The first point of interest that emerges from the analysis of these figures is the frequency of the label directly linked to examples. Out of the 664 occurrences of exemplifying strategies used to make reference to categories, in 226 occurrences (34%) the label is directly linked to examples. That means that even without considering the distinction between lexicalized and non-lexicalized categories, but looking at the whole picture, the linked label pattern is still quite frequent. This figure suggests not only that the label can work as an important cognitive tool in categorization processes, but also, and more importantly, that this is true even when the speaker actually uses exemplification to make reference to the category. This means that not only labels and examples can coexist, but that they can actually work together in a single unified linguistic construction. Therefore, it is pivotal to understand the role and the contribution of labels in the inferential processes, especially because labels make the defining property of the category clear.

As for the linked label pattern, at first glance there seems to be no particular differences between the pattern that involves the use of a language connector and the one that uses a comma between label and examples. Instead, both strategies seem to indicate the necessity of making clear the different status of examples and label. So, in the absence of any linguistic connector that encodes the relation "X is an example of Y", it is necessary to at least use a comma, thus indicating the change of status through a pause. It is possible that the occurrence of one pattern instead of the other is linked to readability issues depending on the syntactic complexity of the label or of the examples. In other words, if the label or the examples (but particularly the label) are syntactically complex, the creation of a single connected construction may be considered problematic. In these cases, the use of a comma is the best strategy, especially in a written text. This issue will be further investigated in the following sections 3.2.1 and 3.2.2, where we will consider the different syntactic types of labels.

Another interesting fact that emerges from Figure 3.1 concerns the free label pattern, that is, whenever the label is not directly linked to the examples. Unlike the linked label patterns (both the one with a linguistic connector and the one with just the comma), in the free label pattern there are no structural constraints regarding the order of constituents, that is, the label does not have to occur after the examples. Despite this fact, the tendency is to express the label first, often in the previous sentence or in the previous turn. Out of 41 occurrences of the free label pattern, there are no cases where the label is placed after the examples

without being directly connected to them. In other words, in my corpus data, free labels are always placed first. While trying to understand this tendency, it is worth noting that this is the traditional pattern of usage ascribed to illustration by means of examples (e.g., the pattern of constructions such as *for example* and *for instance*), that is, where a certain issue is presented, and then it is specified or clarified by means of some concrete examples. Let us consider again (3.9). The article is about underground radio broadcasts which are struggling to broadcast to the point that some of them shut down. While the first sentence presents the issue (i.e., the decreasing of underground broadcasts, *chikahōsō* "underground broadcasts" is the label), in the second sentence, some concrete examples are provided to clarify the previous point. Therefore, in this type of pattern, we may even say that examples are ancillary to the label, which represents the issue into question, therefore the position of the label prior to the examples does make sense from the perspective of the underlying cognitive process (see also Barotto and Mauri 2016).

The opposite pattern (that is, examples before the unconnected label) is possible, albeit not represented in my corpus data. Yet it seems to denote a different type of cognitive processes, which may indicate even more pragmatic motivations. Like Perelman and Olbrechts-Tyteca noted in rhetoric, exemplification can be used to establish a rule (1969: 359). In this sense, the rule is yet unknown and it has to be recognized by means of presenting particular instances. Linguistically, we can borrow this cognitive pattern to understand a linguistic uncanonical order. For example, a speaker can refer to a category whose label she doesn't know (especially in cases of complex ad hoc categories), or can't recall, and through the mentioning of some concrete instances, she can identify a suitable label to describe the category. Or, a speaker may mention some concrete instances and then realize that the category might be summarized by means of a label. Let us consider the following example in Italian:

(3.10) *Basta andare nelle stazioni,   sotto ai   porticati, nelle mense   degli*
       Just   go   in the station:PL   under the   porches in the soup.kitchen:PL of the
       organi caritatevoli,   insomma   nei   posti   dove   i poveri   disperati
       charity:PL   so   in the place:PL   where poor people   desperate
       si radunano.
       gather
       'Just go in the stations, under the porches, in the soup kitchens of charities, so in the places where desperate poor people gather.'

In (3.10), the speaker provides a list of the places where poor people can be found easily, then she realizes that it is possible to describe and summarize all those places just by means of a complex label "places where desperate poor people gather". In Italian, *insomma* is not a linguistic connector and can be paraphrased as "to conclude, to summarize", thus it makes clear not only that the following label encompasses all the previous concrete instances, but also that through the mention of concrete instances, the speaker has come to a conclusion represented by the summarizing label.

We cannot exclude that there might be other reasons behind the choice of using a separate label after the examples. Unfortunately, being this pattern not attested in our corpus, we can just propose some hypotheses without analysing the actual phenomenon.

So far, our considerations implied a uniform notion of "label", without any internal substantial differentiation. However, lexicalized categories are far from being a homogeneous group and it is indeed possible to recognize different types of labels. In particular, different constructions are attested according to one main parameter, that is, the syntactic type of the category label. The presence of a label leads to two types of constructions. In the first case, the label is a single general noun:

(3.11) *Ijō-ga*   *kenchisareta*   *baai wa,*   *mēru*   *niyoru*   *tsūchi-ya*
        strangeness-NOM   detect:PASS:PAST case TOP   mail   by   notice-YA

        sesshon-o   *setsudansuru*   *toitta*   *taiō*   *mo*   *kanō*   *to*   *iu.*
        session-ACC   disconetting:do   such as support also   possible QT   say

        'In case strangeness is detected, (it is said that) software support such as cutting the session or notification by mail is possible as well.'

In the sentence above, the author refers to security management software that can detect and block computer viruses and spyware, filter spam messages and so on. The word *taiō* means "software support" and, from the point of view of the categorization process, it can be interpreted as a label which encompasses all those services offered by a software in moments of difficulty regarding a computer. However, while the category identified just by means of the label *taiō* is general and wide, the actual range of items that might be included in the category is constrained by contextual factors. More precisely, (3.11) describes the specific situation where the software detects and blocks immediately unauthorized accesses. It follows that without any other reference to the category apart from the label, the reader should have deduced the actual members of the category by interpreting co-textual and contextual factors. This process not only requires a cognitive effort, but it also depends on

the ability of the reader to correctly interpret the context and on her previous knowledge regarding the matter, that is, what types of software supports usually take place whenever the device faces difficulties.

In the second type of constructions, the label is communicated by means of complex expressions. Usually, the label consists of a nonspecific[26] super ordinate noun (i.e., the general label of the category) and some sort of linguistic adjuncts (e.g., adjectives, relative clauses, genitive phrases) which provide a higher degree of contextualization and a more precise reference to an ad hoc category. For example:

(3.12) *Doraggu&doroppusuru*   *dake*   *de*   *shashin-o*   *appurōdoshitari,*
drag&drop:do   only   STR   photo-ACC   upload:do:TARI
daburukurikku-de   *suraidoshō-o*   *saiseisuru*   *nado, shoshinsha*   *demo*
double click-STR   slideshow-ACC   play:do   NADO   beginner   GDR:also
kaitekini   *riyōdekiru*   *kantanna*   *sōsa.*
simply   use:POT   simple:AGG   operation
'Simple operations that can be used comfortably even by beginners', such as playing a slideshow with a double-click, uploading photos by simply dragging and dropping and so on.'

Here the author refers to a software which allows users to watch and share photos and videos. In this particular instance, the label is a complex expression and it consists of a simple noun (i.e., *sōsa*, "operation"), an adjective (i.e., *kantanna,* "simple") and a relative clause (i.e., *shoshinsha demo kaitekini riyōdekiru,* "that can be used comfortably even by beginners"). It is important to note that, unlike in (3.11), in this case, the speaker makes a first attempt to contextualize more precisely the category through the label itself: she clarifies that we should not consider all the possible operations of the software, but just those that even beginners can use easily. In other words, just by means of the label, the speaker tries to suggest what kind of items should be included inside this specific category.

We should note that from the perspective of categorization, there is no substantial difference between the ad hoc category in (3.11), namely 'software supports to block immediately unauthorized accesses', and the one in (3.12), that is, 'simple operations that can be used comfortably even by beginners'. They are both context-relevant categories built in the discourse to achieve some communicative goals. What changes is the lexicalization process, in the sense that speakers may choose different types of labels to make reference

---

[26] Regarding the distinction between specific/unspecific labels, see section 2.3.1.3.

to the same type of category. This reveals a different picture than the one postulated regarding common categories (cf. Rosch 1975, Overstreet 1999), which can be expressed only by means of words, or at least simple short linguistic expressions.

We argue that this is motivated by the notion of "ad hoc category label" itself. The concept of category label has never been directly linked to that of ad hoc categories, to the point of defining the latter as non-lexicalized categories (cf. Channel 1994, Overstreet 1999), that is, categories without a proper name. This should be ascribed mainly to the recognized lack of a stable association between this type of categories and fixed linguistic expressions: being categories built on the spot whenever they are needed (cf. Barsalou 1983), they are too inherently volatile to create stable representations of any kind in memory, and thus to allow to be tightly linked to words or other linguistic expressions as with taxonomical common categories. Therefore, to designate this type of categories, Barsalou (1983: 211) proposes a specific linguistic construction which functions more like a paraphrase, rather than a real label. He suggests a structure consisting of nonspecific superordinate nouns (things) and infinitival purpose clauses (to do X), de facto emphasizing the expression of goal orientation of these categories following the way they have been conceptualized by Barsalou himself. Hence, examples of ad hoc categories can be expressed in phrases such as "things to take on a camping trip" or "places to look for antique desks". Nevertheless, when it comes to the spontaneous creation and communication of ad hoc categories in discourse, it seems that the speakers can rely on other linguistic strategies besides the one suggested by Barsalou, that range from simple nouns (to be interpreted according to the context) like in (3.11), to the creation of "ad hoc" labels by adding further linguistic material like in (3.12). Even if, most of the times, these labels are not detailed as the paraphrases proposed by Barsalou, they are still more functional in a discursive situation, where the speaker faces a constant struggle between the principle of economy and the efficiency of the reference.

Having recognized that the lack of stable representations at the cognitive level seems to be linked, at the linguist level, to the lack of a designated linguistic construction to communicate context-relevant categories, in the following section, we will examine the different types of ad hoc category labels attested in our corpus.

3.2.1 SIMPLE LABELS VS. COMPLEX LABELS

In the introductory section, we deliberately chose two examples at the edges of the continuum to highlight the heterogeneity among occurrences; however, it is important to note that – syntactically speaking – simple labels and complex labels do not constitute a

strict and discrete distinction, but two edges of a linguistic continuum. In our corpus, different degrees of complexity are attested, from simple noun, to noun phrases containing relative clauses, to noun phrases containing a mix of different other syntactic adjuncts like in (3.12).

   More specifically, through the analysis of the syntactic structure of the label (cf. section 2.3.1.2), we identified six structural patterns: 1) simple nouns (3.13), 2) noun phrases containing adjectives (3.14), 3) compounds (3.15), 4) noun phrases containing genitive clauses (3.16), 5) noun phrases containing relative clauses (3.17), 6) noun phrases consisting of two or more adjuncts (3.12).

(3.13) *Sofutobanku mobairu wa, meiwaku mēru taisaku-no kyōka-ya kakin*
Softbank mobile TOP trouble email against-GEN reinforce-**YA** billing
taikei-no henkō toitta shisaku-o happyōshita.
system-GEN change such as measure-ACC announce:do:past
"Soft bank mobile announced measures such as a reinforcement of the anti-spam policy and a change in the billing system."
Label: *shisaku* "measures"

(3.14) *Makaron-ya, gimōbu nado no ryūkō-no suītsu-o kikakushitekita*
macaroon-YA marshmallow NADO NML fashion-ADJ sweet-ACC plan:do:STA:PAST
no desu.
NML COP:POL
'(We) have planned (to prepare) fashionable sweets such as marshmallow and macaroons.'
Label: *ryūkōno suītsu* "fashionable sweets"

(3.15) *Sōsharunettowākingusaito wa fisshingu-ya onrain sagi toitta sagi*
social.network.website TOP phishing-YA online fraud such as fraud
kōgeki-no kakkōno hyōteki tonatteimasu.
attack-GEN easy target become:STA:POL
'Social networks have become easy targets of fraud attacks such as phishing and online fraud.'
Label: *sagi kōgeki* "fraud attacks"

(3.16) *Kiniro-ya giniro nado toitta torofī-no iro*
Gold-YA silver NADO such as trophy-GEN colour
'Colours of the trophy such as gold, silver and so on.'
Label: *torofī no iro* "colours of the trophy"

(3.17) *Oiwai*   *messēji-ya*   *chikoku-no*   *renraku*   *nado* *fudan-no*   *seikatsu-ni*

celebration   message-YA   lateness-GEN message   NADO everyday-ADJ life-DAT

awaseta   *sozai-o*   *teikyōsuru.*

match   material-ACC offer:do

'(It) provides materials tailored to everyday life such as messages for being late and congratulation messages.'

Label: *fudanno seikatsu ni awaseta sozai* "materials that are tailored to everyday life"

Frequencies of different syntactic types of category labels are given in Table 3.2.

Table 3.2: Distribution of syntactic types of category labels.

	N	Adj N	Compound	Gen N	Rel N	Mix
ya	27	7	15	4	7	2
nado	44	11	27	6	11	2
tari	20	3	3	2	7	1
toka	23	16	4	4	19	1
Total	113	37	49	16	45	6
in %	(42%)	(14%)	(18%)	(6%)	(18%)	(2%)

At a first glance, it appears that the notion of "ad hoc category label" encompasses different syntactic constructions depicting a more composite picture than the one related to common categories. This means that the speaker provides the quantity of contextualization that it is believed as necessary in each context by adding different types of lexical adjuncts. Moreover, without a stable association between categories and linguistic expression, it is also plausible to assume that different speakers in different discursive context could create different labels for the same category simply by using different linguistic material.

However, even though all structural patterns are attested, the data in Table 3.2 reveal a clear tendency to use more frequently simple labels expressed by simple nouns (43%) than any other syntactic structure. Therefore, being aware that this is not a discrete distinction but a continuum, we may still identify two main patterns of usage: 1) the speakers may use general nouns to be contextualized by drawing on the situational context and/or on the mentioned examples like in (3.13), or alternatively 2) the speaker may build more specific labels by adding further linguistic material to simple labels, in order to identify more precisely the reference like in (3.12) and (3.14) – (3.17). Henceforth the former strategy will be referred as 'simple label' while the latter as 'complex label'.

From one side of the spectrum, a simple label directly designates a broader but less specific category than the one the speaker wants to communicate. Being more general, it requires a lower cognitive effort and less encyclopaedic knowledge to retrieve possible members of the category. Nevertheless, at the same time, because its reference is often too broad or too distant from the context (cf. section 2.3.1.3), it requires some further cognitive effort in linking and tailoring the category to the specific context; that is, to correctly understand what type of members are indeed relevant to the specific context (cf. ad hoc concepts, Wilson and Carston 2007).
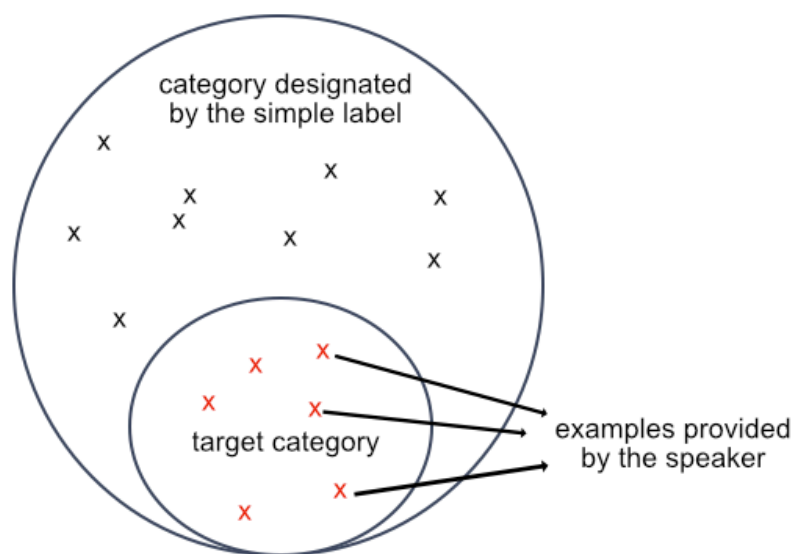
Figure 3.2: Simple labels

The speaker chooses a label designating a category that encompasses different members (the red and black crosses in Figure 3.2). However, the target category the speaker wants to actually communicate is a specific context-dependent subset. Therefore, to correctly direct the interpretation, he or she provides some examples chosen from the target category (the red crosses). Consider the following example.

(3.18) *Kōgekisei-o*   *mashi**tari***   *suru*   *yōna*   *fukusayō-ga*   *deru*

aggressiveness-ACC   increase:**TARI** do   like   side effect-NOM   go out

kanōsei-ga   *aru*   *koto-ga*   *wakatta.*

possibility-NOM   exist   thing-NOM   understand:past

"It was found that there is the possibility of experiencing side effects such an increase of aggressiveness."

In (3.18) the author describes antidepressant drugs which can cause violent behaviour. For example, it is reported the case of a man who beat his wife with a metal object while under the effects this type of drugs. Therefore, while *fukusayō* "side effects" makes reference to a well-known broad category, the entire exemplifying construction designates a context-dependent sub-category of side effects, namely 'side effects such an increase of aggressiveness'. Therefore, the author does not refer to common side effects of medications like, for example, nausea or fever, but with a specific type of side effects, that is, instances of aggressive and violent behaviour. It is possible to come to this conclusion because both the examples and the context direct the interpretation towards the construction of the correct category. In fact, despite being part of the broader category "side effects", instances of aggressive and violent behaviour are not the best choice to represent it, that is, they are not prototypical examples of it (see Taylor 1995: 40). The fact that the author deliberately chooses an example that is not prototypical of the broader category designated by the label shifts the attention of the readers to a particular sub-set of members, around which the new category is thus created. Moreover, also the context provides some other clues to infer the defining property of the category (e.g., the description of the type of drugs the article is describing, the episodic case of the man attacking his wife), and consequently to understand that the author makes reference to a more specific category that the one designated by the explicit simple label.

Another example was provided in (3.13), repeated here as (3.19).

(3.19) *Sofutobanku*   *mobairu wa,*   *meiwaku*   *mēru*   *taisaku-no*   *kyōka-**ya***   *kakin*

Softbank   mobile   TOP   trouble   email   against-GEN   reinforce-**YA**   billing

taikei-no   *henkō*   *toitta*   *shisaku-o*   *happyōshita.*

system-GEN   change   such as   measure-ACC   announce:do:past

"Soft bank mobile announced measures such as a reinforcement of the anti-spam policy and a change in the billing system."

Label: *shisaku* "measures"

Here, the author describes implementing measures introduced by Softbank (a Japanese telecommunications and Internet corporation) regarding the email service. The simple label *shisaku* "measures" encompasses different types of policies taken to achieve or alternatively avoid something. However, the article refers only to those measures that are relevant in the context of email service, such as, precisely, *meiwaku mēru taisaku no kyōka* "the reinforcement of the anti-spam policy" and *kakin taikei no henkō* "change in the billing system". Therefore, the mentioned examples and the context direct the reader to correctly interpret the target category, identifying what members are part of the category, and what members encompassed by the label *shisaku* "measures" are not relevant in that specific context.

The analysis of the context is particularly crucial whenever simple labels referring to broad categories are used, in order to grasp correctly the reference. Nevertheless, this is especially true in those cases in which the examples seem to be insufficient to correctly direct the interpretation. Let us consider the following example.

(3.20) *Messēji-o   jushinsuru   to,   namae, denshimēruadoresu, jūsho   **nado***
Message-ACC receive:suru   when   name   email address   address   **NADO**
no   kojinjōhō-o   kinyūsuru   web   fōmu-ga   hyōjisareru.
NML   personal information-ACC   fill:do   web   form-NOM   display:do:pass
"When a message is received, it is displayed a web form to fill out with personal information, such as name, email address, address and so on."

If we merely consider the exemplifying construction, (i.e., the label and the members taken as examples of the category), we may be tempted to interpret it as a reference to the broad category "personal information". However, if we consider the entire context – as the reader would do in a real-life situation – the result will be quite different. The article in (3.20) describes different phishing systems, that is, fraudulent websites that steal sensitive information. Specifically, in (3.20) a fraudulent web form requiring personal information is described. Then, the next sentence (3.21) explains that a further form asking for credit card is displayed, thus providing further details on the kind of sensitive information the article is actually describing.

(3.21) *Sarani   kurejittokādo-no   bangō,   kigen,   sekyuritikōdo-o*
Further   credit.card-GEN   number   expiration date security.code-aCC

kinyūsuru	*fōmu ga*	*hyōjisareru.*
fill.in:do	form-NOM	display:do:PASS

'Further, it is displayed form to fill out with the number of the credit card, the security code and the expiration date.'

Hence, taking into consideration the broader context, the reader is actually able to tailor the category "personal information" focusing only on those personal data that are relevant to that specific context: personal information that could be useful to hijack credit card details online.

However, being a continuum, the simple label pattern is not a prerogative of simple nouns. Consider the following example.

(3.22)	*Indoneshia-ya*	*Kankoku, Chūgoku nado*	*Ajia*	*kakkoku-o*	*rekihōsuru.*
	Indonesia-YA	Korea, China NADO	Asia	country-ACC	tour:do

"(Hillary Clinton) tours Asian countries such as Indonesia, Korea and China."

In the sentence above, the label *Ajia kakkoku* "Asian country" is not a simple noun, but a compound [27]. Yet, it refers to a broader category than the one the author wants to communicate. The examples and the context help to identify the Asian countries the article is actually referring to: on one side, examples focus only on far east Asian countries, on the other, the context describes the visit of Hillary Clinton to Asian countries that are major trading partner with the US.

Generally speaking, simple labels are based on words (or at least small phrases) whose referent is broad and encompasses a larger variety of members than the ones the speaker wants to designate. Therefore, while the elaboration of the label and the retrievement of its members may require a smaller effort, the real issue is to anchor and tailor the referent of the label to the actual context.

On the other side of the spectrum, more detailed labels (or complex labels) are the result of the speaker's deliberate effort to create a label to directly designate the category he or she wants to communicate. As stated before, the lack of a stable association between contextual-dependent categories built on the spot and fixed linguistic expressions allows the speaker to choose from a wide range of linguistic constructions in order to designate the

---

[27] *Ajia* is actually a so-called "no-adjective", that is, nouns which are typically translated to adjectives in English and other languages. In this case, it is used without the adjective marker *no*, thus it should be considered as a noun.

specific category she has in mind. With every added linguistic adjunct, the reference to the category grows more specific, less open to different interpretations.
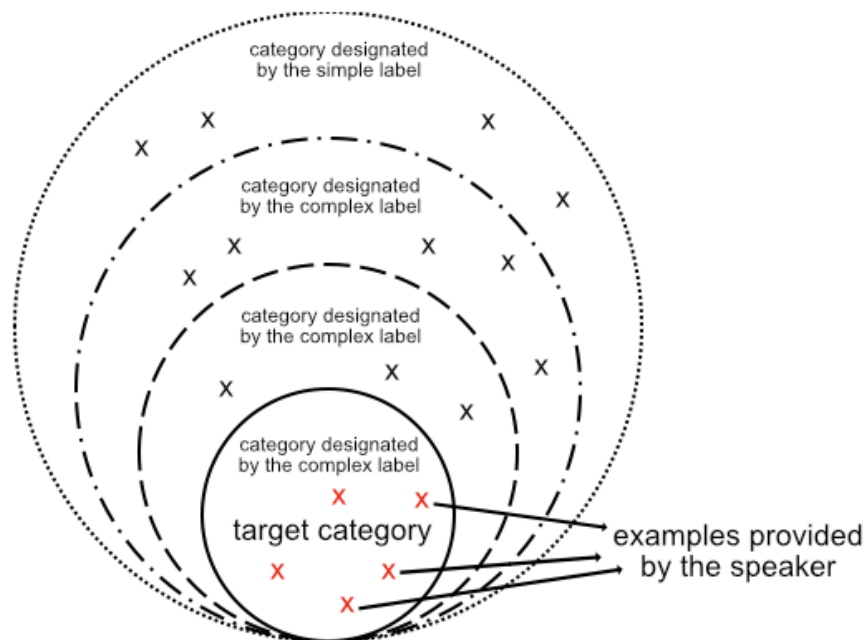
Figure 3.3: Complex labels

Providing more complex and thus detailed labels, the category they designate gets more and more precise, helping the hearer to understand what members are relevant in the context (the red crosses in Figure 3.3), and what members are not relevant and should thus be excluded (the black crosses). In some cases, the (complex) label designates exactly the target category, in other cases, the label is broader than the target category, but still more precise than a simple label.

Let us consider the following example.

(3.23) *Imēji-gata*   *supamu-no*   *hassei*   *nado* *supamu*   *haishin*
Image-model spam-GEN   occurrence   NADO spam   distribution
gijutsu-ga   *kōmyō*   *kashiteoru.*
technique-NOM   ingenious   change.into:STA
"Spam delivery techniques, such as (the occurrences of) image-type spam have grown more sophisticated"
Label: *supamu haishin gijutsu* "spam delivery techniques"

Here, the simple label *gijutsu* "techniques" with the addition of a concrete example (*Imēji-gata supamu no hassei* "occurrences of image-type spam") could have been sufficient to designate the category, especially in an article referring to spamming issues. However, the writer chooses to be more specific, creating an equally specific label to designate the category by means of a compound: *supamu haishin gijutsu* "spam delivery techniques".

(3.24)   *Iryōyōgu ya*   *gakkō-no*   *karada*   *sokutei*   *kiki*   *nado no*
         medical.use.tool-YA school-GEN   physical measurement   equipment NADO NML
         hoken   *yōgu.*
         health   equipment
         'Health gears such as school physical measurement tools and tools used by doctors.'

In the sentence above, the label *hoken yōgu* (health gears) makes reference to a specific sub-category than the simple label *yōgu* (tools), nevertheless, its referent is still broader than the target category (i.e., the types of medical tools manufactured by the company Nittokagaku), therefore the mention of concrete examples facilitates the process of exclusion of irrelevant members.

In other cases, there is a perfect match between the target category and the referent category of the complex label, like in (3.16) or in the following example.

(3.25) *Femininna*   *pinku*   *wa,*   *orenji-ya*   *ierō*   *nado*   *no*   *bibiddo karā-ni*   *wa*
         Feminine   pink   TOP   orange-YA   yellow NADO   NML   bright colour-DAT   TOP
         matchisuru   *bannō*   *karā.*
         match:do   all-purpose   colour
         'Feminine pink is an all-purpose colour that match well with bright colours such as orange and yellow.'

Here the author refers to pink as a colour that can easily go with many other different colours and styles. In particular, it is said that it goes well with other bright colours (i.e., *bibiddo karā*) such as yellow and orange. In this case, the target category and the category designated by the complex label match. In fact, there are no clues in the context to suggest that we are dealing with a further context-dependent sub-category of bright colours (i.e., hypothetically, warm bright colours). The target category is indeed 'bright colours'. In these cases, we may wonder what is the purpose of mentioning examples, since the label already

designates successfully the target category. We will return to this issue in section 3.5, after the analysis of the examples.

Complex labels designate specific set of items that are very similar (or may even coincide) to the target category the speaker has in mind. Therefore, since the referent is more precise, the tailoring of the category requires a minor effort. In other words, having a more precise instruction, it is easier for the hearer to exclude the irrelevant members and to focus only on those that are relevant. Nevertheless, the identification and elaboration of the category designated by the complex label may prove to be more difficult depending on how detailed the label is. In fact, more detailed labels may turn out to be too complex or opaque to be easily comprehensible. Moreover, lacking a stable association between linguistic expression and category, the choice of what and how many linguist adjuncts to add is completely up to the speaker's subjectivity. This issue is well described by the example in (3.12), repeated here as (3.26).

(3.26) *Doraggu&doroppusuru*   *dake*   *de*   *shashin-o*   *appurōdoshitari,*
     drag&drop:do   only   STR   photo-ACC   upload:do:TARI
     daburukurikku-de   *suraidoshō-o*   *saiseisuru*   *nado, shoshinsha*   *demo*
     double click-STR   slideshow-ACC   play:do   NADO   beginner   GDR:also
     kaitekini   *riyōdekiru*   *kantanna*   *sōsa.*
     simply   use:POT   simple:AGG   operation
     'Simple operations that can be used comfortably even by beginners', such as playing a slideshow with a double-click, uploading photos by simply dragging and dropping and so on.'

It is true that 'simple operations that can be used comfortably even by beginners' well specifies what members are relevant and what members are irrelevant, however without the help of the concrete examples, it may be difficult to understand what it stands for, that is, what set of members should be considered in the first place (even before understanding what it is relevant and what is not).

In this sense, we can then identify two competing strategies that act on the continuum. On one hand, the simpler the label used, the easier the elaboration and the identification of the category designated by the label, but also the harder the interpretation and contextualization towards the target category. On the other hand, the more complex the

label used, the easier the process of contextualization, but the harder the elaboration and identification of the category designed by the label[28].
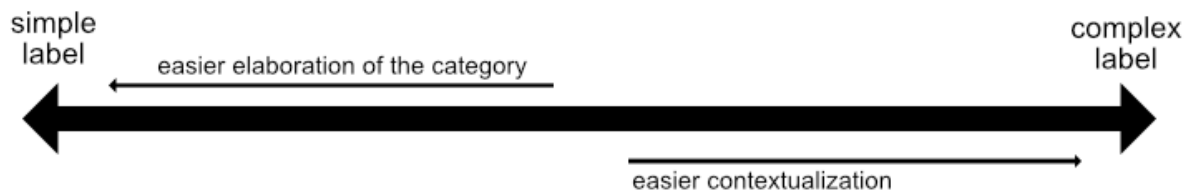
Figure 3.4: The continuum between simple and complex labels

These differences in the elaboration and contextualization of simple and complex labels seems to have consequences at the linguistic level, specifically in the position of the labels with regards to the mentioned examples (cf. Figure 3.5)
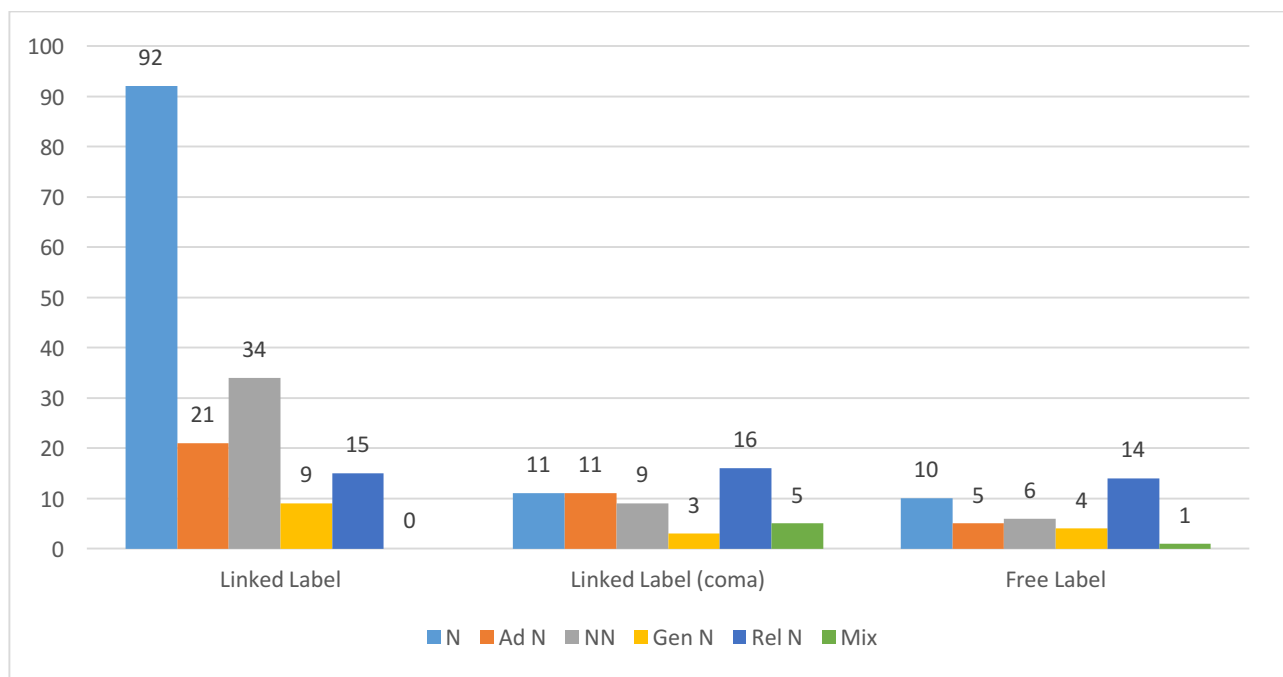
Figure 3.5: Category labels: position and syntactic types

---

[28] This relation between simple and complex labels (and their inference) exhibit some correlation with the notion of "hidden complexity". Bisang (2009, 2013) refers to the notion of "hidden complexity" to describe extremely simple surface structures that need more inferential effort from the perspective of the hearer. This type of complexity is motivated by economy, while the other opposite strategy (i.e., using overt grammatical markers) is motivated by the need of explicitness.

Data on the correlations between the position of the labels and their syntactic structure show a rather interesting picture. Nearly the totality of simple labels expressed by simple nouns frequently occur directly linked to the example(s) by means of a linguistic connector (81%). The other two patterns are much less frequent (the free label pattern represents only the 9%). This tendency becomes less strong with the increase of syntactic complexity. In fact, on the other side of the continuum, all three patterns are equally attested with label expressed as noun phrases containing relative clauses: the linked label pattern represents the 32% of the total, as well as the free label pattern.

This tendency becomes more evident if we consider the distinction between simple labels (i.e., simple nouns) and complex labels (i.e., noun phrases), as shown in Figure 3.6 below.
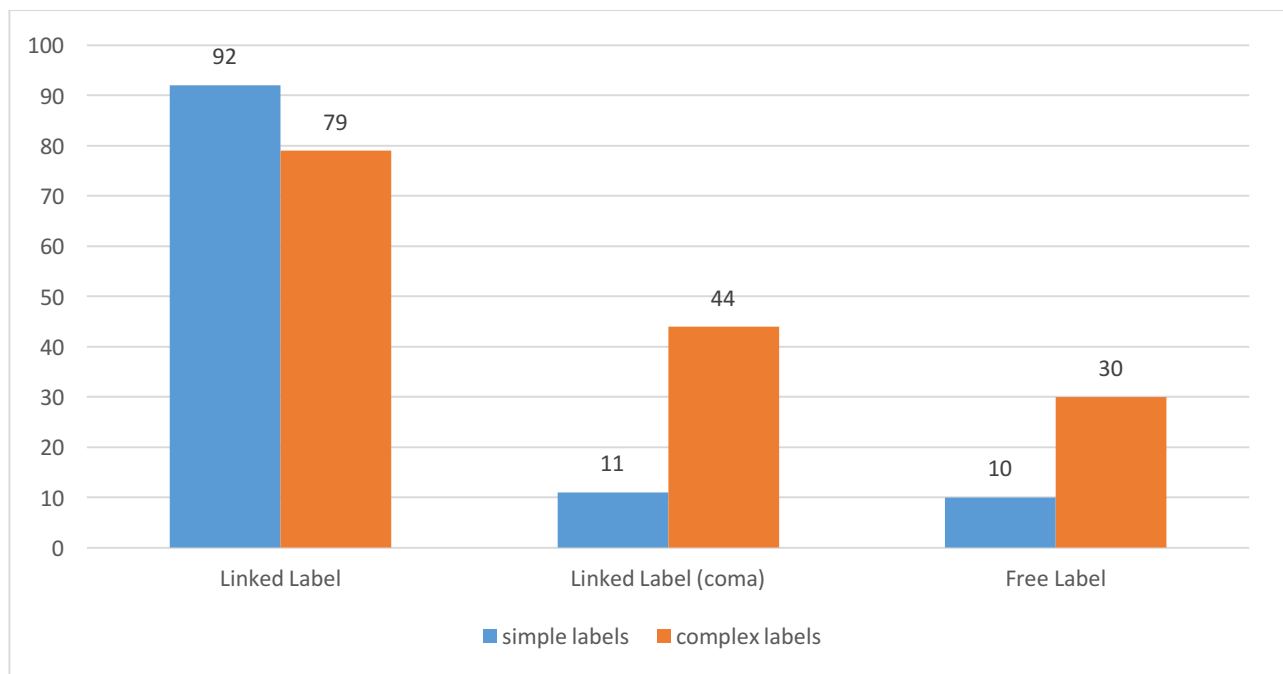
Figure 3.6: Positions of complex labels in relation to examples.

To explain this tendency, we should take into consideration the categorization processes underlying these types of labels. Simple labels refer directly to broader categories which should be reinterpreted and tailored according to the context (and, in case they are provided, to the examples). For this reason, we may say that there is a (sometimes, strong) discrepancy between the category actually represented by the label and the one the speaker wants to communicate. Simple labels alone can successfully make reference to the targeted and context-relevant category only if the context is clear enough to ease the proper interpretation (cf. Wilson and Carston 2007). Otherwise, the reference might be ambiguous. Therefore, simple labels are less independent and examples are a crucial factor in their

interpretation, because they actively help to identify more precisely the actual members of the category and its defining features. Accordingly, firmly connecting the label and the example(s) by using a linguistic connector that highlights the relationship "X is an example of Y" points out more clearly that the targeted category is not the one represented by the label alone, but by the entire construction and that the interpretation should be mediated by the linked examples. On the contrary, data show that complex labels can function also in isolation, that is, without being directly linked to the example(s).

The higher frequency of the pattern which consists of a linked label by means of a comma can be explained at the structural level: attacking a list of examples to a complex label by means of a linguistic connector may result in the creation of a very complex single construction, one that may be too complex to be elaborated correctly. On the contrary, the comma allows to differentiate the status of the two, keeping them close enough to be processed together, but without them become a single construction. Nevertheless, at the same time, the use of a comma can also be motivated at the cognitive level, further highlighting the potential independence of complex labels. In fact, there is no need to elaborate the label and the examples as a single construction strictly linked to each other, because the label does not need the mention of examples to refer to the category: the label alone is designed by the speaker specifically for representing a specific context-relevant category. Therefore, in this case, there are two potentially independent elements that work together to make the reference more precise and specific.

At this point a question may arise: if it is possible to create specific ad hoc labels to designate ad hoc categories, why would a speaker provide also examples? Again, we argue that the reason lies in the fact that context-dependent categories do not reside as knowledge structures in long-term memory, with all that entails at the linguistic level. Without fixed representations, there cannot be stable association between categories and linguistic expressions that designate the categories. In other words, there are no designated words or other linguistic expressions to unequivocally and specifically refer to those categories. These labels are created on the spot according to the speaker's ability to summarize the defining feature of the category, in order to allow the hearer to successfully infer the category. Therefore, while the speaker might think that simple labels are too general to describe a specific contextual-dependent category, the speaker might also think that "ad hoc" labels are not clear enough to designate them without any risk of misunderstanding.

Nevertheless, the identification of complex labels is a further demonstration of the need to revise the notion of lexicalized category, because it shows that even in the absence of a

stable association between categories and lexicon, the speaker can identify and use a specific label to make reference to a functional category built on the spot. This is not the case for simple labels, because, as we have seen, when they are used in isolation, they simply refer to common taxonomic categories and it is only through the mediation of the context and of possible examples that they can be used to represent ad hoc categories. In other words, with the mere identification of simple labels, we might be tempted to think that labels are actually a prerogative just of common categories, and that they can be used to represent more specific context-relevant sub-sets acting as abstract models around which to reshape the new category. This interpretation would in fact support the idea that only common categories can be lexicalized and that all other categories can at most be inferred through their cognitive and linguistic mediation. On the contrary, the evidence provided by the identification of complex categories becomes effective proof of the possibility of lexicalizing any kind of categories, in the sense that every category can be described by a specific label (and thus, not only by means of other linguistic strategies, e.g., general extenders). Of course, since these are categories built on the spot and that are not characterized by permanent representations in our memory, also their lexicalization suffers from the same instability: the speaker will identify the best label to represent the category depending on the specific context and the discursive goal.

3.2.2 GENERAL LABELS VS. SPECIFIC LABELS

The distinction between simple labels and complex labels cannot be considered solely in terms of syntactic structures; it must also be evaluated in terms of semantic properties, specifically general-specific semantic relations (cf. section 2.2.1.3). However, while the syntactic structures of the labels can be judged regardless of the examples (and thus of the category) provided, the distinction between general and specific labels strongly relies on the nature of the actual target category.

As we are dealing mainly with context-dependent categories (and not with natural categories), it is unlikely that a specific superordinate noun (that is, the one on the immediately higher level in the taxonomic hierarchy) is available. More frequently, there are semantic gaps that should be filled with other linguistic strategies, such as more complex expression or, moving higher up the hierarchy, more general simple labels. Therefore, as it was stated in the previous sections, intuitively, we may argue that simple labels expressed by simple nouns tend to be more general as they refer to a general (that is, broader) category, while more complex labels expressed as (different types of) noun phrases tend to be more

specific are they refer to specific sub-categories of the general categories. In other words, *karā* "colours" is more general (and thus less specific) than *bibiddo karā* "bright colours", *gijutsu* "techniques" is more general (and thus less specific) than the compound *supamu haishin gijutsu* "spam delivery techniques", and so on. While this general observation may be rather uncontroversial, some specific points should be raised as well.

First of all, as previously pointed out, the general-specific semantic relation is a relative and not an absolute concept, as in our case, it depends strongly on the category the speaker wants to communicate. For this reason, it cannot be taken for granted an exact correspondence between the syntactic complexity of the label and the general-specific parameter. Consider the following examples.

(3.27)   *Femininna*   *pinku*   *wa,*   *orenji-ya*   *ierō*   *nado*   *no*   *bibiddo karā-ni*   *wa*
    Feminine   pink   TOP   orange-YA   yellow nADO   NML   bright colour-DAT TOP
    matchisuru   *bannō*   *karā.*
    match:do   all-purpose   colour
    'Feminine pink is an all-purpose colour that match well with bright colours such as orange and yellow.'

(3.28)   *Kakei-o*   *sekkyokutekini*   *minaoshitari,*   *yoking-o*   *fuyasu*
    family.finance-ACC   aggressively   review:do:TARI   deposit-ACC   increase
    iyoku-ga   *waku, nado*   *omowanu*   *kōka-ga*   *umareteiru*   *yō*   *desu.*
    desire-NOM   grow NADO   unexpected   effect-NOM   born:STA   like   COP:POL
    'Unexpected effects appear such as reviewing aggressively the family finances and being fill with the desire to increase the bank account.'

In (3.27), the world *bibiddo* "bright" is an adjective, but it is used without the adjective marker *na* (that is, *bibiddo-na karā*), as a simple transliteration of English expression "bright colours". In (3.28), the word *omowanu* "unexpected" falls into the category of nouns and verbs acting prenominally (i.e., *reitaikei*), that is, they directly modify the nouns they are attached to. The suffix *–u* in *omowanu* indicates the attributive form. Hence, from the point of view of syntactic complexity, the label in (3.27) and the one in (3.28) are comparable and we may expect the same degree of specificity. However, when they are examined taking into consideration also the examples (that is, the members of the target category and thus the category itself), they appear to behave differently. While in (3.27) the label is specific to the point of designating a category that coincides with the one the author wants to

communicate, in (3.28) the label is very broad and does not specify the target category. In fact, the label only suggests how the mentioned examples should be conceived (that is, as unexpected effects prompted by having money issues), rather than identifying precisely the defining property. It follows that the general-specific semantic relations should be always considered in relation to the specific target category they are called to designate.

Secondly, it is important to consider also the discursive goal of the speaker. Consider the following:

(3.29) Tōshin-de wa,   tetsugaku-ya shinri,   keizai,   hōgaku   nado

report-LOC TOP   philosophy-YA psychology   economics   law   NADO

no   hiroi   bunya-de,

NML   wide   field-LOC

'In a wide range of fields such as law, economics, psychology, philosophy.'

Here the speaker deliberately wants to refer to a very broad category, therefore, the label has been chosen accordingly. In other words, while the label *bunya* "fields" is a simple general label[29], it designates exactly the category the speaker wants to communicate.

Finally, even more than the syntactic complexity of the labels, the general-specific parameter is a continuum and not a strict distinction. This is especially true while considering context-dependent categories, since the semantic taxonomic hierarchies are full of semantic gaps, forcing thus the speaker to create alternative expressions which are not always easy to compare in terms of general/specific.

For all these reasons, we will use this parameter carefully while examining the properties of the examples, investigating each occurrence in isolation, thus without trying to create an arbitrary scale. Generally, a label will be considered general or unspecific if, in a hypothetical taxonomic hierarchy, it is possible to place the same label in one of the highest nodes (or in the highest mother-node, cf. section 2.3.1.3). Labels that are placed in daughter-nodes close to the examples will be considered specific.

3.2.3 SIMILARITY-BASED VS. FRAME-BASED CATEGORIES

Regarding the overall coding of the category, we also need to address whether the nature of the shared category property can play an important part in the lexicalization process.

---

[29] Adjectives that do not qualify the category but only focus on the number of elements (i.e., quantity) are not considered as part of the label (cf. section 2.2.1.3).

As explained in section 1.2.4, the defining property can depend on some intrinsic similarity between the linked elements (cf. Joosten 2010: 32) or on some extrinsic contiguity within some accessible frame (cf. Lakoff 1987). We labelled[30] the former as similarity-based categorization like in (3.30) and the latter as frame-based categorization like in (3.31).

(3.30) *Dejitaru-kamera-ya*   *pōtaburumediapurēya,*   *dendō-haburashi*   *toitta*
        digital-camera-YA   portable-media-player   electric-toothbrush   such as
        minsei-kiki-ni   *mo*   *tekiyō-ga*   *susundeiru.*
        consumer-products-DAT   also   adaptation-NOM   progress:STA
        'The adaptation is progressing also in consumer products such as electric toothbrushes, portable media players and digital cameras.'
        Label: *minsei-kiki* "consumer products"

(3.31) *Mata,* *"yōfuku"*   *"seikatsu shikin"*   *"pasokon"*   *nado,* *jisseikatsu-ni*
        also   clothes   living expenses   PC   NADO   real.life-DAT
        chokketsushita   *bōnasu-no*   *tsukaikata*   *mo*   *medatteimasu.*
        direct.connection:do:PAST   bonus-GEN   way.to.use   also   stand.out:STA:POL
        'In addition, it also stands out usages of the bonus directly connected to real-life such as "clothes", "living expenses", "PC".'

The question that naturally arises is: do these different types of categories exhibit different patterns of lexicalization? To answer this question, we need compare this distinction with the degrees of syntactic complexity of the attested labels, as shown in Figure 3.7 below.

---

[30] Regarding the difference between similarity-based and frame-based categories, see also Mauri 2016.
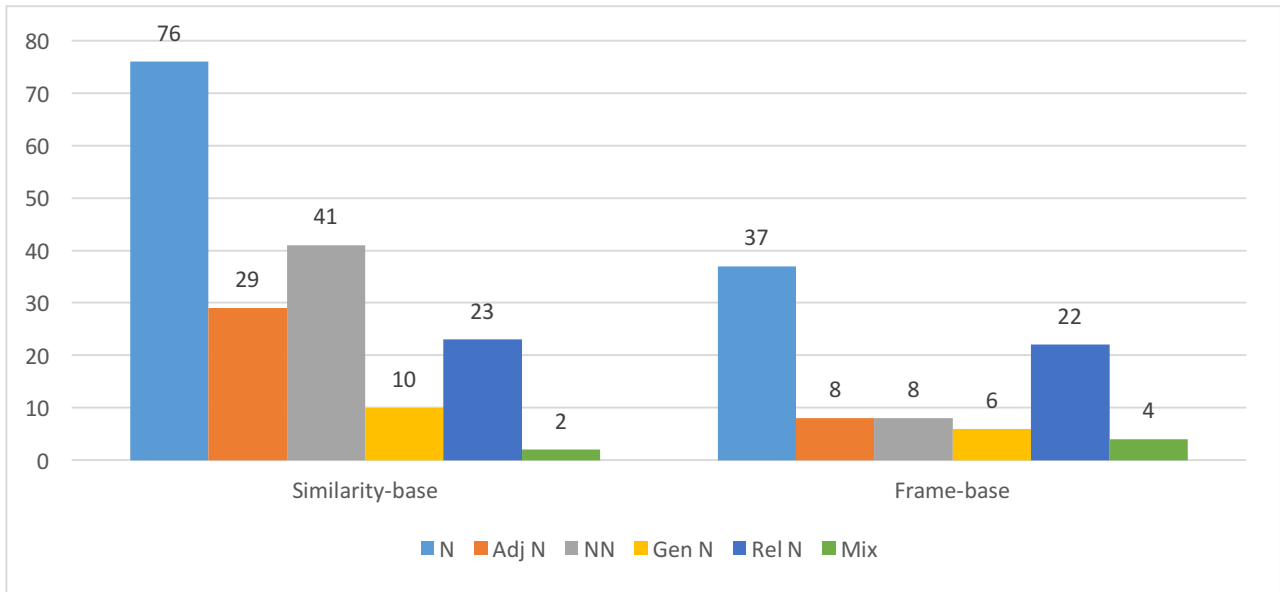
Figure 3.7: Types of categorization and syntactic types of labels

Similarity-based categories are frequently lexicalized by means of simple labels (that is, simple nouns) or short complex expressions such as compounds or noun phrases containing adjectives, like in (3.30). On the contrary, frame-based categories are mainly lexicalized by means of simple labels (42%) or noun phrases containing relative clauses (25%) which resemble the paraphrases indicated by Barsalou (1983). Other strategies are less frequent.

While in both cases, simple labels represent a widespread strategy, the semantic properties of these labels change accordingly to the type of category. Simple labels that designate similarity-based categories tend to be less general than those used to designate frame-based categories. For instance, in the first case, simple labels tend to highlight at least part of the property that characterizes the category, like in (3.32)[31].

(3.32) *Kōsui   nado   no   fureguransu   niyotte   nioi-o   kakusu.*
        perfume   NADO   NML   fragrance   by means   smell-ACC   hide
        '[It] hides the smell by means of fragrances such as perfume.'

In case of natural categories[32] (cf. Rosch 1973), the category designated by the simple label may even coincide with the target category. In our corpus, there are no instances of

---

[31] Other examples were provided in the previous sections, in (3.35), (3.36) and (3.37).

[32] As we noted in section 1.2.4, natural categories are prototypical instances of similarity-based categories.

natural categories, however there are instances of more specific sub-set in which the reference of the simple label almost coincides with the target category:

(3.33) *Konchū-rui   kara   kogata-no   honyūrui-ya   chōrui nado no*
insect-kind   from   small size-LK mammal-YA   bird   NADO NML
sekitsuidōbutsu   made, samazamana dōbutsu-o   esa-ni   shiteimasu.
vertebrate   till   various   animal-ACC   feed-DAT   do:STA:POL
'[It] feeds on various animals, from insects to vertebrates such as small mammals and birds.

Here, the target category is not 'vertebrates', but a more specific sub-set, namely 'small vertebrates'. Nevertheless, the simple label is still more precise in its reference than in most cases attested in our corpus.

On the contrary, when they designate frame-based categories, simple labels are very general, so much so that, in some cases, they do not specify concretely the shared property, but only how the examples should be conceived in that specific situational context. Consider the example provided in (3.34).

(3.34) *Konzai-kankyō-de   riyōsuru   baai wa,   yūzā-akaunto-o*
mixed-environment-LOC   use:do   case TOP   user-account-ACC
sōgoni   tōrokusuru   nado no   kufū-ga   hitsuyō-ni   naru.
mutually   register:do nado NML   schemes-NOM necessary-DAT   become
'When you use [it] in a mixed environment, schemes such as to register mutually the user accounts are required.'
Label: *kufū* "schemes"

Here, the author refers to a network system called NAS, a file-level computer data storage server connected to a computer network and providing data access to a heterogeneous group of clients. He states that whenever this system is used in a mixed environment (e.g., using devices made by different brands), some precautions are necessary, such as registering mutually the user accounts. In other words, he refers to a category of actions in which the exemplars are strongly associated by their contiguity with a frame, that is, the use of this type of network system in a mixed environment. To refer to this category of actions, he uses a broad label *kufū* "schemes", which helps him to avoid the identification and lexicalization of the shared property of the examples since it is strongly dependent on the context.

Another example is provided below.

(3.35) *Yappari*　　　*gakkō-no*　　*tesuto-de*　　*tensū-ga*　　*warui to*　　*o-kāsan-ni*
　　　after all　　　school-GEN　test-LOC　　score-NOM　　bad　when　HON-mother-DAT

　　　miseru-no-ga　　　　*iya*　　　　*da*　*to*　　*omottari,*　　*misenai*　　*yōnishitari*
　　　show-NML-NOM　　　　unpleasant　COP　QT　　think:TARI　　show:NEG　　try:TARI

　　　toka　*sōiu*　*keiken-ga*　　　　　*atta*　　　　*node,*　　　*Takane-no*　*kimochi*
　　　TOKA　sort of　experiences-NOM　exist:PAST　because　　Takane-GEN　feeling

　　　wa　　*yoku*　*wakarimasu.*
　　　TOP　　well　understand:POL

　　　"After all, because there were experiences such as when the score in the school test is bad
　　　and you think that showing [the score] to your mother is unpleasant and you try not to
　　　show [the score to your mother], I understand well the feelings of Takane"

The utterance above is part of an interview to a Japanese dubbing actress regarding her
last performance. In the previous sentence, the interviewer asks her if she can resonate with
the character she dubs, an average 10-year-old boy, given the fact that she has worked as
a voice over actress since she was eight. Here, she explains that despite her work in the
show business industry, she had a very regular childhood. To do this, she makes reference
to a particular narrative frame: when kids get bad scores at school and are afraid to tell the
parents about it. Therefore, she creates and refers to a category of situations that should be
typical of childhood and she labels it with the very broad and unspecific word *keiken*
"experiences". In this case, relationship between the exemplars strongly depends on the
narrative frame rather than on some intrinsic similarity among the events.

Interestingly, this applies even for slightly more complex constructions, such as noun
phrases containing adjectives. Consider the example in (3.28), repeated here as (3.36).

(3.36) *Kakei-o*　　　　　　*sekkyokutekini*　　*minaoshitari,*　　　*yoking-o*　　*fuyasu*
　　　family.finance-ACC　aggressively　　　review:do:TARI　　deposit-ACC　increase

　　　iyoku-ga　　　*waku, nado*　*omowanu*　　*kōka-ga*　　　*umareteiru*　*yō*　　*desu.*
　　　desire-NOM　　grow　NADO　unexpected　effect-NOM　　born:STA　　like　COP:POL

　　　'Unexpected effects appear such as reviewing aggressively the family　　finances and being
　　　fill with the desire to increase the bank account.'

Not only, the reference of the noun *kōka* "effects" is broad and just directs the hearer to
conceive the mentioned examples as effects, but also the adjective *omowanu* "unexpected"
does not qualify concretely the shared property of the target category.

Therefore, we can identify tendencies regarding the lexicalization of different types of categorization processes:

1. Similarity-based categories tend be lexicalized through simple nouns or short linguistic expressions (mainly compounds). More specific sub-set are lexicalized by means of complex constructions, such as mainly noun phrases containing relative clauses.
2. Frame-based categories tend to be lexicalized through complex expressions (mainly noun phrases containing relative clauses) or, alternatively, general simple nouns with a very broad reference. In some cases, the simple label merely indicates how the examples should be conceived in the specific context.

These tendencies can be explained if we consider that it is much easier to identify and lexicalize a common property based on an intrinsic similarity, especially in the case of categories of concrete objects (as we will see in section 3.3), where people may rely on physical features. In this regard, it is not surprising that natural categories - which are prototypical similarity-based categories (cf. section 1.2.4) – are also the more lexicalizable type of categories.

On the contrary, to identify and to lexicalized the shared property of frame-based categories is a much less straightforward process, since they strongly rely on the context in its entirety. For this reason, speakers may prefer complex and thus detailed labels, or very general labels that suggest how the examples should be conceived.

3.3 LINGUISTIC PROPERTIES OF THE EXAMPLE(S)

After examining the attested types of labels, at this point it is necessary to shift our focus on the other half of the exemplifying construction, namely the examples. In this section, we would like to provide an overview of the modality of usage and linguistic coding of exemplars whenever they are used in combination with an explicit category label.

Before examining the analysis of the parameters that related to the linguistic encoding of examples, we feel it is necessary to provide some clarifications as to why some issues that will be raised in this section fit into the discussion of categorization in an interesting fashion. In particular, in cognitive psychology, most experiments and actual examples of categorization provided in studies almost exclusively focus on categorization of objects, excluding almost completely the categorization of actions or situations (more generally,

events). For example, Rosch herself describes a category as "a number of objects which are considered equivalent" (Rosch et al. 1976: 383). In general, this is particularly true for natural categories studies, whose main concern is the way in which the brain categorizes concrete objects that surround us (from animals to inanimate objects). Thus, being primarily concerned with "categorizations which humans make of the concrete world" (1976: 382), this approach inevitably ends up excluding items that affect a more abstract level.

Indeed, a type of categorization that does not exclude a more abstract level was introduced by Barsalou (1983, 2010) through his notion of ad hoc categories. Despite the fact that he himself is mainly concerned with categories of objects (e.g., by way of illustration, he cites "things to take on a camping trip," and "possible costumes to wear to a Halloween party" which both encompass object-type members), the notion of ad hoc category allows to expand the discussion, by moving from a type of categorization mainly aimed at describing and organizing the concrete world, to a goal-driven type of categorization, which does not imply any exclusion of a more abstract dimension. On the contrary, a greater capacity for abstraction is required, as they are associated with items that sometimes seem to have little in common at first glance (e.g., the already cited example "things to use to kill a roach"), but that can correlate on a more abstract level (i.e., a final purpose). Therefore, from the perspective of ad hoc categorization, there is no difference between "things to take on a camping trip" and "things to do on a camping trip": even if the first represents a category of objects and the second represents a category of actions, both are indeed types of categorization aimed at achieving a specific purpose.

Therefore, although - at least in theory - there are no constrain on what types of categories people can actually create and make reference to, research in the field of cognitive psychology seem to have focused solely on a certain type of categorization process, namely, the one that is driven from concrete objects as exemplars.

This shows some consequences also at the linguistic level. Studies on lexical semantics (cf. Cruse 1986) investigate thoroughly the taxonomic lexical hierarchies with regards to categories of objects (e.g., products, animals, furniture). However, the same cannot be said with respect to categories of events. This is primarily due to the influence of categorization theories within cognitive psychology (cf. Rosch et ali. 1976). However, this approach to lexical taxonomy may give the false impression that the process of lexicalizing a category is possible only with categories of entities. In this regard, it is important to keep in mind that the process of labelling a category requires a precise cognitive effort by the speaker in order to identify 1) the defining feature that characterizes the members of the category, 2) a

linguistic word or expression that can encompass all of them accordingly and effectively. So, while the simple act of categorizing may just require a perception of similarity to group together different elements, the process of lexicalizing the category requires a much more precise idea of the similarity among items, to the degree of being able to identify and codify it by means of linguistic material.

Taking all these issues into account, we believe that observing the syntactic and semantic properties of the examples that appear in our occurrences can provide useful insights on the process of selecting the exemplars in order to build categories.

3.3.1 Syntactic properties of the example(s)

The first parameter concerns the syntactic properties of the members taken as examples, specifically whether the examples are encoded by noun phrases or by verb phrases.

Total Categorization | Total Lexicalized categories

216; 33% | 448; 67% (NP / VP) | 77; 29% | 189; 71% (NP / VP)
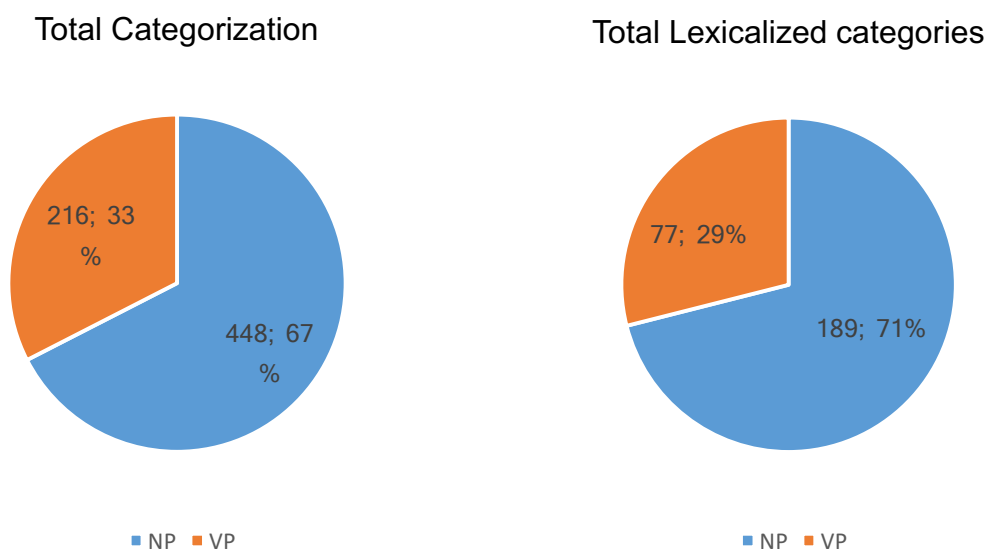
Figure 3.8: Distribution of examples expressed by noun phrases and verbal phrases in lexicalized categories

The issue is even more accurately described by the following figures which show the distribution for each exemplifying strategy. As can be seen, the percentages of distribution tend to be constant, despite being influenced by the fact that the presence of a label appears less frequently in combination with *tari* (36 of 148, or 24 percent of the total number of occurrences).

Table 3.3: Distribution of examples expressed by noun phrases and verbal phrases in lexicalized categories

	NP	VP
ya	62 (170)	0
nado	86 (170)	15 (24)
tari	0	36 (148)
toka	41 (108)	26 (44)
Total	189 (448)	77 (216)

The figures above suggest that both strategies are indeed possible, but examples – generally – tend to be coded by noun phrases. This result, however, should be interpreted with caution because of the modalities of usage of each exemplifying strategy (cf. section 1.3.2). First, it is important to notice that two strategies that we are examining exhibit structural constraints, namely, *ya* is used only to joint noun phrases, while *tari* can be used solely to connect verbal phrases. On the contrary, *toka* does not seem to follow any specific grammatical rule or tendency: not only it is used as non-exhaustive connective to join noun phrases and verbal phrases indistinctly, but it is used as a general extender or as a connector as well. While it is important to consider these issues, we think that this trend in favour of noun phrase examples is still worthy of consideration, especially in the light of the fact that the total occurrences of *ya* (168 occurrences) and those of *tari* (148 occurrences) offset each other, avoiding any strong bias towards a strategy rather than the other. Similarly, apart from the occurrences in combination with *ya*, *nado* still shows a high versatility, occurring also as general extenders at the end of lists of verbal phrases or even clauses, sometimes even in combination with *tari*.

The tendency towards noun phrase examples may have an impact on the categorization process, because different syntactic types of examples may represent different types of categories (i.e., typically nouns designate entities, while verbs designate events or processes) and more importantly, different modalities of encoding categories. Consider the combination of the data regarding the syntactic types of examples and the syntactic types of labels:
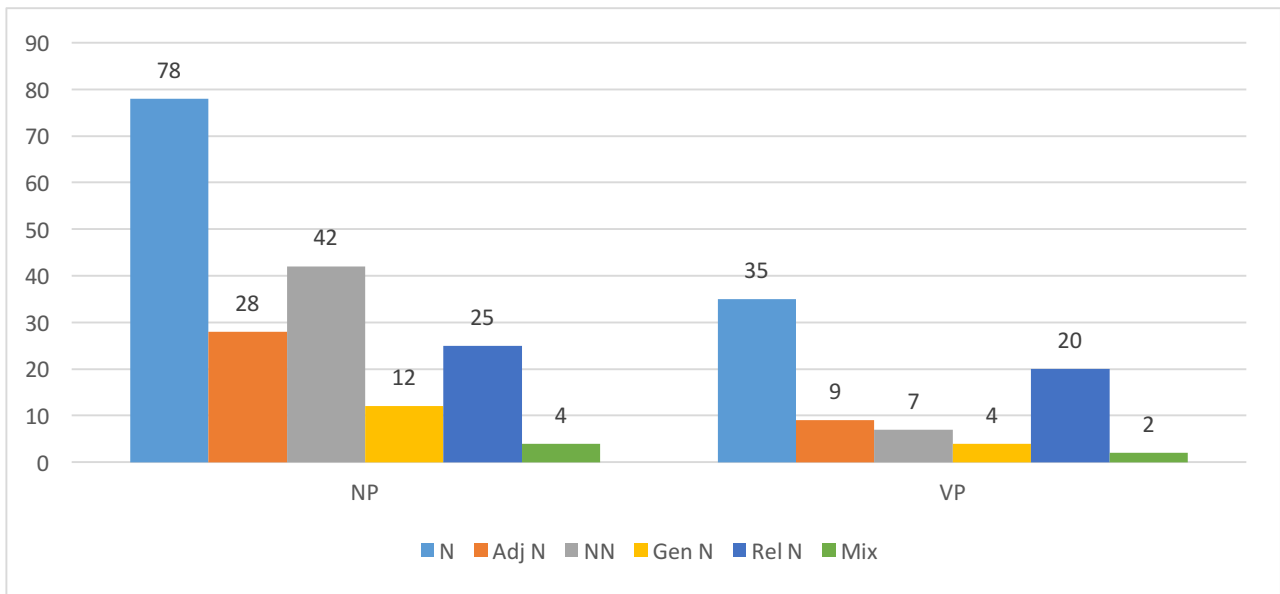
Figure 3.9: Lexicalized categories: syntactic types of examples and syntactic types of category labels
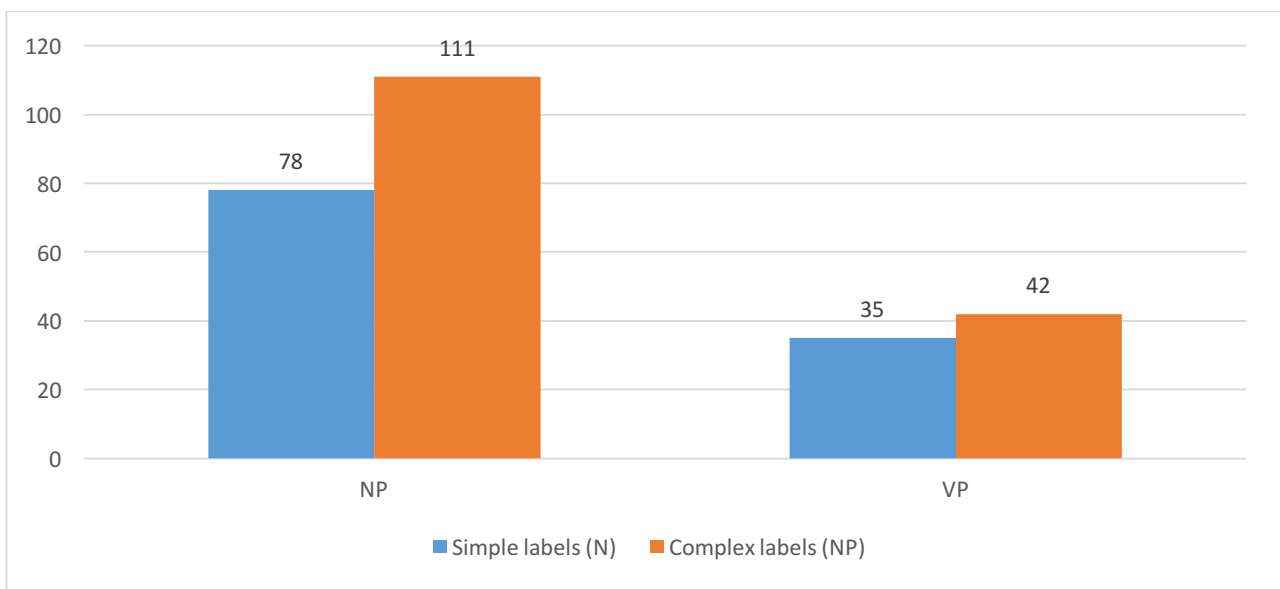
Figure 3.10: Lexicalized categories: syntactic types of examples and syntactic types of category labels (simple vs. complex labels).

While examples expressed as noun phrases exhibit a more varied general picture, categories constituted by examples expressed as verbal phrases are more likely to be lexicalized through a very simple label (i.e., simple noun) or alternatively, through a very complex label (in most cases, a noun phrase containing relative clauses). More specifically, out of the 36 occurrences of examples encoded by verbal phrases and lexicalized through a simple label, in 30 occurrences shows unspecific labels (e.g., "needs", "schemes",

"experiences", "effects", etc., cf. section 2.3.1.3), compared to the actual category designated by the examples, as shown in (3.37) and (3.34), repeated here as (3.38):

(3.37) *Saisho wa   bijinesuyūsu-ga   chūshin   desu   ga,   sonogo,*
at first TOP   business-NOM   center   COP   but   after that
kojin-de   mo   dōga-o   mitari,   netto-o   tsukattari
individual-STR also   video-ACC   see:TARI   net-ACC   use:TARI
toitta   nīzu-ga   dete   kuru   no   da   to   omoimasu.
such as   need-NOM   appear:GRD   come nml   COP   QT   think:pol
'At first business was at the center, but, after that, I think that some needs has emerged, such as using the net or watching videos also in private.'
Label: *nīzu* "need"

(3.38) *Konzai-kankyō-de   riyōsuru   baai wa,   yūzā-akaunto-o*
mixed-environment-LOC   use:do   case TOP   user-account-ACC
sōgoni   tōrokusuru   nado no   kufū-ga   hitsuyō-ni   naru.
mutually   register:do   NADO NML   schemes-NOM necessary-DAT   become
'When you use [it] in a mixed environment, schemes such as to register mutually the user accounts are required.'
Label: *kufū* "schemes"

This pattern does not seem to be a prerogative of simple labels. Even examining syntactically more complex labels, we may note that frequently the noun at the head of the nominal group is broad and unspecific. Furthermore, in some cases, not even the linguistic adjuncts add true qualitative specification:

(3.39) *Eria-o   48-jikan inaini taiōshitari,   denchi-no   muryō torikae-o*
area-ACC   48-hour within respond:do:TARI   battery-GEN   free   replacement-ACC
yattari toitta,   jimichina   doryoku-ga   sukoshizutsu okyakusama-ni
do:TARI   such as   steady   effort-NOM   little by little   customer-DAT
tsutawatteiru.
be.introduced:STA
'Little by little, steady efforts are introduced (to help customers), such as support the area within 48-hour and free replacement of the batteries.'
Label: *jimichina doryoku* "steady efforts"

111

We argue that this fact can be ascribed mainly to two reasons. The first one is a mere terminology issue. Typically, verbal phrases designate events, that is, processes (cf. Givón 2001, Langacker 1987a, 1991b) As it was noted, while there are plenty of taxonomic lexical hierarchies regarding categories of entities, this is not the case regarding categories of events. In other words, languages tend to lack superordinate terms to designate categories of events. This does not seem to discourage or prevent the process of lexicalization of the category (cf. Figure 3.7), but it affects the modalities in which this type of category is lexicalized. Speakers tend to use highly complex labels (curiously, they mainly use the type of label that encompasses another verb) or, alternatively, very broad and unspecific terms that need to be interpreted depending on context.

The second reason is a cognitive issue. Givón (2001) notes that, contrary to nouns, verbs often exhibit considerable complexity. This is due to the fact that (prototypical) events or actions involve several distinct participants (e.g., the agent, the patient, the dative), "all distributed over space and each an individuated, spatially compact, temporally durable entity in its own right" (2001: 52). Similar remarks have been made by Langacker (1987a, 1991b). This observation is even more interesting when we consider that in most cases (74 out of 77 occurrences), what we have called examples encoded by verbal phrases are not merely simple verbs, but groups consisting of a verb and at least one modifier designating a participant to the event. It follows that if we define categorization as a process based on the ability to find similarities between elements in order to group them together, identify the shared property in order to shape an appropriate category label may require a greater cognitive effort, since events are multi-dimensional elements that take into account many different correlations (e.g., the distribution through time). For this reason, their lexicalization requires more detailed labels as shown in (3.40):

(3.40) *Go-hōbi-o*   *erabu toki-ga*   *ichiban*   *tanoshī!*
       hon-reward-ACC   choose time-NOM   best   enjoyable
       Tatoeba,   *shikaku*   *shiken-ni*   *gōkakushitari,*   *dai*   *purojekuto-o*
       For example   qualification   test-DAT   success:do:TARI   big   project-ACC
       seikōsasetari.
       success:do:CAUS:TARI
       'When you choose a reward are the most enjoyable! For example, when you pass a qualification exam, when you make a big project succeed or other similar occasions.'
       Label: *gohōbi o erabu toki* "times when you choose a reward"

In the sentence above, the author refers to a category of situations when people succeeded in something and thus may reward themselves with the purchase or consumption of indulgent products. In order to lexicalize this category, the author chooses a complex label, namely *gohōbi o erabu toki* "times when you choose a reward".

Otherwise, the speaker may also use very broad unspecific labels which do not qualify concretely the category, but *i)* they seem to act as placeholders to stress the presence of other non-mentioned elements (e.g., the non-exhaustivity) like in (3.38), or *ii)* they simply emphasize not the shared property of the examples, but the way the examples should be conceived in that specific context, like in (3.28), repeated here as (3.41):

(3.41) *Kakei-o   sekkyokutekini   minaoshitari, yoking-o   fuyasu*
       family.finance-ACC   aggressively   review:do:TARI deposit-ACC increase

       iyoku-ga   waku, nado   omowanu   kōka-ga   umareteiru   yō   desu.
       desire-NOM   grow   NADO unexpected   effect-NOM   born:STA   like   COP:POL

       'Unexpected effects appear born such as reviewing aggressively the family finances and being fill with the desire to increase the bank account.'

       Label: *omowanu kōka* "unexpected effects"

Here the author does not use a label to identify and communicate the common shared property of the examples, but how they should be conceived in this specific context, that is, they should be considered as (unexpected) effects.

It is important to note that this is peculiar of examples expressed by verbal phrases, and not of simple labels. When examples encoded by noun phrases are lexicalized by means of simple labels, these labels are almost never broad and unspecific. Despite needing contextualization, they seem to concretely specify the category, as shown in the following example:

(3.42) *Fukuzatsu   buki   shisutemu   kara   kōhīmēkā-ya   hifuku*
       complex   weapon   system   from   coffee.machine-YA   clothing

       nado no   juhin   made, arayuru   buppin-ni   oyondeimasu.
       NADO NML   supplies   till   every   article-DAT   extend:STA:POL

       '[The standard] covers all articles, from a complex weapon system to supplies such as clothes and coffee makers'.

       Label: *juhin* "supplies"

Although the label *juhin* "supplies" needs to be contextualized (i.e., supplies that armies may need in war-zones), it is still more specific and characterizing than the labels used in (3.37) or (3.38), since in a hypothetical taxonomic hierarchy, the label would occupy the immediately higher level than the members of the category (cf. section 2.3.1.3).

A final point of interest is the correlation between types of categories (i.e., similarity-based vs. frame-based category) and syntactic types of examples (i.e., noun phrases vs. verbal phrases), since, as we saw in section 2.2.3, both factors can influence the choice of a category label. Generally speaking, we may say that there is at least a loose correlation between them, since categories of entities tend more easily to share some intrinsic similarity, while categories of events often rely on the context to establish a common topic (cf. Lakoff 1971) or property among the examples. This is partially due to the set of interconnections profiled by verbs (cf. Langacker 1987a, 1987b, 1991b), which are strongly connected to the narrative frame (e.g., the distribution over space and through time).

Unsurprisingly, examples expressed by verbal phrases and frame-based categories seem to share the same tendency regarding the choice for category labels, that is, both prefer highly complex labels or general simple labels.

Nevertheless, this correlation is not always perfect. The example in (3.40) is an instance of category of events sharing a property based on an intrinsic similarity (i.e., to succeed in something troubling). Vice versa, also in categories of entities, the shared property may be extrinsic and based on the contiguity of the members recurring within a specific frame (cf. the more prototypical instances of ad hoc categories in Barsalou 1983, e.g., things to use in a trip to Alaska).

Therefore, we may wonder what happens when these two influential factors do not co-occur, as shown in (3.40). To answer this question, we should compare the example (3.40) and the example (3.41). Contrary to (3.40) where the events used as examples share some intrinsic similarity, in (3.41) the shared property is based on the contiguity of the examples recurring within a specific frame, that is, actions that may occur when people are considering a deposit-linked mortgage. In this case, the events can be processed as similar only when the narrative frame is considered (cf. Lakoff 1987). While in both cases, the examples are expressed by verbal phrases, the speakers choose different types of labels: a complex label in the similarity-based case (3.40), and a very unspecific label in frame-based case (3.41). It follows that, while the similarity-based category of events is still lexicalizable with a certain amount of precision (despite requiring a complex label), the frame-based category of events is likely the most difficult type of category to lexicalize, to the point that speakers prefer to

rely on broad and unspecific simple labels, that do not designate precisely the target category.

3.3.2 Semantic properties of the example(s)

As it was mentioned at the beginning, one of the objectives of this section is to determine from what types of exemplars (e.g., entities, events, properties) context-relevant categories may be driven. More specifically, to investigate this fact, we cannot rely only on syntactic parameter, but it is crucial to investigate also the semantic properties of the examples. For example, despite the opposite being far more likely, a noun phrase can still refer to actions by means verbal nouns. This is especially true in Japanese, because of the Sino-Japanese vocabulary or *kango*, i.e., a portion of the Japanese vocabulary that originated in Chinese or has been created from elements borrowed from Chinese. Sino-Japanese words are almost exclusively nouns, some of which are verbal nouns. The latters can be used as verbs simply by appending *suru* (e.g., *kōkan* "replacement", *kōkan suru* "to replace, to change"). Naturally, the usage of these verbal nouns may result in designating a category of actions and not entities. Therefore, while verbal phrases taken as examples represent categories of events[33], the connection between noun phrase examples and categories of entities may not be so straightforward and should be checked further.

The first figure shows the percentages of examples that are entities and examples that are events.

---

[33] In Japanese, it is possible to noun a verb, by means nominalizing makers such as *no* or *koto*. Both these strategies are incompatible with *tari*, which is added directly to the second base of the verb. The *no* strategy is also incompatible with *ya, nado, toka*. The *koto* strategy does not exhibit structural constrains, but it is not attested in my corpus. Being mainly a corpus of written language, Sino-Japanese action nouns are widely preferred.

Total Categorization

160; 24%
255; 38%
249; 38%

Total Lexicalized categories

67; 25%
109; 41%
90; 34%

■ Abstract entities ■ Concrete entities ■ Events
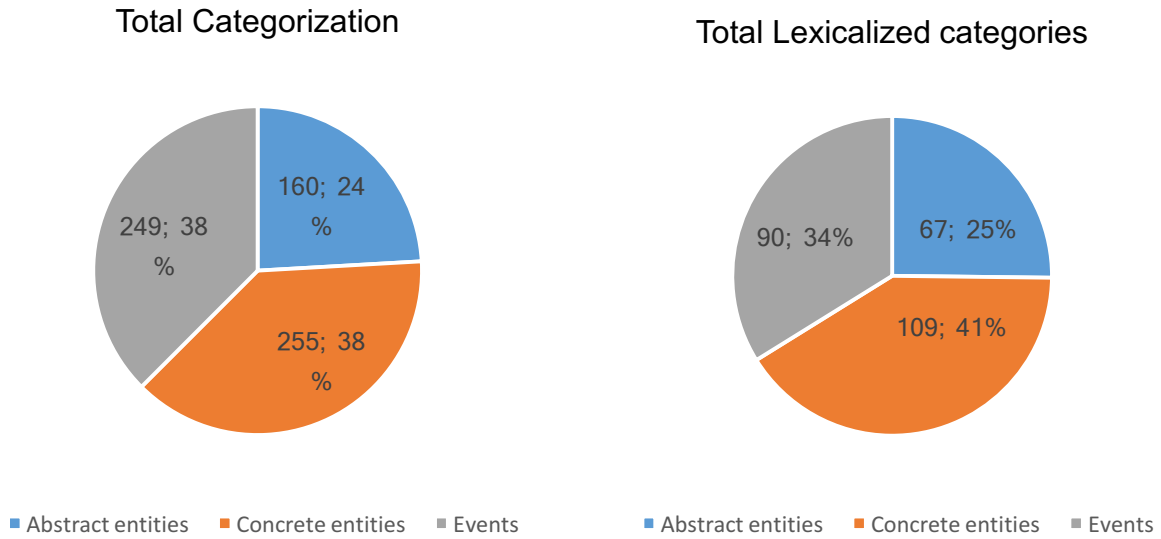
Figure 3.11: Semantic properties of the examples (lexicalized categories)

Table 3.4: Semantic properties of the examples (lexicalized categories)

	Abstract entities	**Concrete entities**	**Events**
ya	19 (58)	37 (95)	6 (17)
nado	32 (63)	51 (97)	18 (34)
tari	0	0	36 (148)
toka	16 (39)	21 (63)	30 (50)
Total	67 (160)	108 (255)	91 (249)

Semantic properties of examples seem thus to confirm what had already emerged, although sketchier, from the analysis of the syntactic properties. First, it proves that categories may be driven from exemplars of any kind: inanimate entities, concrete entities and events[34] (see Givón 2001: 56). This means that compared to what have been described by Rosch and - albeit less strictly - by Barsalou, the concept of categorization should be redefined in a broader sense. Human brain can actually categorize elements of a different nature, the only constraint being the presence of a shared property that justifies their being grouped together, or - to quote Rosch - that justifies why they are "considered equivalent" (Rosch et al. 1976: 383). This constraint is likely one of the reasons for the high frequency of concrete entities used as examples. Even beyond theories concerning cognition, understanding the preference for the categorization of concrete entities over other types of exemplars seems rather straightforward, since they are characterized by features much easier to identify and label. Not only inherent features such as colour, shape, size and so

---

[34] The only exception seems to be properties, which - at least in our corpus - are not attested

on, but also more ad hoc features, which depend on the narrative frame, such as the modality of use (e.g., a slipper and a newspaper may have the same function in the frame of killing a mosquito). This fact is relevant not only in the process of building categories, but also in the process of lexicalizing. In fact, since the shared property can be more easily inferred, it is also easier for the speakers to create more or less specific (but always characterizing) labels for the category they want to communicate. This becomes evident in the ways categories of concrete entities are lexicalized, as shown in Figure 3.12.
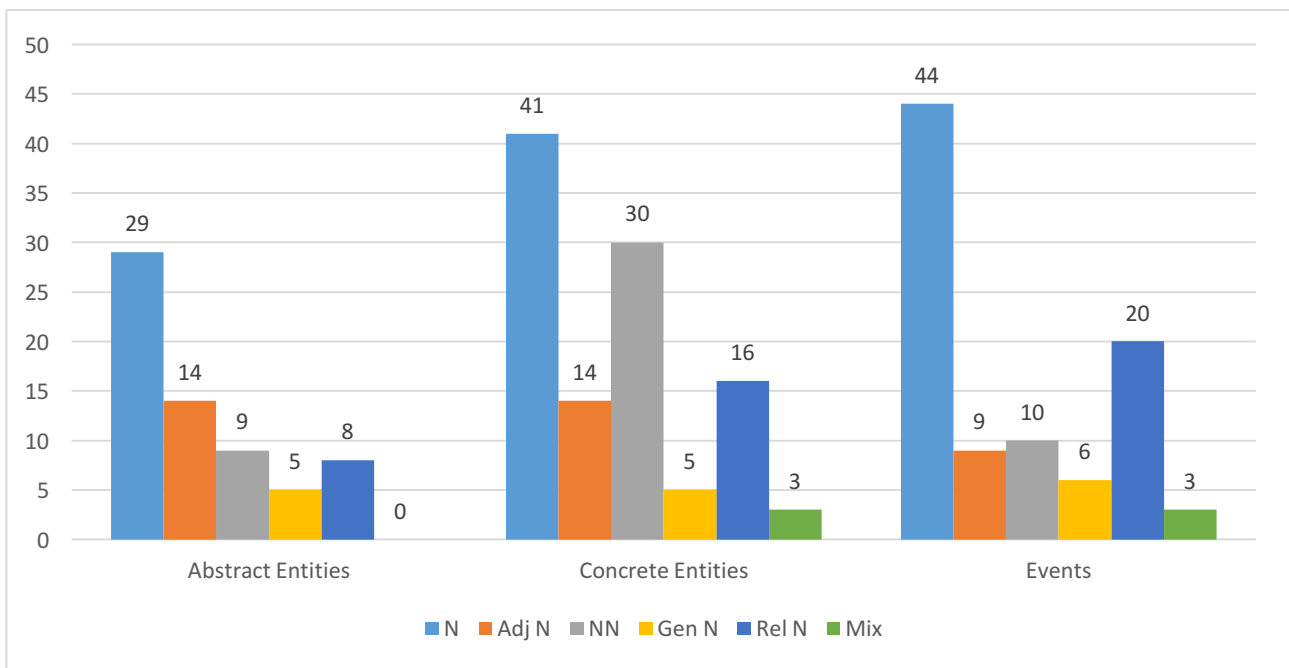
Figure 3.12: Syntactic types of labels and semantic properties of examples (lexicalized categories)

First of all, even without considering the syntactic complexity, the labels used to lexicalize categories of concrete entities are more specific and characterizing than those used especially for events. Consider the following constructions attested in my corpus:

(3.43) *Kaiun-ya*   *tekkō*   *nado no*   *meigara-ni*   *jōshō-ga*
        marine transportations-YA   iron and steel NADO NML   stock-LOC   rise-NOM
        medatsu.
        stand out
        'Rise in stocks, such as iron and steel, marine transportation and so on, stands out.'
        Label: *meigara* "stocks"

(3.44) *Saishin PV   eizō-ya   raivu   eizō   nado   no   dōga   mo   haishinshimasu.*
latest PV   clip-YA   live   clip   NADO   NML   video   also   distribution:DO:POL
'We also distribute videos such as the latest PV clips and live clips.'
Label: *dōga* "videos"

Another example was also provided in (3.42). In all these cases, despite being expressed by simple nouns, the labels always highlight precisely the shared property of the examples.

Moreover, the concreteness of the examples does not just facilitate the actual creation of category labels, but also their interpretation and elaboration. In fact, since examples act as bridges between the concreteness of the hyper-specific context and the abstractness of the labels and the category they designate, the usage of concrete entities as examples may actually help the processes of elaboration and contextualization of the labels and the categories. In fact, cognitive theories on categorization have spent a considerable time focusing on this type of categories (cf. Rosch et ali. 1976, Rosch 1978) because they constitute the most basic stimuli in our experiential environment. For this reason, as it will be shown in detail in section 3.5, they may facilitate the process of ascending from the abstract (i.e., the category) to the concrete (i.e., the context), through the contextualization and actualization of abstract category labels.

The comparison between the syntactic properties and the semantic properties of the examples allows us also to consider further the process of lexicalization of categories of events. In our corpus, 11 occurrences of verbal nouns (i.e., noun phrases that designate events, and not entities) used as examples are attested. Comparing these occurrences with those of verbs or verbal phrases used as examples, it emerges that the way examples are syntactically encoded may affect the process of lexicalization more that the type of exemplars. Previously in this section, we have seen that examples expressed by verbal phrases tend to be lexicalized by means of very complex labels or very broad unspecific simple labels. Based on this observation, we may assume that generally categories of events undergo the same issue whenever they are lexicalized. However, this does not seem to be the case. Out of a total of 11 occurrences of categories of events encoded by noun phrases (i.e., verbal nouns), in 6 occurrences the label qualitatively characterizes the category, as shown in (3.45):

(3.45) *Omoni   genkyū-ya   kaikoku   nado   no   chōkai-shobun-ga*
mainly   pay.cut-YA   reprimand   NADO   NML   disciplinary-action-NOM

kentōsareteiru.

consideration:do:PASS:STA

'Mainly, disciplinary actions are considered, such as pay cuts or reprimands.'

Label: *chōkaishobun* "disciplinary actions"

The use of verbal nouns is peculiar (cf. Langaker 1987) because they include the event structure, but it is embedded in the holistic conceptualization characterizing nouns, which, for instances, does not encode the evolution through time: "explode and explosion are not considered semantically equivalent: nominalization involves a conceptual reification whose character can be explicated with reference to the notional definition proposed for the noun and verb classes" (1987: 22). In this sense, the events thus profiled are much similar to abstract entities than to real events profiled by verbs: they are less complex to be processed, and thus we may argue that it is easier for the speaker to identify and select a characterizing label to designate the label.

Therefore, the process of lexicalization may be schematized by means of a continuum, as shown in Figure 3.13 below.
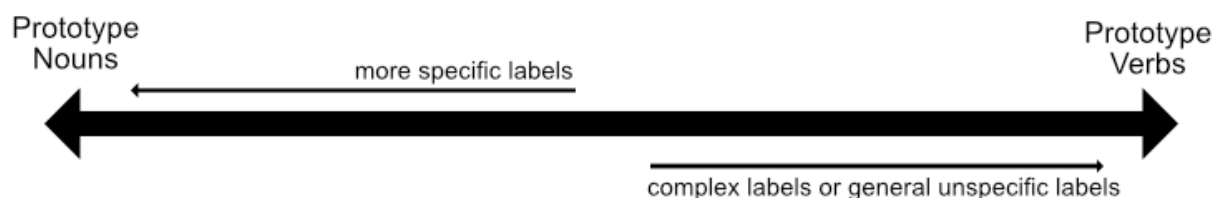
Figure 3.13: Continuum regarding the relation between the syntactic types of label and the syntactic types of examples

The notions of 'prototypical nouns' and 'prototypical verbs' are taken from Givón (2001). He states that prototypical nouns are "concrete, and made out of relatively-durable materials. Their bundled — co-experienced — properties, such as size, color, shape or consistency, thus change relatively slowly as individual features as well" (2001: 51). For this reason, they are also stable and durable through time. On the contrary, prototypical verbs are "most typically events that involve concrete participant nouns" and "code rapid changes in either the state, condition or spatial location of some noun-coded entity" (2001: 52).

Therefore, we may hypothesize that the more the examples resemble to the prototypical nouns, the easier is for the speaker to create specific labels to designate the category. On

the contrary, the more the examples resemble to the prototypical verbs, the more lexicalization relies on complex labels or, alternatively, unspecific broad simple nouns. In the middle of this continuum, less prototypical instances such as abstract entities and events encoded as abstract entities seem to be less predictable, showing different lexicalization strategies.

Unfortunately, we do not have enough data to confirm more substantially these tendencies. We thus hope that they will be further investigated by future research (also on other languages), since they may prove that the way we encode the examples indeed affects the lexicalization process.

3.3.3 Number of examples

Finally, we would like to conclude this section with some remarks on the number of examples. In this regard, we identified three main patterns based on the number of examples provided by the speakers:

1. One example

(3.46) *Musen   kontorōrā-o   saiyōsuru   nado   kinō-o   kōjō*
        wireless   controller-ACC   use:do   NADO   function-ACC improvement
        'To improve functions such as using a wireless controller.'
        Example: 1) *musen kontorōrā o saiyōsuru* "using a wireless controller"

2. Two examples

(3.47) *Chūshajō-o   mitsukeru   kotsu-o   oshie   au   toka,*
        parking.lot-ACC   find   trick-ACC   teach:GRD   do   TOKA
        konsāto-o   yatteru   saichū   no   chatto nado, iron'na
        concert-ACC   do:STA   in the middle of   DET   chat   NADO   various
        mokuteki-ni   tsukaeru.
        purpose-DAT   useful
        '[It is] useful to various purposes, such as chatting in the middle of a concert or teaching tips on how to find a parking lot.'
        Examples: 1) *chūshajō o mitsukeru kotsu o oshie au* "to teach tricks to find a parking lot", 2) *konsāto o yatteru saichū no chatto* "to chat in the middle of a concert"

3. Three or more examples

(3.48) *Namae,   denshi-mēruadoresu, jūsho   nado   no   kojin-jōhō-o*
        name   email address   address   NADO   NML   personal-information-ACC

kinyūsuru   *web-fōmu-ga*   *hyōjisareru*.

entry:do   web-form-NOM   display:do:PASS

'A web-form is displayed to fill out personal information such as an address, name, e-mail address and so on.'

Examples: 1) *namae* "name", 2) *denshi mēruadoresu* "email address", 3) *jūsho* "address"

At a first glance, we may assume that the presence of an explicit label favours the use of fewer examples, since the label itself helps to elaborate correctly the examples in a specific context. Therefore, without the necessity to infer the property by comparison among the examples and because of simple linguistic economy reasons, we expected a higher frequency of the single example pattern. However, this is not the case, as shown in Figure 3.14 below.
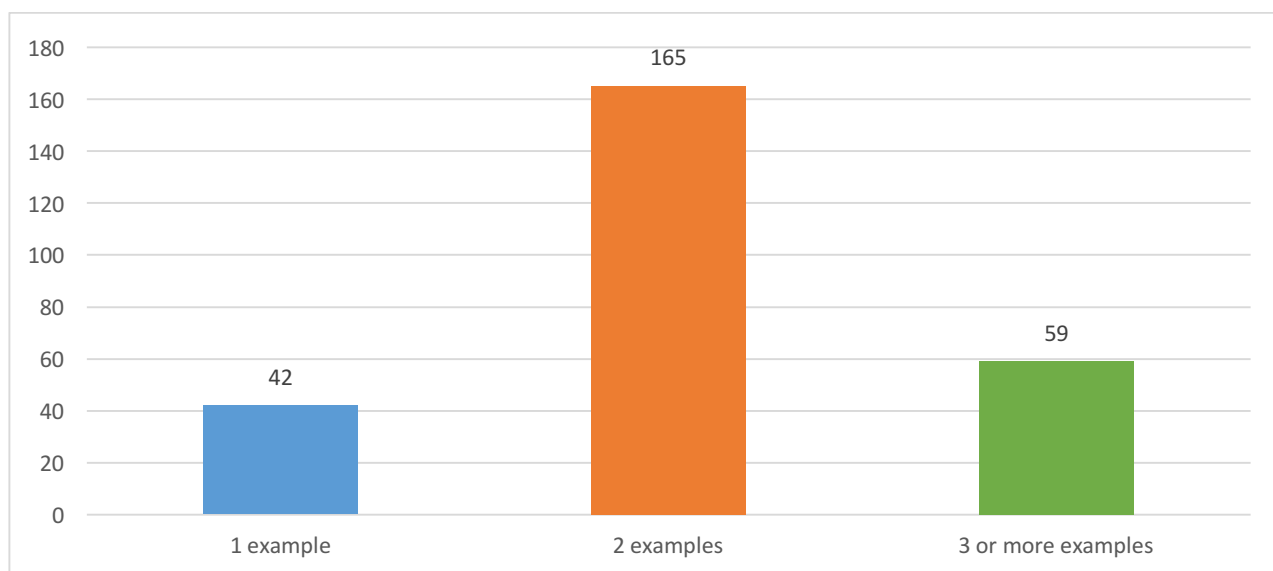
Figure 3.14: Number of examples in lexicalized categories

Despite the presence of an explicit category label, speakers tend to use two examples. Moreover, this tendency does not seem to correlate with any specific syntactic or semantic property of the examples or of the labels. It follows that it should be regarded as a general tendency of the exemplification process.

We argue that the preference for exemplifying constructions containing two examples is due to the necessity of comparing the examples to identify the common shared property in the specific context. While labels represent a help in this direction, they are still a (sometimes not specific enough) summary of the reasons why elements have been grouped together under the same category. On the contrary, the comparison of concrete examples in the

121

relevant context facilitates the hearers to infer more precisely the shared property and the entire category. In this sense, the presence of two examples instead of just one directs the inferential process towards the correct identification of the shared property. This general consideration will be elaborated further regarding non-lexicalized categories (cf. section 4.2), where the easiness of elaborating and comparing the examples is crucial to the construction of the category without the guarantee role of a label.

3.4 LINGUISTIC LINKAGE BETWEEN LABELS AND EXAMPLES

Whenever a category label is directly linked to the example(s), a linguistic connector is used. More specifically, with the eyeshadow term "connector" we encompass all those linguistic strategies that may encode the relation "X is an example of Y", and therefore can be used to connect labels and examples to make explicit their relationship.

While some of these linguistic constructions have been analysed specifically regarding the relation between category and example(s) (e.g., approximators, cf. Mihatsch 2007), there are no comprehensive studies (not even language-specific studies) that show what types of linguistic strategies can be used to perform this function.

In fact, our preliminary survey based only on Japanese shows great variation in the types of strategies that may be used as linguistic connector. Constructions encoding the relation "X is an example of Y" range from similative constructions (i.e., "Y like X", where Y is the category label and X the concrete example) to synthetic means like noun complement markers (see Horie 2000, 2003). Moreover, the varied picture to emerge seems to be peculiar of Japanese and strongly dependent on the different diachronic paths of the attested strategies (cf. Mihatsch 2009, regarding the grammaticalization of similative constructions is some Romance languages). Despite this plausible cross-linguistic variation, there may appear to be some systematic patterns in the selection of the linguistic material to form the set of connectors in a specific language.

This section aims to make a first step towards a more comprehensive study on this set of strategies. We will analyse the intra-linguistic variation as it is attested in our corpus data, while trying to sketch a preliminary typological survey of linguistic connectors.

The repertoire of strategies used as connectors in our corpus is given in Table 3.5.

Table 3.5: Linguistic connectors.

	ya	*nado*	*tari*	*toka*	Total
youna	1	0	4	0	5
mitaina	0	0	0	3	3
(tari) to iu	0	0	3	0	3
(nado) no	27	47	2	4	80
(toka) no	0	0	0	1	1
to itta	15	2	10	9	36
nado	7	21	1	5	34
toka	0	0	1	7	8
to iu youna	0	0	1	0	1
Total	50	70	22	29	171

Among these strategies, we may identify five main groups:

1. Similative constructions (e.g., *youna, mitaina*).
2. Noun complement markers (e.g., *no*, *to iu*) directly attached to a general extender. More specifically in our corpus the following combinations are attested: 1) *nado no* (the most widespread), *toka no*, *tari to iu*. Nevertheless, other combinations are still possible, as they are acknowledged in different grammars (e.g., *nado to iu*, cf. Kaiser et *ali*. 2001). In our corpus, there are no occurrences of these markers directly attached to the example(s) without the mediation of a general extender.
3. Fixed expression that have been grammaticalized as "such as" (e.g., *to itta*).
4. General extenders (e.g., *nado, toka*).
5. Combinations of two different strategies (e.g., *youna to iu*).

Beyond these, we cannot exclude the existence of further types of connectors.

In the following sections, we will examine each of these strategies in some details.

3.4.1 Similative constructions (*youna, mitaina*)

Similative constructions are linguistic constructions that express similarity (cf. Haspelmath and Buchholz 1998, Vanhove 2013, König and Umbach 2016). They are usually simple phrases, consisting of a similative marker (e.g., *like*) and a standard (to which something is compared). Similative constructions differ from equatives in the sense that while the latter expresses true equality, the former expresses an approximate similarity, focusing on quality instead of quantity.

This basic core function can be expanded to explain the relationship between a category and one of its examples. In this sense, the prototypical mentioned example represents the standard around which the category should be interpreted. Consider the following sentence:

(3.49) Psychiatric illnesses like depression.

Here, the similative construction suggest we should consider only the members of the category 'psychiatric illnesses' which are similar to the standard established by the mentioned example (i.e., 'depression').

In our corpus two similative marker used as linguistic connector is attested, namely *youna,* as shown in (3.50), and *mitaina,* as shown in (3.51).

(3.50) *Kōgekisei-o   mashitari   suru   yōna   fukusayō-ga   deru*
aggressiveness-ACC increase:TARI do   like   side effect-NOM   appear
kanōsei-ga   aru.
possibility-NOM   AUX
'It is possible that side effects like the increasing of aggressiveness come out.'

(3.51) *Kakaku-komu toka   attokosume   mitaina   kigyō.*
Kakaku-com   TOKA   at cosme   like   company
'Companies such as Kakaku.com and @Cosme.'

Similative constructions as connectors are not particularly widespread in our corpus (5%). More specifically, the similative marker *youna* is attested with *ya* and *tari* (and it is particular frequent with the latter), but not with *nado* and *toka*; while the similative suffix *mitaina* occurs only twice and in both occurrences, it is used in combination with the connective *toka*. This is not surprising since *mitaina* should be considered the colloquial equivalent of *youna.* Therefore, not only it is less used in a corpus mainly constituted by written texts, but it also occurs with the less formal exemplifying strategy among those here studied. As we will see further in this section, similative markers are not the only connectors that seem incompatible with *nado*.

3.4.2 *Noun complement markers*

In our corpus, only two noun complement markers are attested, that is, *no* and *to iu*. It is unclear if other Japanese noun complement makers (e.g., *koto*) can be used as well as

connectors. Evidences from grammars (e.g., Kaiser et *ali*. 2001) appear to be against this possibility.

Contrary to many European languages, Japanese is characterized by an extensive use of nominalization in forming complement cause [35] (cf. Ross 1973). Consequently, it productively uses "nominalizers" as a means to complementation. Among these, *no* is one of the most commonly used in Modern Japanese. Scholars formulated two hypotheses regarding its diachronic source: 1) it was originated from the genitive marker *no,* or 2) it is a truncated form of the lexical noun *mono* "concrete thing".

Generally speaking, *no* can be used to modify a following noun (N_2) with a preceding noun or clause (N_1 or S). In this sense, N_1/S describes or specifies N_2 in a variety of meanings, such as location or ownership. For instance,

(3.52) N_1 no N_2

> *Ototo*   *no*   *Masao-kun*
> younger brother   NML   Masao-HON
> 'Masao-kun[36], the younger brother.' (Kaiser et *ali*. 2001: 326)

By extension, we may say that, through the use of the nominalizer *no*, a category (i.e., N_2) can be described or specified by its example(s) (i.e., N_1/S). Nevertheless, *no* tends not to be directly attached to the example[37], but to a general extenders. Therefore, the typical pattern of usage is schematized as follow: [Example] [general extender] [no] [Label]. Here are some examples:

(3.53) Example NADO NO Label

> *Supekku*   *nado*   *no*   *shōsai wa*   *kōkaisareteinai.*
> specification   NADO   NML   detail   TOP   pubblish:PASS:STA:NEG
> 'Details, such as specifications, have not been published (yet).'

---

[35] From a typological perspective, in this sense, Japanese behave differently from many European languages, but curiously very similar to Korean (cf. Horie 2000)

[36] Kun is an honorific suffix used by people of senior status addressing or referring to those of junior status, mainly male children or male teenagers. As such, it is untranslatable in English.

[37] There is not a clear answer regarding this tendency. It is possible that without the general extender of the connective, the example(s) would be interpreted as a genitive complement(s) of the label, since *no* can also be used as a genitive marker. Nevertheless, the same tendency is found also with regards to *to iu*.

(3.54) Example[1] YA Example[2] NADO NO Label

Femininna	*pinku*	*wa,*	*orenji-ya*	*ierō*	*nado*	*no*	*bibiddo karā-ni*	*wa*
Feminine	pink	TOP	orange-YA	yellow	NADO	NML	bright colour-DAT	TOP

matchisuru	*bannō*	*karā.*
match:do	all-purpose	colour

'Feminine pink is an all-purpose colour that match well with bright colours such as orange and yellow.'

The combination of *no* with the general extender *nado* is the most widespread pattern in our corpus (47%). In particular, it seems the favourite strategy whenever the general extender *nado* is used at the end of the list of examples.

As for *to iu*, morphologically, it consists of the verb *iu* "to speak" or "to say" and the quotative marker *to*. Beyond its basic function as a quotative marker, *to iu* has undergone a grammaticalization process and eventually became a generalized noun complement marker[38], frequently occurring between a modifying clause or noun and its head noun (cf. Terakura 1983).

Generally speaking, it implies that the modifying clause or noun explains the content of its head noun (cf. Kaiser et *ali*. 2001), as shown in the following examples:

(3.55)

Honsha-o	*doko-ni*	*oku-no-ga*	*ii*	*ka*	*to*	*iu*
Headquarter-ACC	where-LOC	put-NML-NOM	good	Q	TO	IU

mondai	*mo*	*aru.*
problem	also	exist.

'There is also the problem of where best to position the headquarter.' (Kaiser et *ali*. 2001: 535).

In this sense, examples may occur as modifying clauses or nouns to further explain (or better, specify) the content of the category label, that is, the head noun. In other words, examples are considered the specification of the category label. Consider the following occurrence from our corpus:

(3.56) Example TARI Example TARI TO IU Label

Gyōmu-chū-ni	*bon'yari*	*shitari,*	*hyōjō-ya*
business-during-LOC	absent-minded	do:TARI	facial.expression-YA

---

[38] A similar grammaticalization pattern is also attested in Korean (cf. Horie 2003).

kōdō-ni	*genki-ga*	*nakunattari*	*to iu*	*henka-ga*	*miraremasu.*
behaviour-LOC	health-NOM	disappear:TARI TO IU		change-NOM	see:POT:POL

'It is possible to see changes, such as being absent minded during work and being no healthier in look and behaviour.'

Here, the examples (i.e., being absent-minded during work and being no healthier in look and behaviour) actually explain how to correctly interpret the very general noun "changes" according to the context.

If we consider the relationship between examples and category as the former being a specification of the latter, the overall high frequency of noun complement markers (especially *no*) used as linguistic connector (49%) is not surprising.

Nevertheless, differences in frequency between the two patterns are attested: in our corpus, the specific usage of *to iu* as a connector is not particularly widespread (2%) and occurs exclusively in combination with *tari* (and therefore verbal phrases/clauses).

3.4.3 Fixed expressions (*to itta)*

To itta is the second most widespread connector in our corpus (21%). Morphologically, it is the combination of the quotative marker *to* and the past tense of the verb *iu* "to say", so literally we may translate it as "said" or "called". Nowadays it is used also[39] as a fixed expression where the meaning of the verb *iu* "to say" is bleached, and comparable to the English connector *such as*.

A proof of this grammaticalization path is the fact that usually *to itta* is written using kana alone (100% of the occurrences in our corpus), that is, without the kanji of the verb *iu.* In fact, it is a general rule in modern Japanese that lexical items are written in kanji, while grammatical items such as case morphemes and in flectional endings in hiragana. Usually, when lexical items normally written using kanji become grammaticalized in certain contexts, they end up being written in hiragana instead[40]. Therefore, the Japanese writing system can be a good element to check grammaticalization paths. The fact that *to itta* is written using merely hiragana whenever it is used to link category labels and examples, means that in this context it is perceived more like a grammatical element than a lexical item.

---

[39] In many other contexts, *to itta* is the actual past tense of the verb *to iu*, that is, "said". Usually, whenever it is used to convey this meaning, it is written using the kanji of "to say".

[40] For instance, whenever the verb *miru* conveys its lexical meaning "to see", it is written using the kanji. On the contrary, whenever it is used as part of a compound verb with the sense of "try to do", it is written in hiragana.

Here are some examples:

(3.57) Example₁ YA Example₂ TOITTA Label

Keiei   *rinen-ya*   *kigyō*   *bijon*   *to itta*   *mono wa*   *hitsuyō*
management   idea-YA   business   vision   such as   thing TOP   necessary

desu   *ka.*
COP   Q

'Do we need things such as management philosophy and corporate vision?'

(3.58) Example₁ TARI Example₂ TARI TOITTA Label

Eria-o   *48-jikan inaini taiōshitari,*   *denchi-no*   *muryō torikae-o*
area-ACC   48-hour within respond:do:TARI   battery-GEN   free   replacement-ACC

yattari   *toitta,*   *jimichina*   *doryoku-ga*   *sukoshizutsu okyakusama-ni*
do:TARI   such as   steady   effort-NOM   little by little   customer-DAT

tsutawatteiru.
be.introduced:STA

'Little by little, steady efforts are introduced (to help customers), such as support the area within 48-hour and free replacement of the batteries.'

Interestingly, this strategy occurs less frequently in combination with *nado*. This figure is noteworthy because in our corpus *nado* is the exemplifying construction with the higher frequency of lexicalized categories and *toitta* is one of the most used connectors. Nevertheless, they occur together only twice.

(3.59) Example₁ YA Example₂ NADO TOITTA Label

Baisoku-ya   *daburuchūnā nado*   *to itta*   *jōkyū*   *kurasu no*   *kinō*
double.speed-YA   double-turner NADO   such as   high   level   LK   function

wa   *tōsaisareteinai.*
TOP   equip:PASS:STA:NEG

'High level functions such as double-speed and double-turner have not been mounted.'

Finally, contrary to *no* and *to iu*, *to itta* can be attached directly to the examples like in (3.57). This is likely due to the dedicated nature of the fixed expression *to itta* as a connector between category labels and examples.

3.4.4 General extenders (*nado, toka*)

Some general extenders can be directly attached to category labels, functioning as linguistic connectors without any other strategy. More specifically, in our corpus, both *nado* and *toka* perform this function, as shown in the examples below.

(3.60) Example NADO Label

Imēji-gata	*supamu-no*	*hassei*	*nado*	*supamu*	*haishin*	*gijutsu*
Image-model	spam-GEN	occurrence	NADO	spam	distribution	technique

"Spam delivery technologies such as (the occurrence of) image type spam".

(3.61) Example TOKA Label

Yahoo toka	yūmeina	kaisha-ga	ippai	aru
Yahoo TOKA	famous	company-NOM	a lot of	exist

"There are a lot of famous companies such as Yahoo".

The usage of *nado* as a connector is particularly widespread (20%), to the point that in 5 occurrences, it performs this function even though the examples are connected by *toka*, like in the following example:

(3.62) Example$_1$ TOKA Example$_2$ NADO Label

Kontentsu	*wa*	*sarani*	*"sakkā"*	*toka*	*"kujira"*	*nado*	*8-sen no*
Content	TOP	further	soccer	TOKA	whale	NADO	8-thousand-NML

kategorī-ni	*saibunsareteiru.*
category-DAT	divide.do:PASS:STA

'Contents are divided into 8 thousand categories such as "soccer" and "whale".'

This functional extension is likely due to the closeness of general extenders and category labels because of the order of constituents in Japanese, according to which the label must be added at the end of the entire list construction, hence (immediately) after the general extender. This may also explain why *tari* is the only general extender that is not used as a connector in our corpus. Since *tari* requires the addition of the auxiliary verb *suru* "to do" at the end of the list of examples[41], there can be no direct connection with the category label.

---

[41] It should be noticed that in some cases, the auxiliary verb *suru* is indeed elided, especially in speech. Nevertheless, the nature of *tari* as a non-finite verbal form may still be the structural constraint that prevents (or al least restricts) its usage as a connector.

Diachronically, it may be also motivated by the probable omission of the nominalizer *no* in some occurrences:

Example NADO NO Label   >   Example NADO Label
Example TOKA NO Label   >   Example TOKA Label

This type should be investigated further in languages with the same order of constituents as Japanese, to understand if the functional extension from general extender to linguistic connector can be a cross-linguistic pattern or if it is a language-specific strategy, merely motivated by the elision of the nominalizer *no*.

3.4.5 Combinations of connectors

In one occurrence of our corpus, the similative marker *youna* is combined with the noun complement marker *to iu*:

(3.63) Example$_1$ TARI Example$_2$ TARI TOIU YOUNA Label

Gakkyoku-o	*sentakushitari,*	*saisei-o*	*okonattari*	*toiu youna*
tune-ACC	selection:do:TARI	reproduction-ACC	perform:TARI	TO IU like

sōsa	*wa*	*okonau-koto-ni*	*naru.*
operation	TOP	perform-NML-DAT	become

"Operation will be carried out for example selecting songs and performing the reproduction.'

The combination of connectors is neither particularly odd, nor peculiar of Japanese. For instance, regarding Italian, Barotto and Mauri (2016) note the frequency of similative markers co-occurring with exemplifying markers that can link category labels and examples (e.g., *per esempio* "for example"), like in the sentence *luoghi più disparati, ad esempio come stazione di ricerca, location per eventi e molto altro ancora* "a variety of locations, for example such as research station, event venues and much more".

3.5 (RE)CATEGORIZATION: THE DIVISION OF LABOUR BETWEEN LABELS AND EXAMPLES

What has emerged through this chapter is that labels and examples play both an active role in the building and communication of categories. Moreover, whenever they are used in the same utterance, they cooperate to better direct the reference to the target category. In short, there is a division of the labour between labels and examples. This fact is backed up by the

linguistic data provided in section 3.2 which show that in most cases these two cores are tightly joined by means of linguistic connectors. In this final sections, we would like to highlight the different functional tasks that labels and examples may perform in order to improve the reference to the target category, easing the inferential process of the hearer.

3.5.1 THE ROLE OF THE LABEL IN DIRECTING THE INFERENTIAL PROCESS

Generally speaking, the main role of labels is to make explicit (and therefore clear) what the members of the target category have in common, in order to direct and also facilitate the inferential process towards the identification of the shared property. In this sense, the very presence of a category label acts as a sort of guarantee of the fact that the hearer is really called upon to build a category that exhibits certain features. Therefore, the hearer does not need to infer the presence of a category and its defining property by comparing the mentioned examples. Consider the following example:

(3.64) *Saikin   PC-ga   furīzushitari   gamen-ga   utsuranaku   nattari*
Recently   PC-NOM   freeze:do:TARI screen-NOM   project:NEG:ADV become:TARI
toitta   shōjō-ga   deta   tame,   netto   kensakushite
such as   symptom-NOM   appear:PAST   because   net   search:DO:GRD
kaiketsusaku-o   mitsukete   okonaimashita.
solution-ACC   find:GRD   perform:POL:PAST
'Recently, because issues have appeared, such as the PC got freeze and the screen did not project (images) no longer, I looked up for a solution through an internet research.'
Label: *shōjō* "symptoms"

The label *shōjō* literally designates medical symptoms, but here it is used to refer to computer issues. Despite needing a further contextualization (as is usual with simple labels), when the label is properly interpreted, it provided useful insights on the type of category the hearer must infer and build: computer issues that may occur and signal the presence of a larger problem that should be identified and solved. Therefore, even without the mention of concrete examples, the hearer would be able to infer not only the presence of a wider category, but also to sketch it out at least partially.

In addition, the importance of the label in directing the inferential process can be further understood if we consider that the same list of examples can be interpreted differently depending on the context. Consider another example:

(3.65) *Kō-gan-zaichiryō-de*   *wa*   *saketetōrenai fukusayō.*

anti-cancer-drug.therapy-LOC   TOP   unavoidable   side effect

Shokuyoku-ga   *ochitari,*   *hakkekkyū-no kazu-ga*   *sukunaku*

appetite-NOM   fall.down:TARI leucocyte-GEN number-NOM few:ADV

nattari,   *soshite*   *kaminoke-ga* *nuketari.*

become:TARI and   hair-NOM   fall:TARI

'Sides effects that cannot be avoided during an anti-cancer drug therapy. Losing appetite,
leucocytes become less, and hair falls down.'

Label: *Kōganzaichiryō de wa saketetōrenai fukusayō* "sides effects that cannot be avoided
during an anti-cancer drug therapy"

In the article above the author refers to cancer and potential treatments to cure it. If we consider only the mentioned examples (e.g., *shokuyoku ga ochitari, hakkekkyū no kazu ga sukunaku nattari, soshite kaminoke ga nuketari* "losing appetite, leucocytes become less, and hair falls down"), there is a risk of wrongly interpreting them as the (less strong) symptoms of the disease, instead of the side effects of the treatment. However, the label (i.e., *kōganzaichiryō de wa saketetōrenai fukusayō* "sides effects that cannot be avoided during an anti-cancer drug therapy") makes explicit the common property of the examples in this specific context and directs the reader to interpret them correctly as side effects, thus avoiding any possible misunderstanding.

Compared to the label used in (3.64), the label in (3.65) provides a higher degree of specification by pointing out some additional details. As previously noted (cf. section 3.2.1), by means of a complex label, the speaker can designate more accurately the target category (in this particular case, side effects of anticancer therapy that cannot be avoided). It follows that, in some cases, the reference of the label can be so precise, as to help the contextualization of the category along with the example(s).

On this final point, it is important to note that just as the choice between lexicalizing or not lexicalizing a category is an arbitrary communicative strategy, the same applies also to the choice of using a simple label instead of a complex label and vice versa. All categories can be designated by at least one simple label[42] or by one or more complex labels (depending on how the speaker decides to encode all the specification of the target category).

Therefore, depending on the situational context, the target category and the discursive goal, the speaker can choose the strategy considered more opportune:

---

[42] As we saw in the previous sections, especially with regards to frame-based categories, the speaker may decide to highlight through the label how the examples should be conceived.

- In case the speaker chooses to use a complex label, he or she can provide a higher degree of specification and a more precise reference to the target category even without the mediation of examples, thus helping further to direct the inferential process. Nevertheless, lacking a stable association between a mental representation of the category and a fixed linguistic expression to designate the category, the usage of such precise and specific types of labels may trigger a greater cognitive effort not only for the speaker who is required to identify and lexicalize the exact common property, but also for the hearer who is called upon to retrieve the actual members of category.

- In case the speaker chooses to use a simple label, there is a minor cognitive effort both for the speaker and the hearer as the common property is outlined in broad terms. Nevertheless, in this case, the contextualization of the label in order to identify the actual members of the target category is less straightforward.

Therefore, considering what we have outlined so far, we argue that category labels can be encompassed[43] in the notion of *category clues* (cf. Barotto and Mauri 2016), that is, linguistic elements which provide semantic hints towards the defining property of the context-relevant category. More specifically, category labels make explicit reference to the lowest common denominator of the examples in a specific situational context, thus helping to avoid any misunderstandings in the construction of the category.

3.5.1.1 THE SEMANTIC VALUE OF PLACEHOLDER LABELS

Up to this point we analysed the value of labels in directing the inferential process towards the identification of the relevant property of the target category, which can be considered the core function of proper category labels. Nevertheless, we may identify other constructions that seem to function as label, but should not be considered as such, since they do not express the defining property of the category (cf. section 2.2.1.1). These other constructions may be used as proper labels to make explicit some semantic features concerning one category and its members. In particular, we refer to two types of situations. In the first, the speaker uses a very generic noun not to make direct reference to the category, but rather to express certain semantic properties that relate to the examples or category

---

[43] This is true with the obvious exclusion of the semantic functions, which are, however, performed by constructions that are not proper labels.

itself (i.e., non-exhaustivity, highlighting the status of example of the items, or the semantic relations among the examples). In the second, some adjectives are added to proper labels to highlight the non-exhaustivity feature (that is, to indicate that there are other potential members) but also to suggest on which level of the vertical dimension the category should be built. Let us consider in detail both patterns.

In section 2.3.1.1, while providing a working definition of the notion of category label, we propose the exclusion of words like *koto* "thing", *mono* "thing", *toki* "time", *kēsu* "case" whenever used as labels without further specifications, as these words do not provide any significant specific reference to a class of items. Most importantly, they do not specify the defining property of the category. For this reason, we decided to call this type of labels as "placeholder labels". However, they do not lack of any semantic value. Consider the following examples:

(3.66) *Keiei   rinen-ya   kigyō   bijon   toitta   **mono** wa   hitsuyō*
management idea-YA   business   vision such as   thing   TOP   necessary
desu   ka.
COP   Q
'Do we need things such as management philosophy and corporate vision?'

(3.67) *Kono   dantai   wa   ryokō-ya   kenkōkanri   toitta   **koto***
this   organization   TOP   travel-YA   health care   such as   thing
nikansuru   jōhō-o   haifusuru.
related to   news-ACC   distribution:do
'This organization provides information on things such as travel and health care.'

While *mono* (3.66) and *koto* (3.67) hardly add any substantial information about the shared property of the category members, they emphasize the non-exhaustive interpretation of the list. In fact, it seems that the real purpose of these types of constructions is to indicate that the mentioned items should be considered only as a part of a larger group of elements (the generic *mono* or *koto*, "things"), thus highlighting the presence of a wider set or category.

Because in Japanese it is possible to codify non-exhaustivity simply by means of dedicated connectives (e.g., *ya, tari, toka*), the usage of these empty labels does not seem to be particularly important (although it is important to note that placeholder labels may still be found even in a not so wide corpus like ours). However, it might become pivotal in those languages that do not have dedicated non-exhaustive connectives, such as English.

(3.68) *However, be prepared to see **things like** mental instability, depression, pornography, drug abuse and aggression.* (LCC)

In (3.68), "things like" does not offer any sort of specifications on the type of set. It does not make clear the defining property of the category, which should be inferred solely by comparing the mentioned members. Nevertheless, it acts as a clue to indicate the actual presence of a larger set of elements, that is, a category, beyond those explicitly mentioned. Therefore, semantically, it highlights the non-exhaustivity of the list.

Another way in which the speaker may provide further semantic specification through placeholder labels concerns the use of nouns such as *toki* "times", *kēsu* "cases", *baai* "situations". As already explained in 2.2.1.1, we decided to exclude this type of words whenever used as labels without further specifications, because instead of making clear the common property of the category, it seems that their scope is to indicate semantic relations among the mentioned elements. Consider the following:

(3.69) *dōshō*   *niyoruto,*   *kaku*   *jichitai-ga*   *hokenshō-no*
       minister   according to   each   local.governament-NOM   insurance.card-GEN
       hassō-o   *hajimeta*   *3 tsuki-chūjun ikō,*   *tenkyo-ya*   *hon'nin fuzai-de*
       sending-ACC   begin:PAST   3 mid-month   since   moving-YA   person absent-STR
       modottekitari,   *ayamatte*   *sonomama*   *suterarete*
       come.back:TARI   by mistake   without change   throw.away:PASS:GRD
       shimattari   *shita*   **kēsu**-*no*   *hōkoku-ga*   *aitsuida.*
       finish:TARI   do:PAST   **case**-GEN   report-NOM   follow:PAST
       "According to the ministry, after mid of March, each municipality began the shipping of insurance cards, there were numerous reports of cases in which people tossed it away by mistake or the cards came back because of the absence of the person or because the person moved away."

In the sentence above, what *kēsu* "cases" highlights is not a specific feature of the category, but rather the type of semantic relation that occurs among the examples. In other words, in (3.68) the author wants to present a range of options resulting from a specific issue (i.e., the shipping of insurance cards). Moreover, she wants to commit to all these options as being the case, although each happens in a separate situation (cf. the notion of *separative conjunction*, Mauri and Ariel, 2016). In this sense, the role of the placeholder label "cases" is to direct towards this interpretation: the examples are all cases, that is, alternatives that

may happen in different situations. It does not say anything about what kind of cases they actually are but it does say how they relate to each other.

Let consider another example.

(3.70) *Hontai-o aita*   *mama*   *hōchishite*   *oku*   *to,*   *gamen-ga*
body-ACC open:PAST   as it is   leave:do:GRD   do   when   screen-NOM
chiratsuku-koto-ga atsutari, *GPS*   *taiō*   *i-apuri-de*
flicker-NML-NOM   exist: GPS   interaction   i-application-STR
genzaichi-o   *chekkusuru*   *to,*   *"ima iru basho-no*
current.position-ACC   check:suru   when   now exist place-GEN
kakunin-ni   *shippaishimashita" to sokui*   *dekinai baai-ga*   *aru.*
confermation-DAT   failure:do:POL:PAST QT positioning   POT:NEG situation-NOM   exist
'There are cases in which, when you leave the body open and the screen is flickering or when you check the current location with the GPS app and it cannot measure the position "Failed to check where you are now".'

Here the author describes some issues encountered by users while using a particular mobile phone model, later withdrawn from the market. Again, *baai* "situations" should not be considered as a proper label, but as a linguistic tool to express the semantic relation that links the listed events: despite being all problems that actually have occurred, it is not implied that they have happened at the same time, but more likely in separate situations.

A structural element that seems to further validate the idea that these words should not be interpreted as proper labels but as "dummy elements", is the fact that they are often not directly linked to the examples by means of connectors or similative constructions (i.e., X such as Y or X like Y). Instead, they work as heads of relative clauses (e.g., *dekinai baai* "situations in which it is impossible", *shimattari shita kēsu* "cases in which for example it ended up being...").

Finally, in addition to placeholder labels such as *mono* or *koto*, another strategy to further stress the non-exhaustivity feature regards the usage of quantity adjectives. These adjectives do not add any concrete specification regarding the category and its defining property, rather they emphasize the wider nature of the category. In our corpus, the most commonly attested are *iroirona* "various" and *samazamana* "various".

(3.71) *Jinkenmondai,*   *kankyōmondai,*   *kosodate*   *shien*   *nado*

human rights issue   environmental issues   child care   support   NADO

samazamana   *tēma-de*   *sekkyokuteki-ni*   *torikumu*   *kigyō*

various   theme-LOC   active-DAT   make effort   enterprise

'An enterprise that works actively in a variety of themes such as human rights, environmental issues and child care support.'

(3.72) *Yūzā-ga*   *Blockbuster-ya eBay nado,* **iroirona**   *kyōryoku*   *saito-o*

user-NOM   Blockbuster-YA eBay NADO   various   collaboration   website-ACC

burauzusuru   *to,*   *sono*   *kōdō*   *kiroku-ga*   *Facebook-ni okurikaesarete,*

browse:do   when   that   action   report-NOM   Facebook-DAT send.back:PASS:GRD

yūjin-ga   *kyōyū*   *dekiru*   *yōninaru.*

friend-NOM   share   do:POT   reach the point

'When the user browses various partner websites such as Blockbuster or eBay, the action report is sent back to Facebook and friends will be able to share it.'

In section 3.2.1, we have seen how adjectives can be added to simple labels to provide a higher degree of contextualization. However, in (3.71), *samazama* does not offer any specific contribution to the inference of the defining property of the category, that is, in this particular case, what kind of issues are taken care by the company in question. Therefore, in order to build the category, the reader can rely just on the simple label *tēma* "themes" and on the examples provided which contextualize effectively the type of topics (i.e., social issues in which a company chooses to get involved by means of donations). The role of *samazama* is to highlight that the covered issues are many beyond those mentioned, thus reinforcing the non-exhaustive feature. In a similar way, in (3.72), the role of *iroirona* "various" is not further specify what types of websites are considered partners, but rather to emphasize their high number.

Nevertheless, emphasizing the non-exhaustivity feature is not the only function performed by this type of adjectives. Indeed, semantically, adjectives like *iroirona* and *samazama* not only provide information about the number of the items (that is, they are not simply synonymous with "many"), but they also highlight the strong heterogeneity of the set. In (3.71), the social issues are many, but also quite different to each other. Examples support this interpretation: the writer chose to mention issues concerning different fields such as *kankyōmondai* "environmental issues" e *kosodate shien* "child care support", whose only common denominator is to be indeed social issues that may interest private companies in terms of charity. Similarly, in (3.72), the websites taken into account are many and different

in nature: eBay and Blockbuster represents two different services, although both fall within the scope of the category.

Although *iroirona* and *samazamana* are probably the most interesting instances, there are also other adjectives that perform similar functions. For example,

(3.73) *Tōshin-de   wa,   tetsugaku-ya shinri,   keizai,   hōgaku   nado*
report-LOC   TOP philosophy-YA   psychology   economics   law   NADO

*no   **hiroi** bunya-de,   ningen   no   kokoro-no   ugoki-o   umidasu*
NML   **wide** field-LOC   human   NML   mind-GEN   movement-ACC produce

nō-no   fukai   chishiki-ga   motomerareteiru   to   shiteki.
brain-GEN   deep   knowledge-NOM   demand:PASS:STA   QT   pointing out

'In the report, it was pointed out that in a wide range of fields such as law, economics, psychology, philosophy, a deep knowledge of the brain that produces the movements of the human mind has been demanded.'

In the sentence above, *hiroi* "wide" is used in a similar way to *iroirona* and *samazamana*, that is, to indicate that the relevant fields are many and various.

The point regarding the heterogeneity of the category is of particular importance, especially giving the fact that in our corpus there are 18 occurrences of adjectives used to highlight the heterogeneity (and the non-exhaustivity) of the category, but there are no occurrences of quantity adjectives used only to highlight the non-exhaustivity (e.g., *ooi*, "many") of the set.

To understand the importance of this function, we should note that by their very nature, categories are heterogeneous sets of elements grouped together just because they share a defining property. This inherent heterogeneity of categories has been examined by different approach to categorization. Despite taking into account only common taxonomical categories, Rosch (1978) recognizes that the degree of heterogeneity is not uniform across categories, but it varies along the vertical dimension: members of superordinate categories (e.g., furniture, mammals) are much more heterogeneous than the members of categories below the basic level (e.g., dogs). This is due to the fact that superordinate categories' members have fewer common attributes than basic level categories' members (cf. Rosch et ali. 1976).

The heterogeneity feature appears to be even more pivotal when it comes to ad hoc categories. Being constructed spontaneously to achieve a goal relevant in the current situation rather than to represent states of the environment (Barsalou 1983, 2010), ad hoc

categories often encompass members that otherwise have little in common. For examples, items as different as a slipper, a newspaper, your foot, and a can of bug spray all fit within the category things to use to kill a roach (Overstreet 1999:42).

Interestingly, at the linguistic level, we may note that some linguistic strategies that codify categories focuses precisely on the heterogeneity feature, such as associative plurals (Mauri 2016). Associative plural constructions are identified by two semantic properties: referential heterogeneity and reference to groups (Daniel and Moravcsik 2005). As for the latter, it indicates that associative plurals denote sets with a clear internal cohesion. However, this cohesion is not synonymous of homogeneity. Actually, the referential heterogeneity feature distinguishes between additive plurals and associative plurals: while the former denotes a homogeneous set, i.e. "cats" denotes a set in which every member is a cat, the latter denotes a heterogeneous set. Consequently, the cohesion is provided by the defining feature which should be interpreted according to the context.

Let consider again example (3.73), which was taken from a scientific paper about research on the working of the brain. This type of context might influence the reader in the construction of the target category, directing the inferential process towards something like "fields which require a deeper knowledge of the human brain". In other words, the reader may think that we are dealing only with scientific fields, such as medical science. In this sense, the adjective *hiroi* "wide" helps to broad the reference: it indicates that a larger variety of fields should be taken into account. Examples support and further emphasize this interpretation by mentioning very different fields (*tetsugaku* "philosophy", *shinri* "psychology", *keizai* "economy", *hōgaku* "law"). Therefore, we might say that *hiroi* allows the reader to make reference to a superordinate category which lacks a specific name, rather than a more specific sub-category that may come up in the reader's mind at the first impact. So, in other words, adjectives like *hiroi* work on the vertical dimension of categories: by increasing explicitly the heterogeneity of the category, they allow to move upward in the vertical dimension, to include elements that at first glance could have been be excluded. This means that, without any need for creating new labels each time and by simply using linguistic constructions that increase the perceived heterogeneity of the mentioned category, it is possible to make reference to superordinate categories. In fact, the use of these adjectives further confirms Rosch's insights on the role of heterogeneity in the vertical dimension of categories, and it proves how language has an active role in cognitive categorization mechanisms.

3.5.2 THE ROLE OF EXAMPLES IN CONTEXTUALIZING AND ACTUALIZING THE CATEGORY

Examples represents concrete members of the category. Under this respect, their main function is to bridge the gap between the concreteness of the hyper-specific context and the abstract configurations of the labels. To do this, they contextualize and actualize the category by providing concrete members chosen according to the context. In fact, as we noted in the previous section about the semantic properties of examples (cf. section 3.4), they are not chosen randomly but to better represent concrete experiences and situations (i.e., the preference for concrete entities) and to be easily processed by the hearer (i.e., the preference for examples encoded by noun phrases).

More specifically, examples perform two important functions 1) contextualizing and 2) actualizing the category. These functions bear a different weight according to the type of label with which they occur, since they cover any lack in the reference designated by the label: on one hand, simple labels might be too general and abstract, on the other, complex labels might be too specific and/or ambiguous. Let us see these functions in detail.

The first important function is to contextualize the category. This is particularly important with simple labels, as the examples must fill the discrepancy that exists between what is designated by the label and the target category. In other words, in these cases, it is necessary to link the broad abstract notion provided by the label to the actual category relevant in the specific context. Consider the following example:

(3.74) *Shisutemu   kanri-no   hanzatsusa-ya,   sekyuriti-no   kyōka,*
system   management-GEN   complexity-YA   security-GEN   strengthening

kanrikosuto-no   sakugen   nado   no   mondai-wo   kaiketsusuru
administrative.cost-GEN   reduction   NADO   NML   problem-ACC solution:do

soryūshon   desu.
solution   COP

'(It) is a solution to solve problems such as the reduction of administrative costs, the reinforce of the security and the complexity of the management system.

While it is not particularly difficult to understand – generally speaking – what *mondai* "problems" stands for (that is, to retrieve the members of the category designated by the label), understanding what types of problems are relevant in this specific context requires a certain amount of cognitive effort, which is eased by the list of examples. In fact, the mentioning of concrete exemplars facilitates the process of narrowing down the abstract

category designated by the label and tailoring it to the specific context: they signal that we are not dealing with a general abstract category *mondai* "problems", but with a more context-specific category, that is "administrative problems that affect business companies". Therefore, on one side, the label facilities the inference referring to a more general (and therefore more accessible) category, on the other, the examples anchor it to the context, shaping the category in a way that is relevant to the context.

While this function has more weight when combined with a simple general label, it still may be useful even with more specific complex labels, as shown in the example below.

(3.75) *Kigyō   wa seitōna   riyū-ga   areba,   naitei-o   torikesu-koto-ga*
company TOP legitimate   reason-NOM   exist:COND   offer-ACC cancel-NML-NOM

dekiru.   Tatoeba,   naitei-sha-ga   naitei-go-ni   hanzai-o
POT   for example offer-person-NOM   offer-after-LOC   crime-ACC

okashitari,   gakureki-o   sashōshiteitari   shita
commit:TARI   academic.background-ACC   false.statement:do:STA:TARI do:PAST

baai   da.
case   COP

'If there are legitimate reasons, companies can cancel the job offer. For example, in case the nominee commits a crime after the nomination or in case he made a false statement about his academic background.'
Label: *seitōna riyū* "legitimate reasons" (to dismiss job candidates)

Here the author refers to situations in which people got fired before even starting a job without real motivation. He further explains that there are some legitimate reasons for firing someone after the job has been offered and then provides a list of concrete examples. Again, while the label *seitōna riyū* "legitimate reasons" focuses only on the presence of justified reason to fire someone, the examples describe real circumstances in which people get rightfully fired, thus contextualizing the reference towards the target category relevant in that specific context. This is in part due to the fact that lacking a stable association between linguistic labels and context-relevant conceptual categories, speakers may choose to focus on some features of the category deemed as more important for their communicative goal, excluding others which may have been more useful in the contextualization of the category.

The second function performed by examples is to actualize the reference, that is, to shift the focus from abstract to concrete experience to facilitate the processes of elaboration and comprehension of the category. It is noteworthy that this is also the main communicative

function traditionally ascribed to exemplification, as noted in section 1.2, for examples in communication studies (see Zillmann 2002) or in research on discourse coherence (see Hobbs 1985).

As we have seen in the previous section, this function is facilitated by the fact that speakers tend to choose examples that refer to concrete objects, with which the interlocutors may have had (direct or indirect) experience in their everyday life. At the cognitive level, this is particularly valuable whenever a complex label is used: again, since there is no permanent representation of the category, the connection between the category and the label chosen by the speaker is absolutely arbitrary, and therefore potentially opaque or unclear. In this sense, providing some concrete examples resolves any potential ambiguity.

(3.76) *Imēji-gata   supamu-no   hassei   nado   supamu   haishin*
Image-model spam-GEN   occurrence   NADO   spam   distribution
gijutsu-ga   kōmyō   kashiteoru.
technique-NOM   ingenious   change.into:STA
"Spam delivery techniques, such as (the occurrences of) image-type spam have grown more sophisticated"

Even if the reference to the category "spam delivery techniques" by means of the complex label can be considered quite satisfactory, especially in an article that describes specifically the spam issue, it remains an abstract, albeit very specific, construction, which requires a certain degree of encyclopaedia knowledge regarding the subject to be processed successfully. On the contrary, the mentioned example, i.e., "occurrences of image-type spam" makes reference to a frequent concrete situation related to the everyday usage of personal computer and internet, that is, the possibility of coming across to web images that serve no purpose other than to spam. As a result, the fact that examples represent potential direct or indirect experiences ultimately makes the reference to the category easier to process helping the elaboration process. Another example was provided in (3.15), repeated here as (3.77).

(3.77) *Sōsharunettowākingusaito wa   fisshingu-ya   onrain sagi   toitta   sagi*
social.network.website   TOP   phishing-YA   online fraud   such as   fraud
kōgeki-no   kakkōno   hyōteki   tonatteimasu.
attack-GEN   easy   target   become:STA:POL

'Social networks have become easy targets of fraud attacks such as phishing and online fraud.'

Label: *sagi kōgeki* "fraud attacks"

Here too the examples help to actualize the set of issues encompassed under the label *sagi kōgeki* "fraud attacks", in addition to favour the contextual interpretation: the types of fraud that can be found online are different from those encountered in real-life. In this sense, *fisshingu* "phishing" represents a concrete situation: the reader might have had a direct or indirect experience of it, and this accessibility facilitates the elaboration of the ultimate concept of online fraud attack.

In other cases, the speaker might stress some features of the category over other more fitting for communicative purposes, with the result that the label may be perceived as misunderstandable or insufficient, and thus needing further specification. Consider the example in (3.12), repeated here as (3.78).

(3.78) *Doraggu&doroppusuru*   *dake*   *de*   *shashin-o*   *appurōdoshitari,*
drag&drop:do   only   STR   photo-ACC   upload:do:TARI
daburukurikku-de   *suraidoshō-o*   *saiseisuru*   *nado, shoshinsha* *demo*
double click-STR   slideshow-ACC   play:do   NADO beginner   GDR:also
kaitekini   *riyōdekiru*   *kantanna*   *sōsa.*
simply   use:POT   simple:AGG   operation
'Simple operations that can be used comfortably even by beginners', such as playing a slideshow with a double-click, uploading photos by simply dragging and dropping and so on.'

The label "simple operations that can be used comfortably even by beginners" is certainly functional to describe the ease of usage of the software in question, however, it does not identify unequivocally a particular set of functions. On the contrary, the examples describe specific situations, that is, actions that can be easily elaborated at the cognitive level, because they are part of the everyday usage of a photo software. The combination of these two elements allows not only to designate correctly the target category, but also to emphasize the fundamental property of being easy to use.

Interestingly, the actualizing function is possible even in those cases where the category designated by the label and the target category coincide:

(3.79) *Femininna   pinku wa,   orenji-ya   ierō nado no   bibiddo karā-ni   wa*
       Feminine   pink   TOP   orange-YA   yellow NADO NML   bright colour-DAT   TOP
       matchisuru   bannō   karā.
       match:do   all-purpose colour
       'Feminine pink is an all-purpose colour that match well with bright colours such as orange
       and yellow.'

Here, the category designated by the label *bibiddo karā* "bright colours" coincides with the target category the author wants to communicate. In this case, the questions around the role of the examples may follow the one provided by Overstreet (1999) about general extenders used to make reference to lexicalized categories: why would a speaker provide examples to make reference to a lexicalized category, when she could refer to it by its name? Overstreet (1999: 44) argues that there are pragmatic reasons for this, mainly referring to possible misunderstanding between speaker and hearer about the reference designated by the label (e.g., the speaker thinks that the hearer does not know the label). While we agree with Overstreet's analysis, there may be other reasons concerning the relationship between labels and examples. Specifically, these other reasons bring us back to the core functions of exemplification that we delineated in chapter 1. Regarding the vertical dimension of categories, Rosch (1978) identifies the "basic level" as the most culturally salient and, for this very reason, also the more accessible. Therefore, for example, while "colours" is a super-ordinate category, on the basic level we find basic colour categories (see Rosch 1973) such as "blue", "yellow", "orange". We may thus consider "bright colours" as an intermediate step: a super-ordinate category that encompasses only a sub-set of the wider category "colours". It follows that because of this position in the vertical dimension, "bright colours" is more abstract and less salient than "yellow" and "orange". Moreover, even if we are dealing with colours, we might still say that "yellow" and "orange" are more concrete than "bright colours", in the sense that they are part of the personal concrete and sensorial experience of everyone (cf. section 1.2.1.2), while "bright colours" requires a further abstraction process. This inherent accessibility of the basic level is what makes it a perfect source of examples. Examples picked from the basic level allow to make the abstract more concrete, thus shifting the comprehension of the category to a more accessible dimension.

Let consider again example (3.79). The sentence is part of a fashion article that explains which colours and styles match bright pink. Providing concrete examples of the category "bright colours" allows the writer to make reference more precise and more comprehensible. More specifically, the mention of specific colours (i.e., "yellow", "orange") can make the

reader think about specific clothing items of those colours that she owns, thus ultimately facilitating the understanding of the category itself. In this case, rather than making the label more understandable, we argue that examples allow to make the label more concrete and more easily linked to things the hearer/reader has direct experience with in her everyday life, ultimately making the reference to the category easier to process.

Summary

In this chapter, we have seen how labels and examples actively work to construct on-line contextually relevant categories. Not only both of them provide a specific contribution to the inferential process, but we have also seen how they closely collaborate by covering any lacks in the reference provided by other element. So, while the label allows to make direct reference to the property that defines the category, and thus to the category itself, it remains anchor to an abstract dimension which makes it potentially hard to be interpreted in a specific context. On the contrary, examples allow contextualization and actualization of the category, but in the absence of a label, the shared property should be identified by means of a comparison not only among the examples, but also between examples and the context, requiring an extra cognitive effort, which may be relieved by an explicit label.

Having said that, the combination of labels and examples is not an essential element in the construction and communication of categories. Just like labels can function alone (with the help of the context to avoids potential ambiguities in reference), also the examples can make reference to categories created on the spot without the assistance of a label. While the first case is behind the scope of our study (although we believe that it might provide useful insights in the relationship between categorization and language), the second case will be the focus of our next chapter.

4. EXEMPLIFICATION OF NON-LEXICALIZED CATEGORIES

4.1 THE NOTION OF NON-LEXICALIZED CATEGORY

In the previous chapter, we have seen how category labels represent an advantage for the hearer, because they facilitate the inferential process, suggesting the property shared by the category members. Nevertheless, labels are not necessary in the categorization process. In other words, the speaker may refer to a category even without an explicit category label. Indeed, the simple mention of examples is enough to trigger the inferential process. Consider the following instance:

(4.1)   *POS   shisutemu-de-no   kādo   kessai   torihiki   nioite,   jiki*

         POS   system-STR-LK   card   payment   transaction   regarding   magnetic

         kādo   dēta-ya PIN, sekyuritikōdo-nado-o   POS-jō-ni   nokosanai.

         card   data-YA PIN   security.code-NADO-ACC   POS-up-LOC leave:IMP:NEG

         'As for the card payment transactions in the POS system, do not leave the magnetic card data, the PIN and the security code and so on the POS.'

In (4.1), we observe an instance of non-lexicalized category, that is, a category encoded only by examples without any explicit label. Here, the author refers to security of card payment transactions. Thus, to create and communicate the category of those parts of a card payment transaction that contain sensitive information, she does not provide a label, but she lists some concrete members as representative examples of the category. The status of example is signalled by means of the non-exhaustive connective *ya* and the general extender *nado*, both emphasising the presence of other potential members beyond those mentioned.

As already mentioned (cf. section 3.1), the lexicalization of a category (that is, the deliberate act of providing a category label) does not discriminate between types of categories, but it is rather an arbitrary communicative strategy to designate categories in specific contexts. This holds the other way around: not lexicalizing a category is a communicative strategy as well. Even though there are good reasons for providing a label (cf. section 3.5), in some contexts the lexicalization may require a greater cognitive effort for the speaker (e.g., frame-based category or category of examples expressed by verbal phrases). First, he or she must identify a suitable label. To do that, it is essential to clearly

identify the defining property of the category and then to encode it by means of lexical items. Then, the speaker chooses whether to use a simple label (whose reference is generally broad and unspecific), or to provide further specification. Therefore, in some cases, the lexicalization process requires to create *ad hoc* labels, adding further linguistic material (cf. section 3.2.2). Depending on the context and the type of category, this process may demand cognitive effort. Finally, *ad hoc* labels may also end up being too opaque, and thus not easily comprehensible for the hearer. For all these reasons, in some cases, the sole use of examples may be the more effective solution.

Nevertheless, while the label functions as a guarantee for the reference to a category, the lack of it presents us with a further major problem, namely the necessity to demonstrate that the speaker is actually referring to a category. In other words, we still need to consider the possibility that categorization may not be the discursive goal of the speaker, that is, that in certain situations, exemplifying strategies simply encode enumeration of items. For instance, consider the following example.

(4.2) *Tēburu-no*   *ue-ni*   *hon-ya*   *koppu-ya*   *pasokon-nado-ga*   *arimasu.*
       table-DET   up-LOC   book-YA   glass-YA   pc-NADO-NOM   exist:POL
       'On the table there are books, glasses, a personal computer and so on.'

Sentences like (4.2) are typically presented in grammars to explain the functioning of the non-exhaustive connective *ya*. However, the discursive function of non-exhaustive lists like that in (4.2) may be a matter of dispute: is it a case of categorization or just a case of simple enumeration? At first sight, the second option seems more likely. The issue is not trivial: labelling each occurrence as an instance of categorization without further discussion would end up watering down the very notion of category, which, on the contrary, exhibits strong defining characteristics and constraints, such as an internal graded structure and a defining property.

However, the task is not an easy one. First, it is crucial to consider an important terminological distinction between the notion of list and the notion of enumeration. A list is a linguistic construction (cf. Fillmore and Kay 1995) which can be defined as the "junction of two or more elements occupying the same structural position in a dependency structure" (Blanche-Benveniste et ali. 1990, Gerdes and Kahane 2009). Under this respect, we may thus propose that list constructions perform different functions, for example categorization (e.g., (4.1)) or enumeration (e.g., (4.2)). At this point, we also need working definitions for

both these functions, in order to identify some features that may help us to distinguish between instances of categorization and instances of enumeration.

Generally speaking, we define enumeration as the ordered listing of the items in a set. So, at least in theory, pure enumeration, unlike categorization, does not presuppose any kind of constraint on how the set should be organized and what is the relationship among the members of the set. Compare (4.2) with (4.3), which is a sentence from my corpus.

(4.3) *Tsūringu-ya   jitenshatsūkin-nado-ni   habahiro-ku   tsukaeru   rōdobaiku.*
        touring-YA   bicycle.commuting-NADO-LOC wide-ADV   use:POT   road.bike.
        'Road bikes that can be widely used for commuting, touring and so on.'

In both occurrences, the non-exhaustive connective *ya* is used to encode non-exhaustive lists. Yet, the discursive goals are different.

In (4.2), the speaker lists the items on the table, leaving the list open to indicate that there are others which are not explicitly expressed. What the mentioned items have in common is not a specific property, but the fact that they are indeed part of the list. In other words, they are all items that are actually on the table in that particular moment. For this reason, it may be difficult for the hearer to identify with certainty other potential items. Of course, it is possible to make a very rough selection at most, based on the encyclopaedic knowledge that tells us that some items cannot stay on the table (e.g., a fridge, a bookshelf, an elephant). However, beyond this, there is almost no certainty about how to fill the open variable configured by the non-exhaustive tags (i.e., *ya* and *nado*).

In (4.3), the speaker lists several situations in which road bikes are particularly useful. Contrary to (4.2), in this case, it is possible to identify a property shared by the mentioned elements: they are all activities in which road bikes may be used to replace other means of transport. For this very reason, the simple comparison between the mentioned elements in the specific context is enough to identify the shared property. Thus, by means of the property, the hearer can saturate the open variable configured by *ya* and *nado*.

Moreover, the difference can be easily detected also by referring to the concept of graded structure (cf. Rosch and Mervis 1975). Since categories exhibit graded structures, it is possible to indicate good examples and bad examples on the basis of the defining property. In (4.3), it is possible to affirm that 'playing frisbee' is not a good example of the category, whereas 'going shopping' may be a good example. On the other hand, in (4.2), there is no way to assert that 'bottle' is a better example than 'purse'.

The comparison between these two situations allows us to outline and contrast the distinctive features of the two, and, consequently, it allows us to draw some partial conclusions:

1. In case of enumeration, what is shared by the items is the membership to the list itself or - at most - the potential membership (that is, in the grocery list, not buyable items must be excluded). On the contrary, in a category, the items share a property which acts as a defining criterion to identify what types of elements are part of the category and what types of elements are not. This property can be defined simply by comparing each element of the category in a specific context.

2. Category members are organized in a graded structure. On the contrary, in enumerations, the items share the same status: there is no element that represents the list better than others. In other words, they share a (tautological) property that does not allow different degrees of membership: elements are part of the list or they are not part of the list.

3. Because we cannot identify an actual property, it is difficult to create labels for instances of enumeration. The only solution is to create paraphrases that start with expressions like "list of…" (e.g., "list of the things in the table").

What outlined above is a strict theoretical distinction, formulated in isolation, which then needs to be applied to real-life situations. This is the great bias of sentences like (4.2): like all invented sentences, not only do they give a distorted idea of the actual use of a certain linguistic construction, but they also lack a context. In other words, they do not represent the linguistic reality: out of 200 occurrences of *ya*, there is no evidence of sentences like (4.2) in our corpus (that is, existential clauses which describe the presence of certain items in a certain space in a certain moment). The same holds for all the other strategies under study. This virtually eliminates the clearest example of enumeration, namely a list of elements related by pure contingency.

Therefore, a question arises: is it indeed possible to find cases of pure enumeration in real-life sentences? According to what we know about coordination constructions (and thus list constructions) and according to our data, it appears very unlikely. There are at least two good reasons for this unlikeness.

First, in real-life situations, speakers create lists in order to carry out particular discursive purposes. This means that there is always a reason behind the creation of a list, and it

follows that the elements that are part of it are not chosen randomly, but with a very specific ratio. Under this respect, studies on the semantics of coordination (cf. Lang 1984, Mauri 2008) demonstrate that there is always a certain degree of underlying associability among elements in a coordination construction. Let us consider the following examples (Lang 1984:36):

(4.4)   *No entry for dogs and Chinese people!*
        (Sign board at a park entrance in a European settlement in pre-war Shanghai)

(4.5)   *Défense de cracher ou de parler breton*!
        Spitting and speaking Breton prohibited
        (Sign board in schools and offices in 19th century Britanny)

Both signs boards have a strong derogatory effect. In the first case, in addition to the negative effect produced by banning the access to a specific nationality, the issue is further amplified by the semantic nature of the other conjunct, i.e., dogs. In the second case, the request might not have an intrinsic derogatory nature (namely, the prohibition to speak Breton), but it becomes so the moment it is associated with the other conjunct, "spitting". If we assume that enumeration does not imply associability between conjuncts, then we cannot explain the derogatory effect amplified by coordination constructions.

This effect can be explained by referring to the notion of 'common integrator', the conceptual entity which is deduced from the combined conjuncts and which, at the same time, includes them, in that they are instantiations of this common integrator. In other words, the common integrator is what Lakoff (1971: 268) calls 'common topic', that is "that [semantic] part of each conjunct of the sentence that is identical", or the ground on which the two conjuncts are pertinently combined. In this regard, Lakoff explains that the common topic is not overtly present and identifiable in the sentences, but it is a necessary (but non-sufficient) condition to the coordination of elements (1971: 118).

We can understand the notion of common integrator while considering a range of coordination constructions where the first conjunct remains unchanged. Using the examples provided by Lang (1984: 26-27), consider the following:

(4.6)   (a) I need a book or some newspapers or magazines
        (b) I need a book or a record
        (c) I need a book or a cigar-box

Not only does the common integrator change according to the elements of the coordination, but also results in the narrowing down the interpretation of 'book' accordingly to (4.6a) 'something to read', (4.6b) 'entertaining present', (4.6c) 'solid object having the thickness of a book'. Hence, we can say that the lexical meaning of the second conjunct determines the interpretation assigned to the first.

These theoretical premises allow us to properly interpret the examples provided in (4.4) and (4.5), thus explaining the underlying derogatory effect. Lang notes that "the cognitive operation basically involved in the deduction of a Common Integrator is that of pairing the conjunct meanings in such a way that they come to hold an equal rank within a conceptual hierarchy" (1984: 35). In this sense, whenever the conjunct meanings are not equal ranking exemplifications of some common integrator, the achieved result is equalizing things that normally rank differently, creating an ironic (cf. Lang 1984: 35) or a derogatory effect, depending on the conjuncts and the context.

This excursus on the notion of common integrator suggests that, in real-life sentences, it is always possible to find a conceptual entity that encompasses all the elements of the list. This is especially true considering the essential role played by the context. As Lang himself notes "the deduction of the common integrator also involves various other factors obtainable only from either the situational context or the interactional setting of the given utterance or from extralinguistic systems of knowledge, belief-systems etc." (1984: 27).

The second reason for the unlikeness of enumeration is indeed the context itself. As noted in sections 1.2 and 1.3, the context plays an important role in the identification of the defining property underlying categories. It follows that even those cases that may be seen as instances of enumeration, the moment they are inserted in a specific context, they tend to be interpreted as instances of categorization. Consider again (4.2) in the context of an alcoholic novelist struggling to finish his latest novel. We can construe the same list as a category, identifying a defining property, good examples (e.g., balled up pieces of paper, bottles of wine) and bad examples (e.g., a calculator). Then, consider it again in the context of a college student studying for the finals and drinking too much coffee to stay awake. Again, we can construe the list as a category, with a different defining property. In this case, 'bottles of wine' is no longer a good example, 'calculator' may be a good example, etc. Indeed, we can imagine as many different scenario as we want, and every time the category will be different.

This little experiment is designed to demonstrate that 1) the context plays an important part in the interpretation of a list construction and that 2) whenever a sentence is inserted into a context, it prevails the tendency to interpret the list construction as categorization, rather than enumeration, because the context provides the information around which to build the category. Consider the following example:

(4.7)　*Yūdachi-go-no*　　　　　*kion-no*　　　　　　*henka-ya,*　　*uchimizu-no*

　　　　evening.shower-after-GEN　temperature-GEN　　change-YA　　spinkling.water-GEN

　　　　kōka-no　　　*kenshō-nado-ga*　　　　　*okonawareru.*

　　　　effect-GEN　　verification-NADO-NOM　　　perform:PASS

　　　　'The verification of the effect of the watering, the changes of the temperatures after the evening showers and so on are performed.'

Here, we may be tempted to interpret "the verification of the effect of the watering, the changes of the temperatures after the evening showers and so" as an instance of enumeration. However, since the broad context describes an interactive map that provides information about the weather, again the category interpretation seems more likely.

Even those lists that are not inserted in a textual context (i.e., co-text), are still part of a broader and always accessible extra linguistic context. Even the classic grocery list written down on a piece paper left on a table can be interpreted as a category, when it is read in the context of a family with food allergies and/or certain dietary habits. The extra linguistic context defines the property shared by the elements of the list.

Let consider another example from our corpus.

(4.8)　*1210 Man*　　*gaso*　*1/ 1. 7-Gata*　*CCD, kōgakute*　　　*bure*

　　　　1210 million　pixel　1/1.7- inch　　CCD optical　　　　camera.shake

　　　　hosei　　　*X ISO 6400 no kōkando,*　　*dōji*　　*8-ri*　　　　*made*

　　　　correction　　X ISO 6400 LK high.sensibility simultaneous 8-people　　till

　　　　kenshutsu　　*kanōna*　　*kao*　　*kenshutsu*　　*kinō*　　　**nado**.

　　　　detection　　possible　　face　　detection　　function　　**NADO**

　　　　'1210 million pixels, CCD 1/1 inches, high sensibility ISO 6400 of the optical camera shake correction, detectable face detection function up to eight people simultaneously.'

At first glance, this list seems a good candidate for enumeration: there is no (a lot of) textual context, there is no verb at the end of the sentence, which could have provided some

semantic clues to identify the common integrator. Yet, when we consider the extra linguistic context provided in Figure 4.1, the interpretation is totally different.

ソニーDSC-W200 1210万画 ヨット

BY ITTOUSAI ● 2007年05月14日 14時00分 ▶ 0

| | 0 | 0 | | |

f シェア   ツイート   B! はてな

そういえば2月末頃にリークのあった12.1メガピクセル サイバーショットDSC-W200が国内で正式 に発表されました。売りは世界初(※) 有効1210万画素 1/ 1.7型CCD(※民生用レンズ一体型カメラ において)、光学手ブレ補正 X ISO6400の高感度(※EX高感度モード時。画素数は300万相当)、同 時8人まで検出可能な顔検出機能「顔キメ」など。

Figure 4.1: Context of example (4.8)

The picture and the entire website act as extra linguistic context, pointing to the identification of the category 'functions of a latest state of the art compact photo camera equipped with sophisticated manual settings'.

Therefore, we argue that, in real-life situations, there cannot be instances of simple enumeration. Whenever a list construction is considered with respect to the co-text and the broader extra linguistic context, it always activates the presupposition that list members share some common property (or common integrator) and should therefore be considered exemplifications of the category defined by the property itself.

This does not mean that enumeration is an impossible concept, but just a very marked one, e.g., lists of numbers, lists of items in pure isolation without considering the extra linguistic context.

In the light of the above, the purpose of this chapter will be to examine how exemplifying constructions may activate and guide the categorization process, even without a category label. As stated in section 1.3.1, an exemplifying construction comprises 1) the use of a non-

exhaustive tag and 2) the mention of one or more examples of the category. Both these cores play essential roles in the cognitive and discursive process of referring to a conceptual category. In the following sections, we will discuss 1) the non-exhaustivity feature as a crucial linguistic tool to trigger the inferential processes, thus explaining why non-exhaustive tags are essential in exemplification, 2) the linguistic properties of the examples, and 3) the role of the context in directing inferential processes.

4.2 NON-EXHAUSTIVITY FEATURE

Up to this point, we have taken for granted the role of non-exhaustivity in categorization processes. This assumption was built on empirical evidence (cf. section 1.3): investigating the strategies identified by Mauri (2016) and Mauri and Sansò (forthcoming) to construct and communicate context-relevant categories, the common shared feature is that they all encode non-exhaustivity, that is, the presupposition of other further elements beyond those explicitly mentioned. Moreover, since we are indeed examining linguistic strategies that codify further reference, non-exhaustivity has been a pivotal feature from the beginning of this book.

This said, to understand the role played by non-exhaustivity, it seems crucial to challenge it. Is non-exhaustivity really an essential feature in categorization processes?

Consider again example (3.3), repeated here as (4.9), and the same utterance with the exhaustive connective *to* "and" instead of the non-exhaustive connective *ya* in (4.10).

(4.9)   *Koe-ya   ugoki-wa,   dokusha-no sōzō-no   hanchū   da.*
        Voice-YA   movement-TOP reader-GEN imagination-GEN   category   COP
        'Voices and movements (and so on) are categories of the readers' imagination.'

(4.10) *Koe-to   ugoki-wa,   dokusha-no sōzō-no   hanchū   da.*
        Voice-and   movement-TOP reader-GEN imagination-GEN   category   COP
        'Voices and movements are categories of the readers' imagination.'

In section 3.1, we have seen that the construction *koe ya ugoki* "voices and movements (and so on)" (exemplar + non-exhaustive tag + exemplar) is used to make reference to the category of features that remain prerogatives of the reader's imagination in a comic book adaptation. We proposed that this is due to the use of a non-exhaustive tag (i.e., the non-exhaustive connective *ya*) which triggers inferential processes, leading to the construction of the category. However, now that we are arguing against the role of non-exhaustivity, we

should wonder if the same category can still be inferred by the exhaustive construction *koe to ugoki* "voices and movements". In other words, can we infer, build and communicate categories in discourse even without using non-exhaustivity tags?

From our considerations, we must exclude all those constructions which comprise the explicit mention of the category label (i.e., lexicalized categories). In these cases, the presence of a non-exhaustive tags does not seem crucial, because the category is already explicit and in most cases (cf. chapter 3) directly linked to the examples. For instance, lexicalized categories can be expressed as similative constructions (e.g., X like Y), without any explicit non-exhaustive tag[44]. Therefore, we are just considering those cases in which the category must be inferred from scratches, that is, from one or more explicit examples. More specifically, we are mainly concerned with two situations: 1) the speaker provides one example, 2) the speaker provides two or more examples. The first situation is what Wilson and Carston (2007) call "category extension", a specific type of broadening that mainly regards salient brand names, personal names or – more rarely – common names that can evoke a broader category of elements (cf. section 3.1). However, this is not a process that can be applied to any kind of exemplars, but just to those that are culturally salient to the point of being able to represent an entire category. More generally, it is quite difficult, if not impossible, to make the hearer infer an entire category from just one example without any other category triggers in the specific discourse.

The second situation involves the use of two or more examples (i.e., lists of examples). In his discussion about the notion of 'common integrator' (cf. section 4.1), Lang (1984) makes no distinction between exhaustive and non-exhaustive lists. Accordingly, his claim holds for any coordinate construction: it is always possible to deduce (or better, infer) a common integrator, that is, a defining property shared by all the members of the list. As a matter of fact, almost all examples provided by Lang are instances of exhaustive constructions (1984: 25-37), and yet, he demonstrates how it is always possible to deduce the common integrator, or such entity that makes the conjuncts associable among themselves. Therefore, for example, in "I need a book or some newspapers or magazines" the common integrator is 'something to read', while in "I need a book or a record" the common integrator is 'entertaining present' (Lang 1984: 27), and so on.

---

[44] Despite the absence of explicit non-exhaustive tags (e.g., general extenders), also similative constructions encode non-exhaustivity at the semantic level. In other words, the implicit assumption (that is, the presupposition) of constructions like "things such as X" or "issues such as Y" is that the mentioned item is part of a larger set from which it was deliberately selected by the speaker.

Going back to our original question on the role of non-exhaustivity, the very notion of common integrator seems to argue against our assumption that only non-exhaustive constructions can effectively configure conceptual categories. We may even argue that there is always a shared category underlying coordination constructions. For instance, in "I need a book or some newspapers or magazines", the underlying category is indeed 'things to read'. This statement is problematic in the light of what we assumed at the beginning of this section: if there is always a categorization underlying list constructions, then what is the difference between an exhaustive construction and a non-exhaustive one? For example, we can argue that in both (4.9) and (4.10) the common integrator is the same, namely 'features that remain prerogatives of the reader's imagination in a comic book adaptation'. But then, what is the real difference between (4.9) and (4.10)?

Actually, Lang seems to notice the profound difference established by the exhaustivity, when he compares sentences (4.11a) and (4.11b).

(4.11) (a)   I need a book or something.
        (b)   I need a book or some newspapers or magazines.

Specifically, he notes that sentences like (4.11a) are open to a wide range of possible specifications, whenever they are taken in isolation. Or, to put it the other way round, there is a large class of possible contexts in which sentences like (4.11a) would fit into (1984: 26). This does not apply to sentences like (4.11b). In other words, Lang notes that the *or something* element is unspecified and that it assumes a specific semantic referent only when it is inserted and interpreted within a specific context.

Following this intuition, Barotto and Mauri (2016) suggest that the difference between exhaustive and non-exhaustive constructions lies in the distinction between the presupposition (cf. Levinson 1983) and "what-is-said" part of the utterance meaning (cf. Grice 1989, Recanati 2004). While the former represents the inference associated with utterances that generally conveys backgrounded, uncontroversial information with respect to the context of the utterance, the latter refers to the conventional meaning of a sentence and the truth-evaluable representation made available to the speaker.

On the basis of this distinction, Barotto and Mauri therefore propose the following generalization: a list construction always activates the presupposition that its members share some common Property P[45] and should therefore be considered exemplifications of

---

[45] Henceforth, we will use the term Property P instead of common integrator.

the category defined by the Property P; however, the Property P is not always part of "what is said", i.e., it is not always the explicitly communicated content of the utterance (cf. explicatures in Relevance Theory). So, for example, in (4.11b), "I need a book or some newspapers or magazines" activate the presupposition that the book, magazines and newspapers share some common Property P, namely 'things to read'. However, this is not part of the "what-is-said" part of the utterance: the speaker is not saying that she needs something to read, but that she needs something choosing between a book, some newspaper or some magazines. Therefore, the fact that it is always possible to activate the presupposition about the common property P does not imply that the category itself is actually communicated.

On the contrary, considering (4.11a) in the context of a person who needs something to read to kill the time[46], "I need a book or something" not only activates the presupposition that the list members share the Property P 'things to read', but it also indicates that the Property P is part of the explicitly communicated content. In other words, the "what-is-said" part of the utterance can be paraphrased as "I need something to read".

Consider the following (Barotto and Mauri 2016):

(4.12) (a)   Please go to the supermarket and [buy me some milk, flour and artichokes.]

The difference between the presupposition and the "what-is-said" part of the utterance can be schematized as follows:

What is presupposed: [$X_{(milk)}$, $X_{(flour)}$, $X_{(artichoke)}$] share some common Property P.
What is said: buy me the following things: [$X_{(milk)}$, $X_{(flour)}$, $X_{(artichoke)}$].

To understand what is (effectively) said, it is necessary to assign a referent to every list member but not to assign a specific value to the Property P. This means that even if the hearer cannot identify it, she can still go to the supermarket and buy milk, flour and artichokes, ensuring felicitous communication.

Now consider the non-exhaustive version of the same utterance:

(4.12) (b)   Please go to the supermarket and [buy me some milk, flour, artichokes and so on]

---

[46] As Lang notes (1984: 26), in this case, the context is essential to interpret the shared property of the list members.

The difference between the presupposition and the "what-is-said" part of the utterance can be schematized as follows:

What is presupposed: [$X_{(milk)}$, $X_{(flour)}$, $X_{(artichoke)}$, $X_{(unspecified)}$] share some common Property P.

What is said: buy me the following things: [$X_{(milk)}$, $X_{(flour)}$, $X_{(artichoke)}$, $X_{(characterized\ by\ P)}$].

Here, in addition to the presupposition of a shared Property P depending on the exemplars, we need to consider also the presupposition of the non-exhaustive tag *and so on* regarding the presence of further unspecified Xs characterized by the Property P. Therefore, it is necessary to identify a specific value for P in order to saturate the unspecified Xs, otherwise the utterance would be ambiguous. In other words, the hearer must be able to distinguish between possible Xs and impossible Xs. To achieve this, she must have access to the situational context. For example, if the context refers to the cooking of a quiche for dinner, the shared Property P will be 'ingredients for an artichoke quiche', and then X = eggs but *X = beer. However, if the context refers to the weekly grocery shopping, Property P will be 'things that are normally found in the kitchen, to give the idea that somebody actually lives in the house', and then X = beer but probably in a western house *X = sushi. It follows that if the hearer has no access to context and cannot identify a shared property P, she cannot saturate successfully the unspecified Xs (presupposed by *and so on*), thus leading to failed communication and potential misunderstanding.

Therefore, while the presupposition of the shared property 'ingredients for an artichoke quiche' is activated in both cases by the list construction (as Lang theorized), in (4.12a) the property is not necessary to understand the explicitly communicated content of the utterance (i.e., 'milk, flour and artichokes'). On the contrary, in (4.12b), the presence of a non-exhaustive tag activates also the presupposition of an unspecified open variable to be saturated on the basis of the property P, which is therefore essential in order to understand the "what-is-said" part of the utterance. As a result, in (4.12b), the superordinate category is actually part of the explicatures of the utterance, and therefore essential for the communication. We can schematize the difference as follows:

(4.13) (a)   Please go to the supermarket and [buy me some milk, flour and artichokes.]

= buy some milk, flour and artichokes

=/= buy the ingredients for an artichoke quiche

(4.13) (b)   Please go to the supermarket and [buy me some milk, flour, artichokes and so on]
              = buy the ingredients for an artichoke quiche = buy some milk, flour, artichokes and
              other ingredients that are needed for this specific receipt.

To explain further the role of non-exhaustivity, Barotto and Mauri (2016) suggest that non-exhaustivity tags function as indexical elements with two variables, which can be schematized as [Xs characterized by P] where the reference of P depends on the explicit members and context, and the reference of Xs depend on P.

Indexical expressions are strictly related to variables, since they adduce a lack of information. Yet, the reason why the indexical expressions do not communicate blanks, but plainly understandable pieces of information, is that usually they are inserted in a context able to saturate the variables. Specifically, in deictic indexicality the contextual source saturating the variable is extra-linguistic, since it depends on some contingencies of the utterance. For instance, the indexical word 'tomorrow' configures an open variable 'the calendar day that succeeds the time of speaking (X)', to be saturated only through access to the specific context of the utterance. In the same way, non-exhaustivity tags configure two variables, namely 1) unspecified Xs which are characterized by 2) the Property P.

Let us see some concrete examples from our corpus, starting from the sentence (4.9), repeated here as (4.14).

(4.14) *Koe-ya   ugoki-wa,   dokusha-no sōzō-no   hanchū   da.*
       Voice-YA   movement-TOP reader-GEN   imagination-GEN   category   COP
       'Voices and movements are categories of the readers' imagination.'

In the sentence above the list construction [*koe ya ugoki*] activates the presupposition that the list members share a common Property P. We may identify the Property P by examining the list members in the context they occur. In this article, the writer describes the process of adapting a novel in a comic book. Therefore, the shared Property P is 'prerogatives of the reader's imagination in a comic book adaptation'. Up to this point, the process would have been the same even with an exhaustive connective instead of *ya* like in (4.10). However, the presence of the non-exhaustive connective *ya* activates the presupposition of further members X, beyond those explicitly mentioned, characterized by the Property P. Thus, explicit list members *koe* "voices", *ugoki* "movements" and implicit members Xs that are 'prerogatives of the reader's imagination in a comic book adaptation' together constitute the superordinate category 'human features that remain prerogatives of

the reader's imagination in a comic book adaptation', which represents the actual "what-is-said" part of the utterance.

Let consider another example.

(4.15) *Kaitori saito-no*   *ii-tokoro*   *wa,*   *osusume*   *no*   *hon, DVD,*
     purchase website-GEN   good-point   TOP   recommendation   LK   book DVD
     *gēmu-**nado**-o*   *chekkushita*   *nochi,*   *sonomama*   *kōnyūdekiru*   *ten*   *desu.*
     game-NADO-ACC   check:do:PAST after   as it is   purchase:do:POT   point   COP:POL
     'The nice thing about this website is that you can purchase directly, after checking
     recommended books, DVD, games, etc.'

The hearer is actively called upon to infer other potential members of the category and the category itself, due to the presupposition activated by the non-exhaustive tag *nado*. Therefore, as a rule, we may say that whenever non-exhaustive tags are used, the following elements become part of the truth-evaluable representation available to the language users (Recanati 2004, Mauri 2016):

1. the reference to the explicit list members (i.e., books, DVD, games);
2. the context-relevant properties shared by the list members, identified through associative reasoning (i.e., entraining items);
3. the reference to a category comprising the list members and further elements sharing the context-relevant property, identified through abstraction (i.e., entraining items that can be purchased in this website).

For this reason, we refer to the entire process as indexical categorization, that is, an exemplar-driven process that implies a variable to be saturated through access to context (i.e. semantic indexicality). Specifically, indexicality is linguistically encoded by an array of possible morphosyntactic constructions (e.g., associative and similative plural markers, derivational collective morphemes, reduplication, non-exhaustive list constructions, cf. section 1.3.1, Mauri 2016) presupposing unspecified further reference, which should be saturated through access to the context.

To sum up, we argue that only in non-exhaustive lists the common integrator (or Property P) coincides with the "what-is-said" part of the utterance meaning, resulting in an instance of indexical categorization. For this very reason, non-exhaustive tags can effectively work as triggers to inferential processes. On the contrary, in an exhaustive list, the common

integrator (and thus the associated category) is just part of the presupposition. We refer to the latter case as *presupposed categorization* (Barotto and Mauri 2016).

4.3 LINGUISTIC PROPERTIES OF THE EXAMPLE(S)

Without an explicit label, examples are the only hints that direct the inferential reasoning towards the identification of the target category. In this case, the hearer is called upon to identify the shared Property P by means of comparison. As Lang (1984: 30) notes, examples are "mutually determined, weighed up against each other, or integrated with each other". Ultimately, their interpretation narrows down to some shared property that will cover them all in a specific situational context.

Because of this, we may assume that the speaker is compelled to choose carefully the examples, so that they express all the important features of the category (cf. Taylor 1995: 40), thus directing the inferential processes.

However, at the linguistic level, we should also wonder if the language has an active role in triggering and facilitating these inferential processes. In other words, are the examples encoded in such a way as to facilitate the inference of the category? In this section, we will examine some linguistic properties of the examples, with the aim of providing an answer to this question.

4.3.1 Syntactic and semantic properties of the example(s)

The first parameter concerns the syntactic and semantic properties of the example(s).
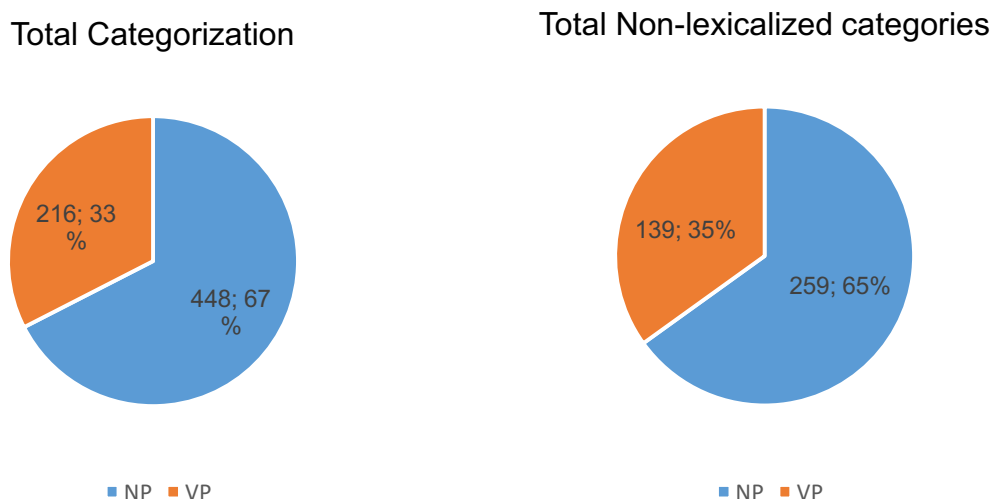
Figure 4.2: Distribution of examples expressed by noun phrases and verbal phrases (non-lexicalized categories)

Table 4.1: Distribution of examples expressed by noun phrases and verbal phrases in lexicalized categories (and in total)

	NP	VP
ya	108 (170)	0
nado	84 (170)	9 (24)
tari	0	112 (148)
toka	67 (108)	18 (44)
Total	259 (448)	139 (216)

The first point of interest in Figure 4.2 and Table 4.1 concerns the numbers of occurrences of lexicalized and non-lexicalized categories regarding *tari*. Comparing the data in chapter 3 with those above, it appears that *tari* tends to occur more frequently in constructions that lack an explicit category label. This can be ascribed to two reasons. The first one is a structural reason. Being a non-finite form of the verb which canonically requires the verb *suru* ("to do") immediately after, it may be more complicated for the speaker to attach a label to this type of exemplifying construction. Nevertheless, we should note that whenever *tari* is attached to a category label, it does not follow a strict pattern. For instance, many different connectors are attested (sometimes even other exemplifying strategies, such as *nado* and *toka*) to link *tari* with the category label. Therefore, while this structural fact may play a part in the low frequency of lexicalized categories in combination with *tari*, it still seems an obstacle that speakers may overcome, especially in the spoken language, when they can act with more freedom.

As for the other reason, *tari* is the preferred strategy whenever the examples are expressed by verbal phrases. Therefore, we may argue that the high frequency of non-lexicalized category in combination with *tari* is linked to the syntactic properties of the attached examples (i.e., verbal phrases). This confirms what has already emerged in chapter 3, namely that the lexicalization of categories designated by verbs is less frequent because the identification and lexicalization of the shared property of verbs is less straightforward.

Nevertheless, the tendency to use examples encoded by noun phrases (65%) instead of verbal phrases (35%) persists even when we consider only non-lexicalized categories. It follows that even beyond the process of lexicalizing a category, the preference for examples expressed by noun phrases is a general tendency of the exemplification process.

As it was noted in chapter 3, nouns and verbs exhibit different degrees of complexity. Givón (2001) describes (prototypical) nouns as "multi-featured bundles of experience [...]

Consequently, when either rapid change or deviance crop up in one feature, the relative stability of the rest insures that a deviant individual remains within a reasonable range (standard deviation) of the population's prototype (mean)." (2001: 51). Therefore, in addition to exhibit easily identifiable features such as size, shape, colour, weight, sound, smell, part-whole composition, behavioural propensities, cultural uses, etc., they exhibit also temporal stability ("nouns change only little over repeated perceptual scans" 2001: 51) and tend to be spatially compact. It follows that it is easier to identify them since they are not scattered all over the perceptual space and through time. Moreover, they are also easier to elaborate, since they are formed by conceptually and sensorially salient features. On the contrary, verbs exhibit low temporal stability, as most of the times they code rapid changes, and are spatially more diffuse. Moreover, "while not quite as multi-featured as nouns, prototype verbs often exhibit considerable complexity" (2001: 52) which consists in the presence of several distinct participants, all of them individuated, spatially compact, temporally durable entity in its own right. In other words, not only verbs do not show sensorially salient features (which are easy to grasp for the human brain) like nouns, but they are even more complex, as they presuppose interconnections among the participants.

These observations have been confirmed and further investigated by Langacker (1987a 1991b). What Givón identifies as temporal stability, for Langacker becomes cognitive stability. He notes that while a noun profiles (i.e., designates) a thing, a verb profiles a process. This has consequences on how the human brain configures both, in the sense that "whereas a noun profiles a thing, a relational predication designates a set of interconnections. A verb, moreover, is an especially complex relation, in that it profiles a series of relational configurations, and further specifies their continuous distribution through time" (1991b: 21-22). In other words, for example, while verbs "represent a higher level of conceptual organization" (1991b: 20) as they relate to interconnections and time, encoding the evolution of a particular event through time, this is not the case for nouns. Langacker agrees with Givón to place nouns and verbs at opposite extremes on the category spectrum (1991b: 19), however he shifts the focus on the level of internal organizational complexity.

These differences in the cognitive configuration of nouns and verbs helped us to explain the difficulties in lexicalizing categories of examples encoded by verbs in chapter 3. However, the identification of a common property is not only the basis of the lexicalization process, but it is also the basis of the entire exemplification process. In other words, having access to the Property P is what allows people to construct context-relevant categories: if the hearer is unable to properly track this property, he or she cannot infer the category. Therefore,

ideally, examples should be simple to understand and elaborate in such a way that makes it easy for the hearer to compare them in order to find what they have in common. According to the insights provided by Givón and Langacker, we argue that it is easier to elaborate and compare things (that is, nouns) instead of elaborating and comparing processes (that is, verbs). Things are stable in time and compact in space, they have features that are sensually salient and easy to cognitively grasp (e.g., the colour, the method of use, the shape...). Processes are complex interconnections of entities (i.e., the participants), they extend through time and space. Consider the followings:

(4.16) *Kasai-ya   hason   nado   niyoru songai-o   kabāsuru.*
       fire-YA   damage   NADO   due to damage-ACC cover:do
       '[It] covers damages due to fire, damage and so on.'

(4.17) *Sarani   dōten-de   wa,   byōki no   petto-o   kakaeteitari*
       furthermore   same.store-LOC   TOP   disease LK   pet-ACC   have:STA:TARI
       anrakushi   nitsuite   nayandeitari   nado   suru   kainushi   to
       euthanasia   about   be.worried:STA:TARI NADO do   pet owner   and
       sono kazoku-ni   kaunseringu-o   okonau   to   shiteiru.
       that   family-DAT   counselling-ACC   perform   LK   do:STA
       'In the same shop, [they] are performing counselling for pet owners and their families who are worried about euthanasia, have pets with diseases, and so on.'

In (4.16), to have access to the defining property, it is necessary to deduce that the two examples are both "things" that affect physical objects, therefore considering their method of use. On the contrary, in (4.17), the hearer must have access to the broader context (i.e., the fact that the article refers to a pet shop that performs also veterinary counselling and support) to understand correctly what 'having an ill pet' and 'being worried about euthanasia' have in common. Moreover, beyond the action/state itself (i.e., *to own* and *to be worried*), the hearer must also consider all the interconnections profiled, such as *the pet* and *the euthanasia*.

Even more without the mediation of a category label, the processes of elaborating and comparing examples are pivotal to infer the category, and this may be the explanation as to why speakers prefer to provide examples that are coded as noun phrases. Furthermore, this preference grows stronger whenever the property is lexicalized by means of a category label.

These observations are supported also by data about the semantic properties of the examples, as shown in Figure 4.3 below.
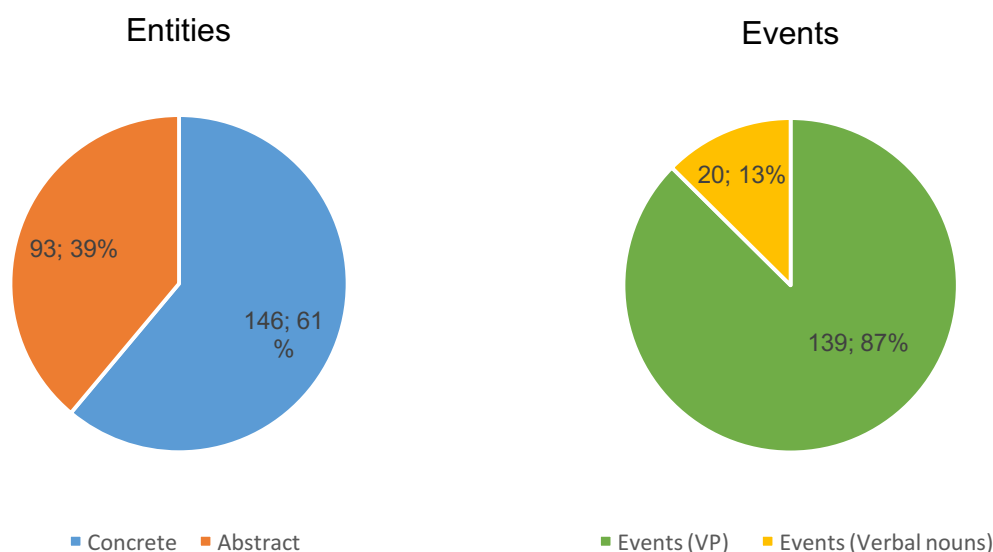
Figure 4.3: Semantic properties of the examples (non-lexicalized categories)

These data confirm that categories of entities are more frequent than categories of events, as already established in chapter 3. Two further points of interest also emerge from Figure 4.3.

The first one regards the general frequency of all semantic properties. Categories of entities are still more frequent than categories of events. Nevertheless, the percentages are not particularly different, unlike in the case of lexicalized categories (cf. chapter 3). For instance, the occurrences of category of events (159) are overall more numerous than those of categories of concrete entities (146). It follows that despite a general preference for entities (in particular, concrete entities), the absence of a category label enables the speakers to make easier reference to a wider range of categories.

The second point of interest is the frequency of verbal nouns used to codify categories of events, as shown in the example below.

(4.18) *Nyūgaku-ya*   *shūshoku-nado-de*   *tan-chōkikan,*   *biza-ga*
studying.abroad-YA   finding.job-NADO-CAUS   short-long period of time   visa-NOM
hitsuyō   *tonaru*.
necessary   become
'(To stay) for a short or long period to study or to look for a job (or something similar), a visa becomes necessary.'

In the sentence above, the speaker refers to a category of actions, yet the examples are expressed by verbal nouns, i.e., *nyūgaku* "studying abroad" and *shūshoku* "finding a job", instead of verbs or verbal phrases.

Verbal nouns represent the 13% of the total occurrences of non-lexicalized categories of events. Despite not being a particularly widespread strategy, this figure is still significant. Japanese exhibits one dedicated strategy to codify categories of examples encoded by verbs (i.e., *tari*) and two strategies that optionally can perform the same function (i.e., *nado* and *toka*). Moreover, out of the 20 occurrences of verbal nouns, 11 are attested with the connective *ya*, which cannot be used with verbal phrases. We thus argue that using verbal nouns instead of verbal phrases is indeed a deliberate choice of the speaker, and not a last resort for lack of better strategies. As it was already noted in chapter 3, this alternative strategy is of particular interest, since the use of verbal nouns "involves a conceptual 'reification' of the designated process" (Langacker 1991a: 20). For example, regarding the semantic features of *–ing* in English, Langacker (1991b: 32-34) notes that it nullifies the sequential scanning which is characteristic of all sort of verbs, thus turning a process into a complex atemporal relation. So, while the verbal noun incorporates the conception of a process, it does not profile interconnections but rather an abstract region (1991b: 37). Therefore, from the point of view of the categorization process, a comparison between actions which excludes any relational configurations and the continuous distribution through time requires a minor cognitive effort, simplifying the process of abstraction.

Taking all these issues into account, what emerges is a general tendency to choose and encode examples in such a way to facilitate the comparison among them and ultimately the inference of common shared property. Moreover, this trend seems to increase whenever the speaker is also called upon to lexicalize the category, although it stands quite strong even when the lexicalization is not needed.

4.3.2 Number of examples

The other parameter to be addressed concerns the number of the examples that are explicitly mentioned by the speaker. Again, we can distinguish three patterns:

1. One example + non-exhaustive tag

(4.19) *Kokkai-shingi-nado-de*   *toriyame-ni*   *natta*   *rei-wa*   *aru*
parliament-deliberation-NADO-CAUS   cancellation-DAT   become:PAST   example-TOP   AUX

ga,   *kono*   *ni-wa*   *kokkai-wa*   *kūtenshiteori,*   *yotei-jikoku-ni*
but   this   day-TOP   parliament-TOP   idling:do:STA:GRD   schedule-time-LOC

raikyaku-wa   *nakatta.*
visitor-TOP   AUX:NEG:PAST

'Although there were instances of [the doorstep interview] being cancelled due to Diet deliberations and such, this day the Diet has been idle and there was no visitor at the scheduled time.'

In the sentence above, the author mentions just one example, namely *kokkai shingi* "Diet deliberations", to which the general extender *nado* has been added. The reference is to a category of situations in which the prime minister cancels doorstep interviews with reporters. In this case, to elaborate correctly the exemplar and then to infer the category, it is necessary to draw on the context. Here, *kokkai shingi* "Diet deliberations" is relevant in the sense of being an important event where the prime minister is required to attend. In particular, it acts also as a counterpoint with the fact that, in this specific case, the interview has been cancelled for no obvious reason, thus implying that these events such as the Diet deliberations can justify the cancellation. Taking into consideration these clues, the reader is then able to properly build the category, inserting in it only politically important events that cannot be cancelled or postponed for an interview.

(4.20) *Mēru-toka-o*   *shihajimeta.*
email-TOKA-ACC   begin:PAST
'I began with emails and such.'

Again, in the sentence above, there is just one example, *mēru* "emails" to which the general extender *toka* has been added. Compared to example in (4.19), in this case it is even more important to draw on the linguistic context and encyclopedic knowledge to infer the category. The speaker is describing how she started using mobile phones. With a little background knowledge of the history of mobile phones in Japan and by reading the entire text, we can guess that she is making reference to the types of emails and other electronic messages that were available in the first mobile phone models. She is not referring to social networks or other online services. Therefore, again, to correctly interpret the mentioned

example, the hearer needs to rely on the broader context. Without having access to the context, it becomes hard to figure out what aspect of *mēru* "emails" is essential to properly build the category (as can be deduced by simply reading the sentence as it was reported above, without any reference to the context).

2. Two examples + non-exhaustive tag

(4.21) *Shiryō-o*   *minagara*   *sagyō-o*   *shitari, fotoretatchi-o*   *shitari suru*
       data-ACC   look:while   work-ACC   do:TARI photo.retouching-ACC   do:TARI do
       toki-nado-wa,   *yuaru-monita nishite tsukatteimasu.*
       time-NADO-TOP   dual.monitor   only   use:STA:POL
       'When you retouch photos or you work while looking the data (or other similar situations), you use only the dual monitor.'

In (4.21), there are two examples: *shiryō o minagara sagyō o suru* "working while looking at the data" and *fotoretatchi o suru* "photo retouching". These examples are linked by means of the non-exhaustive connective *tari*. In addition, the exemplifying construction is incorporated in a temporal subordination encoded by the temporal connective *toki* ("when" but also "time(s)") and to which the general extender *nado* has been added to further highlight the non-exhaustive interpretation.

In this case, the hearer can compare the two mentioned examples in order to understand what they have in common in this specific situational context. Therefore, unlike the previous examples in (4.19) and (4.20), here in (4.21), the examples act as starting points for an associative reasoning, helping the hearer to deduce the common shared propriety, without having to rely solely on the context.

(4.22) *Kenka toka   arashi mo   ōi   ne.*
       fight   TOKA   troll   also   many PP
       'There are also many fights and trolls and such things!'

In the sentence above there are two examples: *kenka* "fights" and *arashi* "(internet) trolls", which are linked to each other by means of the non-exhaustive connective *toka*. Even without knowing the broad context, the simple comparison of the mentioned examples helps to understand what aspect of the two items is crucial to build the category that the speaker wants to communicate. Specifically, the second example *arashi* "(internet) trolls" is much more specific than the first, *kenka* "fights", thus it defines more precisely the relevant

semantic field. In the previous section (4.1) we have seen similar cases indicated by Lang (1984: 26-27), where the lexical meaning of the second conjunct determines the interpretation assigned to the first general conjunct (e.g., in 'book or a record', 'record' narrows down the interpretation of 'book' to ''entertaining present' which is the actual common integrator of the coordination construction). In (4.22), *arashi* "(internet) trolls" determines the interpretation of *kenka* "fights" as 'internet fights'. Therefore, even without any reference to the context, we can still configure the category "negative and annoying experiences people encounter on the Internet".

Indeed, the utterance in (4.22) and the one in (4.20) are from the same interview. The girl describes how toxic Internet can be, while talking about smartphones and other online services in general. While the context is always an important factor to consider as to correctly understand the category, the simple presence of two (or more) examples triggers an associative reasoning that facilitates the identification of the shared property and, ultimately, of the relevant category. Moreover, the semantic relationships between the examples can make a further contribution to the elaboration of the category.

3. Three or more examples + non-exhaustive tag

(4.23) *Shoseki-ni   fusen-o   hattari,   sen-o   hiitari,   orikaeshi-o*
       book-DAT   label-ACC   stick:TARI   line-ACC   draw:TARI   wrapping-ACC

       tsukeru-no-wa   yoku yarimasu.
       attach-NML-TOP   often   do:POL

       'I often use the flap of the book cover, underline, attach a label and so on.'

In (4.23) we identify three examples: 1) *fusen o haru* "to attach a label", 2) *sen o hiku* "to underline", 3) *orikaeshi o tsukeru* "to use the flap of the book cover". The examples are linked by means of the non-exhaustive connective *tari*. Together, they make reference to the category of things that people do in order to study or memorize a book. Like in (4.21) and (4.22), the comparison among the mentioned examples triggers an associative reasoning, as to infer the shared property.

Naively, we may think that the more examples are provided, the easier it gets for the hearer to understand correctly not only the shared property of the mentioned members, but also the category itself. Nevertheless, looking at the data, we can identify a different and very precise tendency.
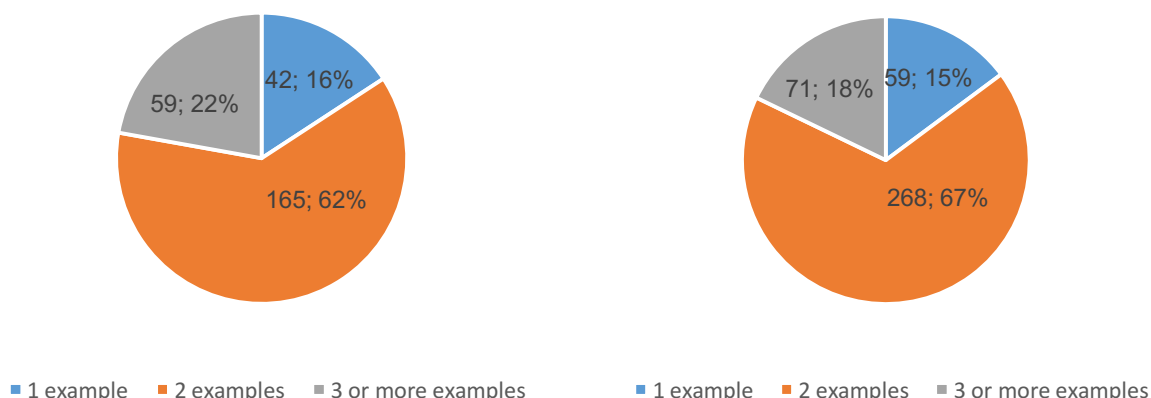
Figure 4.4: The number of examples in lexicalized and non-lexicalized categories.

Table 4.2: Numbers of examples in non-lexicalized categories.

	1 example	2 examples	3 or more examples
ya	0	91	17
nado	30	39	24
tari	6	87	19
toka	23	51	11
Total	59	268	71

What emerges from Table 4.2 is that, while all strategies are indeed attested, there is a strong tendency to provide two examples, rather than providing just one, or rather than providing three or more.

The fact that exemplifying constructions with just one example are not particularly widespread should be addressed carefully. First, we should note that three out of four strategies (i.e., *ya, toka, tari*) are connectives, which implies that, in these cases, at least two explicit examples are needed. Therefore, for example, the fact that there are no instances of just one example with the connective *ya* is consistent with this structural constrain. However, we also need to consider that three out of four strategies (i.e., *nado, toka, tari*) are also used as general extender, that is, they can be used with just one example. Thus, for instance, while *toka* can be used as a connective with two or more examples, it can also be used as a general extender with just one example. This latter strategy seems to be quite common. As a result, while there may be some biases due mainly to the connective *ya*, there are enough data with general extenders to mute the impact on the final numbers.

As for the usage of three or more examples, there is nothing in the modality of use of the analysed strategies that may justify a preference for mentioning two examples, rather than three or more.

These figures suggest that the presence of at least two examples seems to be crucial to the construction of categories. As we have seen above, this is due to the fact that two examples are the minimum to deduce the common shared property by associative reasoning. Without a second element, the hearer is forced to elaborate the only mentioned example according to the situational context. We may thus argue that comparing two elements that are supposed to be similar to each other in some respect requires a minor cognitive effort than analysing a single exemplar in a (multidimensional) situation context in order to grasp what aspect of it is so relevant to the point of building a category around that property. Consider the following variations of (4.9):

(4.24) *Arashi toka mo   ōi   ne.*
      troll   TOKA also   many   PP
      'There are also many trolls and such things!'

(4.25) *Kenka toka   mo   ōi   ne.*
      fight   TOKA also   many   PP
      'There are also many fights and such things!'

Both examples are part of the category the speaker wants to communicate and both have been used by the speaker herself to make reference to the category in (4.22). Therefore, we may think that they are both well qualified to represent the category from which they have been selected.

In (4.24), the example *arashi* "(Internet) trolls" is specific enough to not require further contextualization. However, with no other indications, it can be difficult for the hearer to understand what aspect of the example must be taken into consideration in order to build the category: for instance, people like trolls (e.g., haters, fake accounts and so on) or situations such as when you have to deal with trolls (e.g., flames, hateful comments and so on)? Given the specific semantic meaning of this example, in both cases the inferred category is not too far from the one that the speaker had in mind, but this may change substantially if we consider a more generic term as an example.

In fact, in (4.25), *kenka* "fights" is a generic term that encompasses the act of having a heated argument with strangers but also the type of everyday quarrel between siblings or

lovers. Without any context, it is practically impossible to identify the defining property of the category. And even considering the context, it may still be cognitively taxing. Is she referring to have quarrel with people that she knows on the Internet? That is a type of category. Is she referring to have arguments with fake accounts made with the purpose of bother people? That is another type of category.

All these issues are minimized in (4.22) where the comparison between the two examples help the hearer to grasp immediately the important aspect of both items, that is, the property around which to build the category. As Lang (1984: 27) suggests, the lexical meaning of the second example determines the interpretation assigned to the first, narrowing it down to some common property that cover them from a particular point of view. For instance, *arashi* "trolls" suggests that we are dealing with Internet people that like to have silly arguments with strangers for the sake of it, *kenka* "fights" suggests that we are dealing with a range of situations, and not with a range of people. Combining these elements, we can construe the category in a much easier way, without having to dig excessively into the context.

In fact, without a second example, it is the situational context to determine the interpretation of the example, and therefore the common property of the category. Ideally, whenever the speaker wants to use a single example, he or she should select a very specific and representative item, as to direct correctly the inference. However, this may require a greater cognitive effort and the risk of misconstruing is not completely neutralized.

Still exemplifying construction with just one example are used, and, from our data, it seems that this frequently happens 1) when it is easy to select a good exemplar of the category, 2) when the elaboration of the exemplar in a situational context is not cognitively demanding and 3) when the category is fairly homogeneous. An instance of this was provided in (4.20). Not only it is not difficult to elaborate *mēru* "emails" in the context of a girls describing her first experiences with a mobile phone, but it also represents a homogenous category of not very sophisticated message-type of communication. This is a substantial difference with instances such as (4.22), where the members of the category appear to be more heterogeneous. Consider also the following examples:

(4.26) *Furuhonya-toka-no*   *naka, haraisage no amerikankomikku-ga*
       second.hand.bookstore-TOKA-GEN   inside on sale   LK   american.comics-NOM

       aru   ne.
       AUX   PP

       'In second hand bookstores and such, there are American comics on sale!'

(4.27)   *Hyōko wa*   *hakuchō-rui-nado-no*   *kyūsoku-chi-ni*   *natteiru.*

         hyōko TOP   swan-kind-NADO-GEN   resting-place-DAT   become:STA

         'Lake Hyōko has become the resting place of swans and other similar birds.'

In (4.26), interpreting the exemplar *furuhonya* "second hand bookstores" according to the context of buying American comics is not particular demanding and the category is also quite homogenous (places where you can buy comics). The same applies to *hakuchō-rui* "swan type birds" in (4.27).

This shows some consequences also at the linguistic level. Consider the comparison between the numbers of examples and the syntactic properties of the examples:
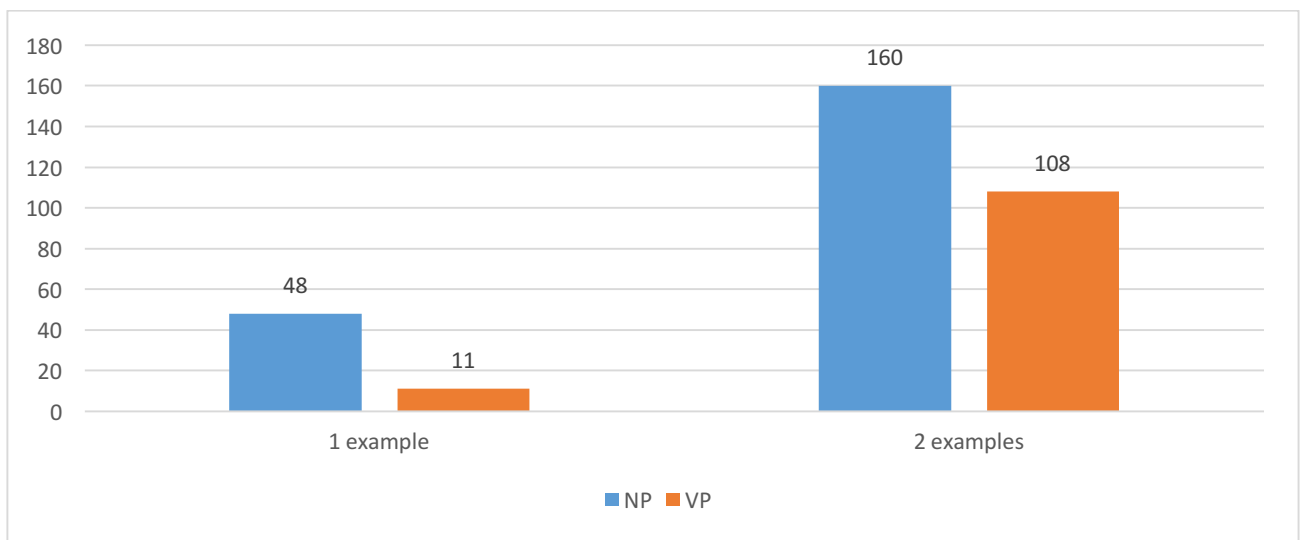
Figure 4.5: Number and syntactic properties of the examples (non-lexicalized categories)

Whenever the speaker provides one example, the tendency of encoding this single example as a noun grows stronger (80%), considering the average for non-lexicalized categories (65%). On the contrary, the tendency is less strong whenever the speaker provides two examples (60%). As it was previously noted, being less complex, nouns tend to be more easy to process and elaborate rather than verbs. Therefore, it is interesting to note that the general preference for examples expressed by noun phrases increases when there is just one example, since it must be interpreted according to the situational context.

What has been said above helps us to explain the tendency towards the use of two example rather than a single one. However, it does not help to understand the preference for two examples instead of three or more. On the contrary, following the same reasoning, we might expect a preference for a higher number of examples, since the more examples

are provided by the speaker, the easier it gets to infer correctly the category for the hearer. Nevertheless, the data suggest a different picture.

This tendency can be explained through the notion of linguistic economy, in the sense of "the principle of least effort", which consists in tending towards the minimum amount of effort that is necessary to achieve the maximum result, so that nothing is wasted. Since two examples are the minimum in order to infer the lowest common denominator by associative reasoning, adding further elements may be considered redundant and ultimately not useful for the creation of the category. Furthermore, we can also consider a purely cognitive reason: if each example should be considered as a further clue to infer the category, it follows that every time a new example is added to the list, all the previous examples should be re-interpreted again on the basis of the new one, as to find the common property that covers them all (cf. Lang 1984, Perelman and Olbrechts-Tyteca 1969). For instance, if we add a *video game* to the list of examples 'a book or a magazine', considering also the context, the common property may shift from 'something to read' to 'something to kill the time'. With a long list of examples, this process may be not only cognitively taxing, but also confusing.

Nevertheless, a long list of examples can be used to achieve different communicative purposes, beyond the construction of the category. For instance, three or more examples can be used to emphasize the number of members in the category or their heterogeneity (cf. Ovestreet 1999: 45).

Consider the following example:

(4.28) A: *4-Nenkan-no daizaichū,   iroirona koto-o   sareteiru   n   desu ne?*
        4-years-gen   stay   various thing-ACC   do:hon:stat   NML   COP PP

    B: *Bāru-ya resutoran-de   arubaitoshitari,   juerī-no   mise-de*
        bar-YA restaurand-LOC   part.time.job do:TARI   jewellery-GEN store-LOC

        hataraitari,   yūjin-to   juerīsukūru-o   uneishite   mitari to,
        work:TARI   friend-COM   jewellery.school-ACC management:do:GRD try:TARI LK

        kōkishin-no   omomuku   mama charenjishite   mimashita.
        curiosity-GEN   move   as   self-challenge.do:GRD   try:POL:PAST

    A: 'During the 4-year stay, have you done a lot of things?'

    B: 'I worked part-time job at restaurants or bar, I worked in a jewellery store, I tried to manage a jewellery school with a friend, I tried to challenge myself engaging my curiosity.'

Here, during an interview about her time in Italy, the speaker answers about the kind of activities she did while she was abroad. In particular, the interviewer asks precisely if she

did a lot of things. Her reply is a quite long list of activities, just to confirm that indeed she did many and various things while in Italy. Therefore, in this case, the long list of examples actually helps to configure a wide and heterogeneous category much better than a simple affirmative answer.

Another example is provided below.

(4.29) *Hobo*   *furītā*   *deshita*   *yo.*

almost   part-time worker   cop:PAST:POL   PP

Honyaku-o   *yattari, web seisaku-o*   *kojin-de*   *uketari,*

translation-ACC   do:TARI web work-ACC   private-STR   receive:TARI

haken   *no*   *SE*   *toka.*

dispatch   LK   system engineer   TOKA

'I was almost a part-time worker. I did translations, I did web works as a private (worker), dispatched system engineer and so on'

Here again the speaker uses a list of three examples to emphasize that he did various jobs as a part time worker in a particular moment of his life.

The same applies also for the examples given in (4.23), repeated below as (4.30):

(4.30) *Shoseki-ni*   *fusen-o*   *hattari,*   *sen-o*   *hiitari,*   *orikaeshi-o*

book-DAT   label-ACC   stick:TARI   line-ACC   draw:TARI   wrapping-ACC

tsukeru-no-wa   *yoku*   *yarimasu.*

attach-NML-TOP   often   do:POL

'I often use the flap of the book cover, underline, attach a label and so on.'

The writer comments on her struggling in remembering books that she has read, so she explains all the things that she tries to do in order to avoid forgetting important passages. The long list of example thus emphasizes the efforts to find a way to resolve her issues with memory.

Of course, in some cases we cannot find a connection between the use of three or more examples and the highlighting of the large number and/or the heterogeneity of the elements that are part of the category. Nevertheless, it is a connection that occurs frequently in our corpus and that should be investigated further, especially considering the differences between spoken and written language.

4.3.3 General considerations about the animacy parameter

A semantic parameter that has not been considered so far is animacy. While our data confirm a general tendency to use concrete entities as examples, yet animate entities are a rarity, as shown in Figure 4.6 below.
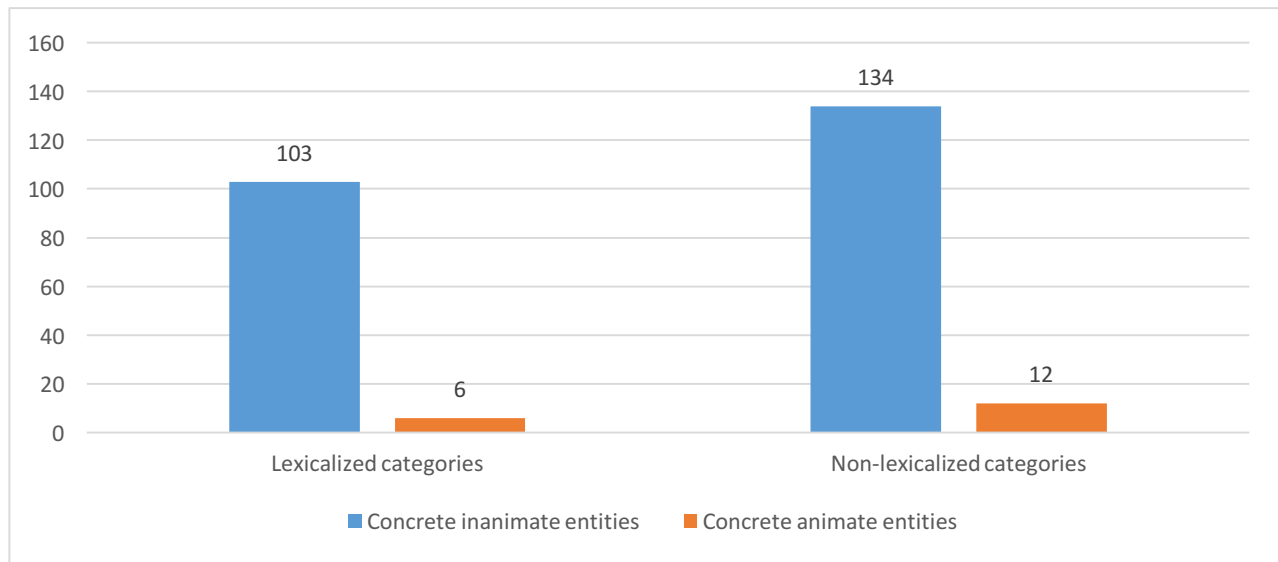
Figure 4.6: Frequency of animate and inanimate entities as examples

We propose two explanations for this low frequency of animate entities used as examples.

First, it is worthy to note that many categories of animate entities are prototypical natural categories, following the definition provided by Rosch (1973), that is, stable categories which reside as knowledge structures in long-term memory and are associated with familiar words (e.g., cats, dogs, mammals, vertebrates). As a consequence, at the linguistic level, there are plenty of rich taxonomic lexical hierarchies regarding categories of animate entities, even more than of concrete objects. Many of the terms (or short expression) to indicate animate entities have been created to establish scientific taxonomies, and while some of them are still prerogative of the scientific lexicon, others are also part of the everyday lexicon (e.g., felines, primates, mammals). Moreover, for all these reasons, there are also less lexical gaps in these hierarchies. It follows that in most cases, to designate these categories, it is more direct to use simple (e.g., insects, poultry) or complex labels (e.g., small mammals) rather than providing concrete examples.

In fact, as Figure 4.7 shows, it appears that exemplification is used just to fill those few gaps in the taxonomic lexical hierarchies. Consider the distribution of 1) proper names, 2) human common nouns and 3) animate non-human common nouns.
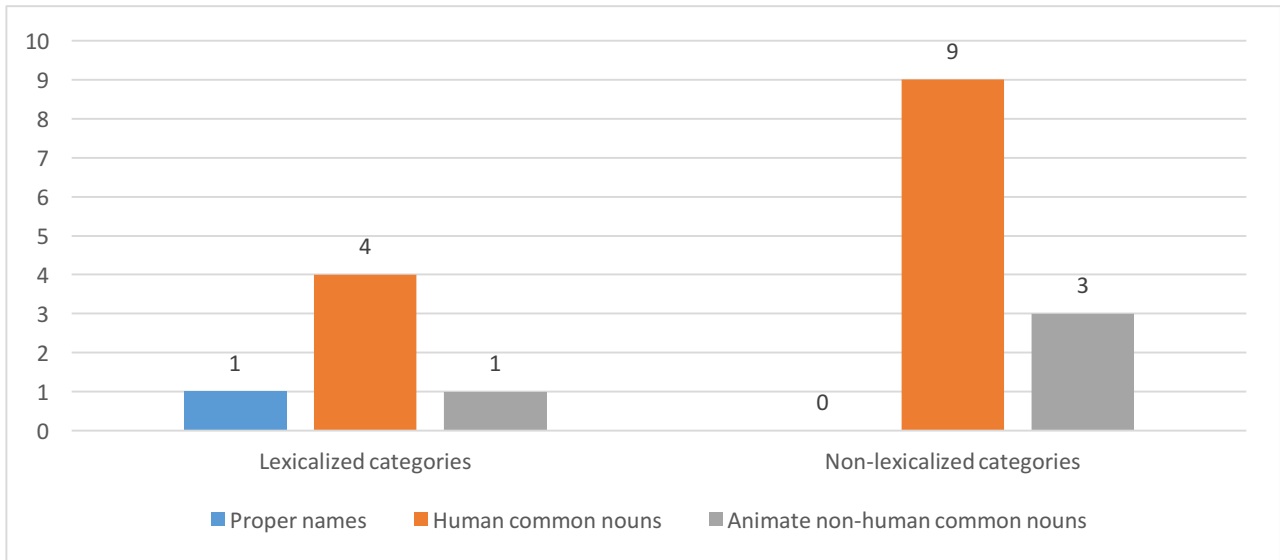
Figure 4.7: Frequency of animate entities as examples.

Exemplification is more frequently used with human common nouns. More specifically, out of 13 occurrences of human common nouns used as examples, 11 occurrences designate superordinate categories of professions, as shown in the following examples:

(4.31) *Ishi   ya   kankoshi,   yakuzaishi-ra yaku 70-ri   ni   sesshushita.*
　　　doctor YA   nurse   pharmacist-PL about 70-person   DAT inoculation:DO:PAST
　　　'About 70 doctors, nurses, pharmacists and so on have been inoculated [with the vaccine].'

(4.32) *Yūmeina   kagaku-sha-ya   supōtsu-senshu,   ātisuto nado, sekai-ya*
　　　famous   scientists-YA   athlete   artist NADO world-YA
　　　jikoku-de   toppu-de   aru   koto-o   shōmei   dekireba,
　　　country-LOC top-COP   AUX   NML-ACC   proof   POT:COND
　　　kono hōhō-de   eijū-ken-o   eru-koto-ga   dekiru.
　　　this way-STR   permanent-residence-ACC   obtain-NML-NOM   POT
　　　'If you can prove to be the top in the world and/or in your country, such as famous scientists, athletes, artists, and so on, in this way, it is possible to obtain a permanent residency.'

There are not a lot of hypernymic terms to designate superordinate categories of professions, and while the creation of ad hoc complex label is always an option, the professions used as examples are still more salient since they mostly reside in the basic level of categorization and refer to human activities: in (4.31), 'doctor' and 'nurse' are much more accessible and easier to process than 'people who work in the medical field'. Unsurprisingly, in most cases, labels are not used at all.

On the contrary, there are plenty hypernymic terms to designate superordinate categories of animals, therefore the need of providing concrete examples decreases accordingly.

Finally, the second reason behind the low frequency of animate entities regards the fact that the animacy hierarchy is closely associated with the definiteness hierarchy (Croft 2003: 132). More definite referents tend also to be the higher in the animacy hierarchy (i.e., pronouns and then proper names). This can be an issue, since categories are the result of grouping together items that share some properties in a situational context. Therefore, we may argue that in the moment these elements are categorized together, they stop having a specific individual definiteness and become simply representative of a wider set. This can at least partially explain why we tend not to create categories of definite referents, such as pronouns or proper names. Unsurprisingly, the only occurrence in which a proper name has been used as an example is more a similative construction, rather than an exemplifying construction:

(4.33) *Sawanobori Midori-san toka, iron'na hō-ga sekai-o butai-ni*
Sawanobori Midori-mrs TOKA various person-NOM world-ACC stage-LOC
katsuyakushiteirassharu n desu yo.
play.a.role:do:STA:HON NML COP PP
let: 'Various people such as Mrs Sawanobori Midori are playing an active role on the world stage.'
id: 'The whole world is the stage of various people such as Mrs Sawanobori Midori.'

While the explanations we have proposed are corroborated by linguistic data, we firmly believe that this correlation between animacy/definiteness and categorization processes calls for further investigations, since it could constitute another confirmation of the close interaction between the linguistic level and the cognitive level.

4.4 CONTEXT DEPENDENCE AND CATEGORY CLUES

Following the definition provided at the beginning of this chapter (cf. section 4.1), we have conceptualized non-lexicalized categories as those occurrences in which the speaker constructs categories relying solely on an exemplar-driven inferential path, that is, a process through which categories are abstracted from specific exemplars in context. Therefore, it may seem that beside the mentioned examples, the hearer is not provided with any other cues as to correctly infer the category. However, this perspective suffers from several shortcomings, most notably the fact that people do not learn and elaborate concepts in

isolation (cf. Barsalou 1982, Murphy 2002). Therefore, while studying labels and examples in (partial) isolation is useful to analyse in depth their linguistic properties, we should always consider the key role of the context in which they are processed.

We have already seen that the context may influence the identification of the common property of the category members. In fact, the associative inference (through which categories are constructed) is anchored in and depends 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" (Mauri 2016: 6). It follows that, in some cases, without having access to the context, it may be almost impossible to understand the category. Consider the following example:

(4.34) *Hoteru-de*   *yaru*   *geki-toka*   *jikan-o*   *kakete*   *isshōkenmei*
       hotel-LOC   do   drama-TOKA time-ACC   take:GRD   very hard
       yōishita   *noni...*
       preparation:do:PAST   and yet
       'We prepared very hard spending time over the drama to play at the hotel and such, and yet...'

In the utterance above, the speaker provides a single example: *hoteru de yaru geki (toka)* "drama to play at the hotel (and such)". In this case, it would be impossible to infer correctly the category without knowing that 1) the speaker is making reference to a cancelled trip to Disneyland, and 2) Japanese schoolchildren often prepare some activities to do during school trips (e.g., little dramas to play and such). In other words, the access to the context is essential to understand the category.

However, is this enough to consider the entire context as a cue towards the category? For instance, it is not an explicit cue since it does not make overt and clear reference to the category or to its defining property. In a sense, it seems to play more a passive role, like a cognitive background to which examples and labels are compared to infer correctly the category. On the contrary, examples and labels show an active role, since they provide explicit (semantic) hints to identify the category and its defining property. Therefore, the question should be: beyond representing the cognitive background where categories are elaborated, can the context (or even better, contextual elements) play an active role in the inference of categories?

Consider this example:

(4.35) *Takuhaibin-ga   kuru   hi-toka,   tomodachi-ga   asobinikuru   hi-toka.*

courier-NOM   come day-TOKA   friend-NOM   visit   day-TOKA

Sōiu   hi   wa,   genkan-no   nioi   wa   yōchūishitai

such   day   TOP   entrance-GEN   smell   TOP   need.special.attention:do:DES

basho desu.   Gaiki-no   nioi-ni nareta   hito   ga,

place COP:POL   open.air-GEN   smell-DAT   be.used:PAST people NOM

saisho-ni   haitte   kuru   basho da   kara   desu.

first-ADV   enter:GRD   come place COP   because   COP:POL

'For example, the day the courier comes, the day friends visit you, and so on. In such days, the smell of the entrance is the place you want to be careful about. Because it is the place people who are used to the smell of open air first enter.

The paragraph above is part of an article about how to eliminate the bad smell of pets. This is the context that acts as a background and that the reader should consider in order to correctly identify the common property. However, in this case, the context also provides semantic clues towards identification of the defining property of the category. For instance, the expression *gaiki no nioi ni nareta hito* "people who are used to the smell of open air" suggests what kind of people participate in this relevant category of situations. Moreover, *genkan* "entrance" and *haitte kuru* "come entering" suggest the common place and common action of the category of situations. When the reader puts together all these cues, the identification of the relevant category gets easier. The existence of instances like this one indicate that the context may influence the inferential process at different levels.

As we have seen in section 2.3.4, the context is not merely an inert setting, but it is a multi-dimensional element encompassing different components, such as the shared knowledge of the participants and their interpersonal relations. Moreover, it also encompasses pure linguistic components, such as the preceding discourse, the immediately adjacent co-text, and the type of speech (Croft and Cruse 2004).

All these components can actively work to influence the construction of the category by providing contextual clues that direct the inferential process. Nevertheless, as linguists, we should focus our investigation on the type of context that pertains the most to our field of study, that is, the linguistic context (Croft and Cruse 2004: 102). For instance, as shown in the example in (4.35), the immediately adjacent co-text may provide other hints to help the hearer to identify correctly the category. In particular, in our corpus, two types of hits are

attested: 1) linguistic elements that establish a contrast with the members of the category and thus by comparison help to identify the common property of the examples, and 2) linguistic elements that are not proper labels (according to the working definition we established at the beginning of our study, cf. section 2.2.1), but nonetheless may provide semantic specification about the category.

Consider the following example:

(4.36) *Haisen*   *wa*   *narubeku*   *saitankyori-de,*   *ippitsugaki-de*
      wiring   TOP   as much as possible   shortest.distance-LOC   single.stroke-STR
      okonau   *yōnishiyou.*   *Mudana ukai-ya*   *edawakare*   *nado*   *wa*
      perform   to.be.sure.of:VOL   useless detour-YA   branching   NADO   TOP
      shingō-ni   *akueikyō-ga*   *aru*   *node*   *kyokuryoku*
      signal-LOC   negative.effect-NOM   AUX   because   as much as possible
      yokeru beki   *da.*
      avoid   IMP   COP.
      'Make sure to perform the wiring at the shortest distance and in one single stroke. Useless detour, branching and such should be avoided as much as possible because there are negative effects in the signal.'

In the sentence above, the explicit reference to the category is based on a list of examples: *mudana ukai ya edawakare nado* "useless detours, branching and such". Here, the author wants to highlight some actions and measures that people may take during the wiring installation that can have negative effects to the entire system. It is noteworthy that in the immediately previous sentence, the author presents the correct steps to take as to install correctly the wiring system. In this sense, the category represents the exact opposite and the contrast thus created actually helps the hearer to better understand the category itself (i.e., the category encompasses all actions and measures that are the opposite of those mentioned above).

Here are other examples:

(4.37) *Kenkō-ga*   *daiichi.*
      health-NOM   first
      Kega-ya   *byōki-o*   *shinai yōni*   *onegai*   *shimashita.*
      injury-YA   illness-ACC   do:NEG for   request   do:POL:PAST
      'Health [comes] first. I asked not to get hurt or sick (or something similar).'

In the utterance above, the speaker describes her wishes for the next year. She explains that even if she is having some troubles with money, she thinks that health always comes first. Therefore, she wished to avoid any kind of health-related issues (e.g., injuries, illnesses and so on). In this case, *kenkō* "health" facilitates the interpretation of the following category, whose defining property is to encompass everything opposite to the notion of being healthy.

(4.38) *Taitei   no   jōhō   wa   intānetto-o   kensakusureba   shunji-ni*
most   LK   information   TOP   internet-ACC searching:do:COND in a moment
kotae-ga   mitsukaru   yōninari,   hon-de   shirabetari,   hito-ni
answer-NOM   be found   become:GRD   book-LOC   check:TARI   people-DAT
tazunetari   suru   koto   wa   hettekiteimasu.
ask:TARI   do   NML   TOP   decrease:STA:POL
'As for most information, if you search on the internet, you get to find the answer in a moment, and things like searching in books or asking to people are getting less frequent.'

Here the author refers to a context-relevant category by means of a list of examples: *hon de shirabetari, hito ni tazunetari suru* "checking on books, asking to people". The process of comparison that leads to the identification of a common property (and thus to the category) is facilitated by the contrast between the category itself and an information provided in the co-text, namely the expression *intānetto o kensakusur[u]* "searching on the internet". Through this opposition, the reader can identify the members of the list construction as traditional ways to access information in contrast to the modern way, that is, using the internet. This ultimately smooths the inferential process towards the construction of the category. Therefore, even without an explicit category label, we can see that there are still other elements in the linguistic context that actually help the identification of the defining property of the category and therefore the construction of the category itself.

In other cases, the speaker can provide not only an element of contrast, but also some other explicit information about the category itself, thus making the opposition even more evident. Consider the following example:

(4.39) *Kaigai-ni   ikanai   hito-ya   sumanai   hito-ni wa   amari*
abroad-LOC   go:NEG   people-and   live:NEG   people-DAT   TOP   not much
ennonai   biza.   Daga, ryūgaku-ya   shūshoku   nado   de
unrelated   viza   but   studying-YA   finding a job NADO   CAUS
tan-chōkikan,   kaigai-ni   sumu baai, biza-ga   hitsuyō   tonaru.
short-long.period   abroad-LOC   live   case   visa-NOM   necessary   become

'The visa is not related to people who don't go abroad or don't live abroad. But, if you live abroad for a short or long period of time for studying abroad, looking for a job and so on, a visa becomes necessary.'

In the sentence above the contrast between not living abroad (*kaigai ni […] sumanai hito* "people who do not live abroad") and actually living abroad (*kaigai ni sumu* "to live abroad") facilitates the identification of those activities that discriminate between people who do not actively live in a place (e.g., tourists) and people who are effective residents and part of the society of that specific country. Therefore, the activities that characterized the second group are, for example, studying at the university, finding a job, working, attending to other types of schools, etc., and they all require a visa.

Contrary to (4.37) and (4.38), here the author does not just provide an element of contrast, but also an explicit hint about the category, i.e., *kaigai ni sumu* "to live abroad". While it would be a stretch to consider it as a proper category label (following the working definition provided in section 2.1.1), at the same time it is true that it helps the inferential process because it suggests the defining property of the category. More specifically, we can say that it works as a category clue (cf. Barotto and Mauri 2016), that is, an optional linguistic element that guides further the inferential process towards the identification of the relevant property shared by the examples. Unlike category labels that make an explicit reference to a category, clues do not automatically involve the identification of a category (therefore, they cannot work in isolation without other strategies), since they perform an ancillary function.

(4.40) *Tokuni   ōkuno hito-ga   muzukashiku kanjiru no   wa   jibun-no   ie*
in particular   many people-NOM   difficult:ADV   feel   NML   TOP   oneself-GEN house
no   naka-de   no   satsuei   dewanaidarouka.
LK   inside-LOC   LK   photographing   perhaps
Utsushitakunai   mono-ga   utsutteshimattari,   hikari-ga   fushizen dattari
photograph:DES:NEG thing-NOM   photograph:ASP:TARI light-NOM   unnatural COP:TARI
kage-ni   natteshimattari.
shadow-DAT   become:ASP:TARI
'In particular, I think that many people feel difficult to take picture inside their own houses. Ending up photography things that they do not want to photograph, lights are unnatural or end up becoming dark.'

Here, the author refers to a category of problems relating to photography: 1) *utsushitakunai mono ga utsutteshimattari* "ending up taking picture of things that do not want

to photograph", and 2) having trouble with the lights, as *hikari ga fushizen dattari* "lights are unnatural" or *kage ni natteshimattari* "end up becoming dark". The identification of the defining property of the category is facilitated by the statement in the preceding sentence about the fact that people may feel that taking picture inside of their own houses is difficult. Due to this category clue, the reader can infer that the following list of issues (and thus the category) represents a specific type of difficulties that people face while taking picture indoor.

It is interesting to note that, in isolation, this statement would not explicitly invoke the identification of a precise category of issues. However, in this context, that is, followed by a list of concrete examples, it acts as a category clue.

In some cases, contextual clues can also be used to further specify category already designated by means of labels as to make more clear the property shared by the examples.

(4.41) *IPX 5/ IPX 7*   *no*   *bōsui*   *seinō.*
        IPX 5 / IPX 7   LK   waterproof   ability
        Kitchin-ya   *basurūmu*   *nado,*   *tsukau basho-o*   *kini-sezuni*   *mēru-ya*
        kitchen-YA   bathroom   NADO   use   place-ACC   mind-witout   mail-YA
        wansegu-ga   *tanoshimeru*   *yōninatteiru.*
        1-SEG-NOM   be able to enjoy   become:STA
        'Waterproof ability IPX5/IPX7. You will be able to enjoy email and 1 SEG[47] (and so on) without having to worry about the place of use such as the kitchen or the bathroom.'

In the sentence above, category members are not only defined as *tsukau basho* "places of usage" (i.e., the category label) of the described mobile phone, but also as related to the notion of *bōsui* "waterproof". Therefore, they are places of usage and at the same time they are also related to water, like a bathroom or a kitchen.

In our corpus, this last situation is far less frequent (12 occurrences) than category clues used to specify non-lexicalized categories (a total of 74 occurrences). This is mainly due to the fact category labels themselves are a special type of category clues (cf. section 3.6), so in most cases the speaker can refer to the category by means of them, without needing other specification. In case further specification is needed, the speaker may simply choose to provide a more complex and detailed label by adding some linguistic adjuncts.

While category clues highlight the fundamental role of context in a very straightforward way, it is still noteworthy that they are not necessary for indexical categorization and do not

---

[47] 1 SEG is a mobile terrestrial digital audio/video and data broadcasting service in Japan.

even occur in the majority of instances of non-lexicalized categories. Figure 4.8 shows the frequency of three types of constructions: 1) lexicalized categories, 2) non-lexicalized categories with at least one category clue, 3) non-lexicalized categories with no category clue.
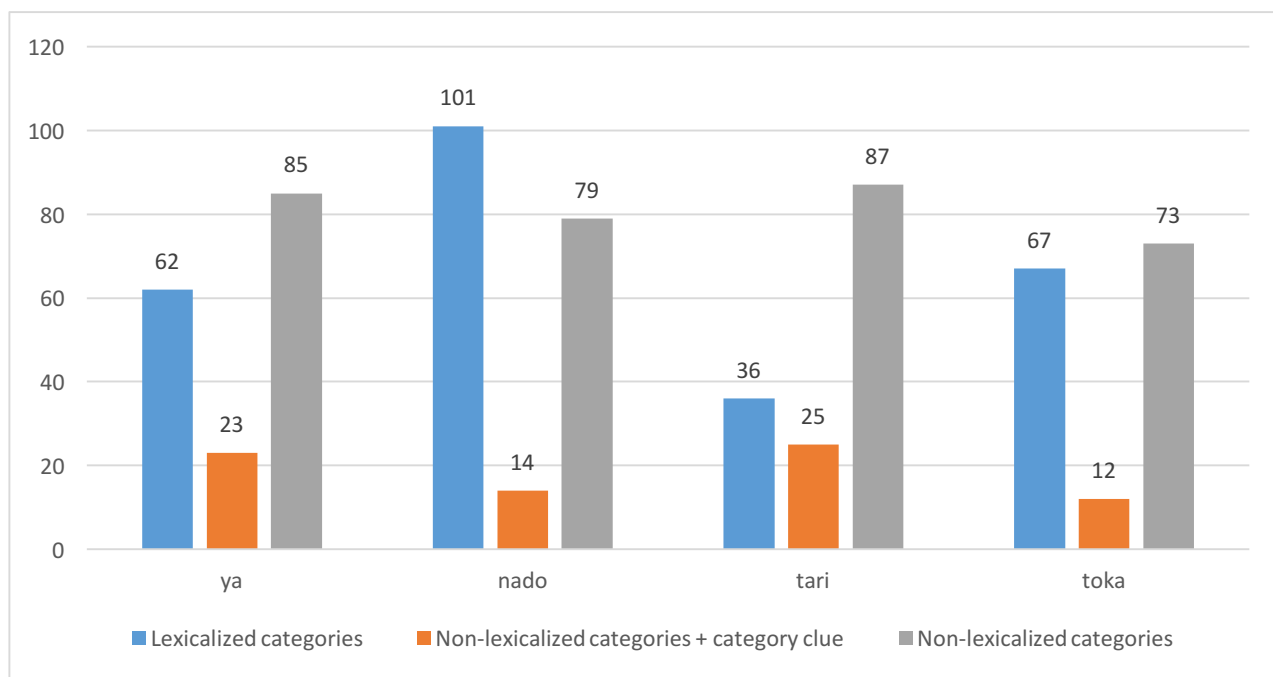
Figure 4.8: Distribution of lexicalized categories, non-lexicalized categories with category clues, non-lexicalized categories.

The first point of interest is the frequency of lexicalized categories compared to that of non-lexicalized categories with at least one category clue. It emerges that between a category label and a category clue, usually speakers tend to prefer the former, probably because 1) they make explicit reference to a category and 2) often they are much more specific in identifying the common property of the examples, thus they can direct more precisely the inferential process.

The second point of interest concerns the distribution of these pattern in correlation with *tari*. In this case, the frequency of lexicalized categories compared to that of non-lexicalized categories with at least one category clue, is very similar. This figure is especially interesting given the fact that *tari* can be used only with examples expressed by verbal phrases. In chapter 3, we have seen that the frequency of category labels is lower whenever the speaker provides examples encoded by verbal phrases. It follows that, in some cases, speakers may still want to give some semantic specification about the category, but they find it easier to

provide other types of categories clues to direct the inferential process, rather than providing a proper category label (cf. section 3.3).

The third point of interest is that, in our corpus, quite frequently speakers provide only examples, without any other type of clues (category labels included). This appears even more evident in Figure 4.9 below.
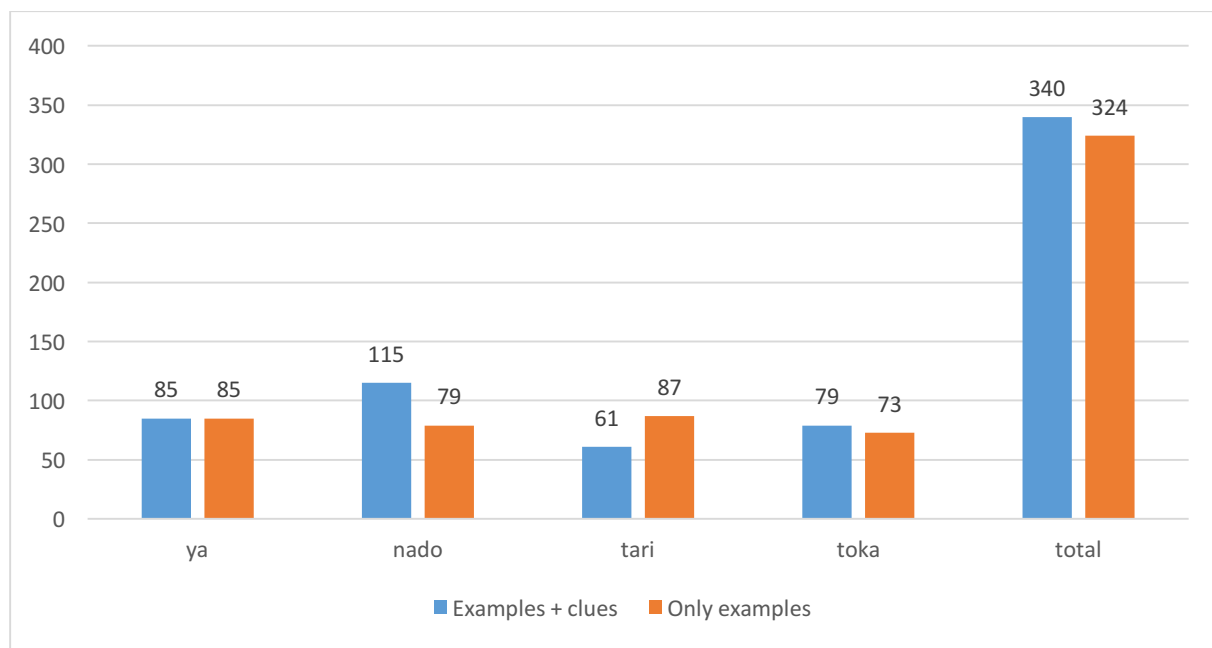
Figure 4.9: Constructions exhibiting only examples vs. constructions in combination with category clues

This means that beyond the usefulness of category clues in directing the inference, it is indeed possible (and also quite frequent, especially with regards to some types of categories, cf. chapter 3) to communicate context-relevant categories only by means of concrete examples.

To conclude, what emerges from this final section is that the inference of context-relevant categories is not merely the result of the comparison among the examples in isolation, but a much more dynamic process where also other (linguistic) elements in the co-text can make a significant contribution to the inferential process. This ultimately confirms the centrality of context and the fact that it should not be considered merely as some inert background to the cognitive process. Moreover, categorization does not simply involve the abstract manipulation of data, but it is something that takes place in a lifeworld consisting of social expectations, actions and goals. In other words, it is "*something we do*, in talk, in order to

accomplish social actions" (Edwards 1991: 517). This means that categorization not only cannot be detached from the context, but can also be actively driven by it.

5. EXEMPLIFICATION WITHOUT CATEGORIZATION: THE PRAGMATIC LEVEL

5.1 TOWARD THE CONSTRUCTION OF THE FUZZY EFFECT

Esemplificare, nel costruire un testo, vuol dire fornire all'interlocutore dei casi particolari (tra i molti altri possibili), cioè uno o più elementi di un insieme più vasto (dato o potenziale) di entità, di attività, di situazioni, di problemi, ecc.

(Manzotti, 1998: 108)

Exemplification implies the generalization of a single instance. It signals that the following or preceding elements, that is, the examples, should be conceived as representative of a wider set. As we will see in this chapter, drawing out this characterizing functional core, exemplifying strategies can perform a series of different communicative functions.

As it has been discussed in the previous chapters, one of these functions is the on-line construction of contextually relevant categories. In this sense, exemplification is a linguistic strategy chosen by speakers whenever the label of the category is regarded as insufficient for the communicative purpose (cf. chapter 3) or whenever the construction of a label requires a cognitive effort that the speaker decides not to exert (cf. chapter 4).

Therefore, whenever examples are used to categorize, they should be conceived as arrows to the category they represent, rather than as bearing an independent and discourse relevant reference. Consider the following:

(5.1)   *Nemuku   naranai   niwa   koohii toka   nomimasu.*
        sleepy   become:NEG in order   coffee TOKA   drink:POL
        "In order to stay awake, I drink coffee and so on".

In the sentence above, the example *koohii* "coffee" serves the only purpose of identifying a category of similar items "caffeine drinks that help people to stay awake". Therefore, the actual "what-is-said" part of the utterance is the category 'caffeine drinks that help people to stay awake', not the single mentioned item 'coffee'.
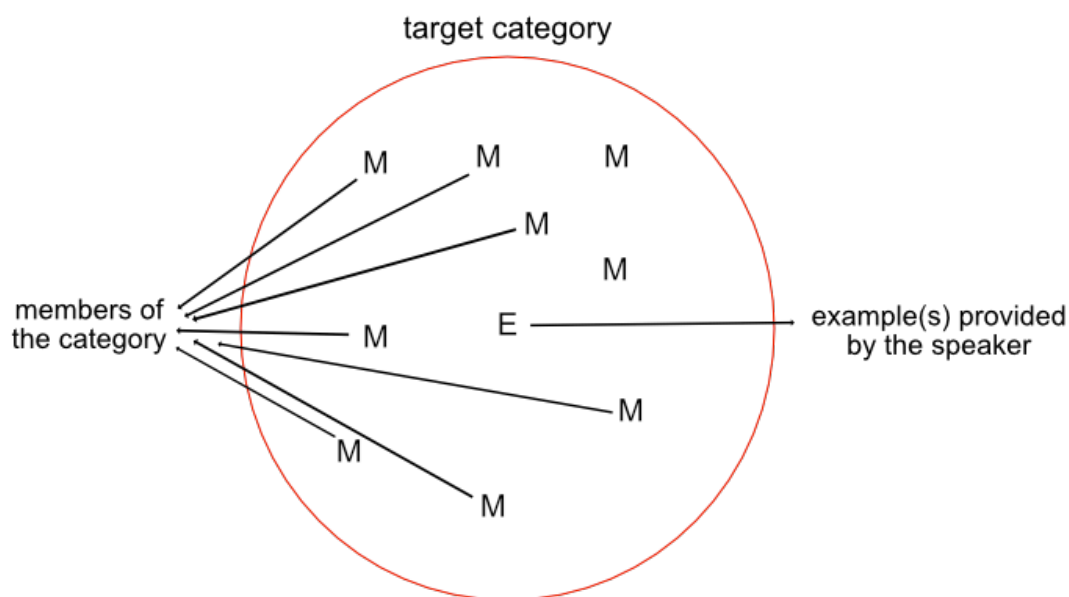
Figure 5.1: The roles of the example and the set in the creation of categories

Generally speaking, we may say that whenever exemplification is used to create a category, the real focus of the speaker is the target category (in red in Figure 5.1), and not the element marked as an example (in black), which has to be interpreted as being merely a potential exemplar of the higher-level category, chosen among many other members of the same category (in black as well).

In this regard, Manzotti (1998: 121) notes that one of the components of the notion of exemplification is the interchangeability of particular cases: examples illustrate and specify the abstract by giving a more concrete and specific item out of a set of other equally illustrative items. The further implication is that examples are fundamentally arbitrary choices, since the selected item is just one of a number of possible other examples.

It follows that, in different situations, speakers may use this implication of the presence of a wider set just to explicitly open up "a paradigm and thereby show that other neighbouring expressions would be equally possible, neighbouring points on a numerical or chronological scale as well as semantically related lexical items" (Mihatsch, 2010b: 108). Consider the following:

(5.2) *Kōkō*   *3 nen toka, benkyō*   *shitenakatta*   *wa*
       high-school   3 year TOKA   study   do:STA:NEG:PAST   PP
       'Around my third year of high school, I wasn't study.'

Above, the speaker makes an informed guess about the period of time when she was not very studious. The important information is not the exact period of time, but the fact that during high-school she had difficulties with her studies. Here, the exemplifying strategy *toka* denotes a more or less symmetrical interval around the exemplar item (cf. Jucker et ali. 2003: 1758). It signals an open list of possible candidates other than the one mentioned (i.e., *kōkō 3 nen* "third year of high school"), which represents just a possibility among many others, thus indicating "that there is a slight discrepancy between the linguistically encoded concept and the concept that the hearer is expected to pragmatically infer" (Ghezzi 2013: 163).

In both (5.1) and (5.2) the basic function of the exemplifying strategy *toka* is to widen the extension of the modified expression (Mihatsch, 2010b). Nevertheless, while in (5.1) this extension allows to make reference directly to the target category from which the element has been taken, in (5.2) *toka* signals that the lexical choice is less than perfect, that is, it is less than prototypical. We may therefore argue that in this second case the actual role of the exemplifying strategy is to add *fuzziness* (cf. Lakoff 1973, Prince et ali. 1982), more precisely 'fuzziness within the propositional content proper'. This is indeed possible because of the set di others potentially available choices implied by the nature itself of exemplification. Therefore, we may add that exemplifying strategies can open up the paradigmatic axis stressing that the mentioned element is selected from a larger set of alternatives, and because of this, it should not be taken as certain.

The widening process can also be employed to hedge strong assertions representing the speaker's point of view, which are relativized by being presented as arbitrary choices among many potential others (cf. Mihatsch 2010a).

(5.3)   *Koohii toka   nomitai.*
        coffee TOKA   drink:DES
        'I want to have a coffee (or something).'

In the sentence above, on the one hand, the speaker presents the mentioned element (i.e., *koohii* "coffee") as a demand, thereby directly affirming his personal desire. On the other hand, he downgrades his directive suggesting the presence of wider set of alternatives beyond the one mentioned, regarded as an example of what the speaker would like. Therefore, here, the presence of *toka* urges to conceive the modified element (i.e., *koohii* "coffee") just as an arbitrary choice among many potential others, as to attenuate the strength of the speaker's assertion. In this case, the exemplifying strategy does not operate

on the propositional and semantic level as in (5.2), rather it operates at the illocutionary level, hedging the illocutionary force of the speech act and the speaker's commitment towards the utterance and thus establishing a 'fuzziness in the relationship between the propositional content and the speaker' (Prince et al. 1982: 85).

At this point we should go back to our first claim about the definition of exemplification, and therefore, analyse it in the light of what have been pointed out right above. Exemplification implies the presence of two basic elements: 1) the single instance, that is, the example, and 2) the wider set of equally representative instances, which represents the generalization of the provided example. These two elements represent the conceptual core of exemplification, and its cognitive and functional basis. Depending on the communicative function, the speaker may use the relationship between these two elements to achieve different effects.

In one case, e.g., (5.1), the speaker may use the particular case(s) just as an arrow to the set (or better, the category), which is the real focus of the discourse. In other words, the examples are only tools to achieve the categorization process. However, in (5.2) and (5.3), the relationship between the single instance and the set changes noticeably. In (5.2) the purpose of the set is to allow the speaker to widen the range of possible values in order to loosen the category boundaries and imply a discrepancy between the linguistic choice of the speaker to represent a concept and the concept itself (cf. Mihatsch, 2010b, Ghezzi 2013). In (5.3) the purpose of the set is to limit the assertive force and the speaker's commitment towards the utterance.

In both these last cases, the set is not the real focus of the speaker anymore. On the contrary it is merely a tool to achieve a communicative effect. More precisely, the implication of a set of arbitrary choices results in a fuzzy effect, that is, the creation of intentional fuzziness to being less precise (i.e., semantic approximation) or to be less assertive (i.e., pragmatic hedging). In other words, there is exemplification without categorization.

Nevertheless, even between these two situations, the set assumes slightly different roles. In (5.2) the speaker is actually taking into consideration other potential values in the set, as shown in the Figure 5.2 below:
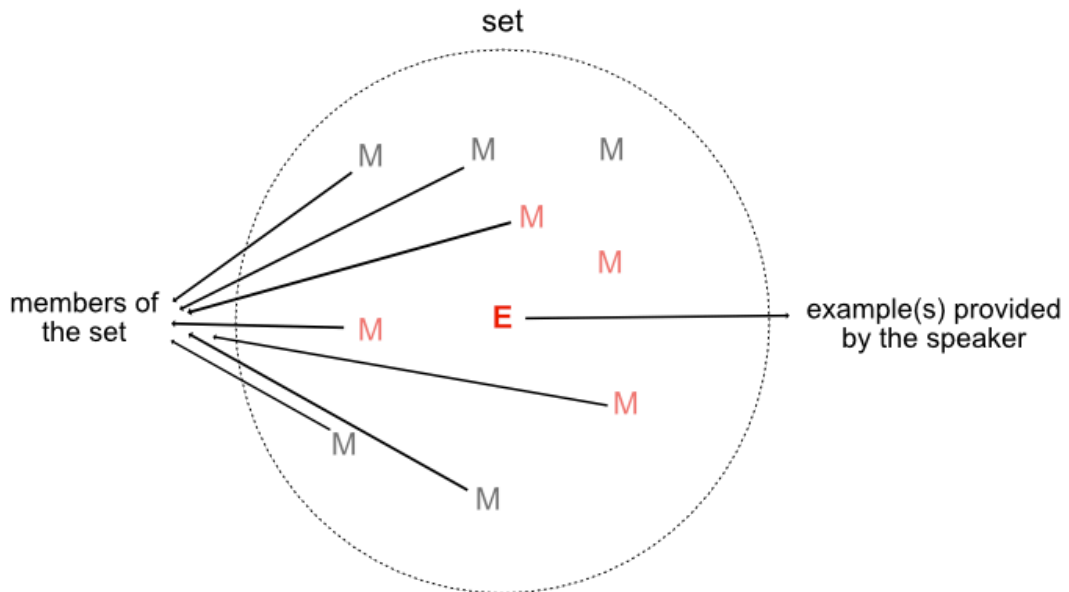
Figure 5.2: The roles of the example in the expression of semantic approximation

On the contrary, in (5.3) the speaker is not taking into consideration other potential values. In fact, her only interest is to refer to her desire about drinking coffee; however, she chooses to present it as a possible member of a set of similar drinks in order to make the utterance less strong and direct. We can schematize this process as follows:
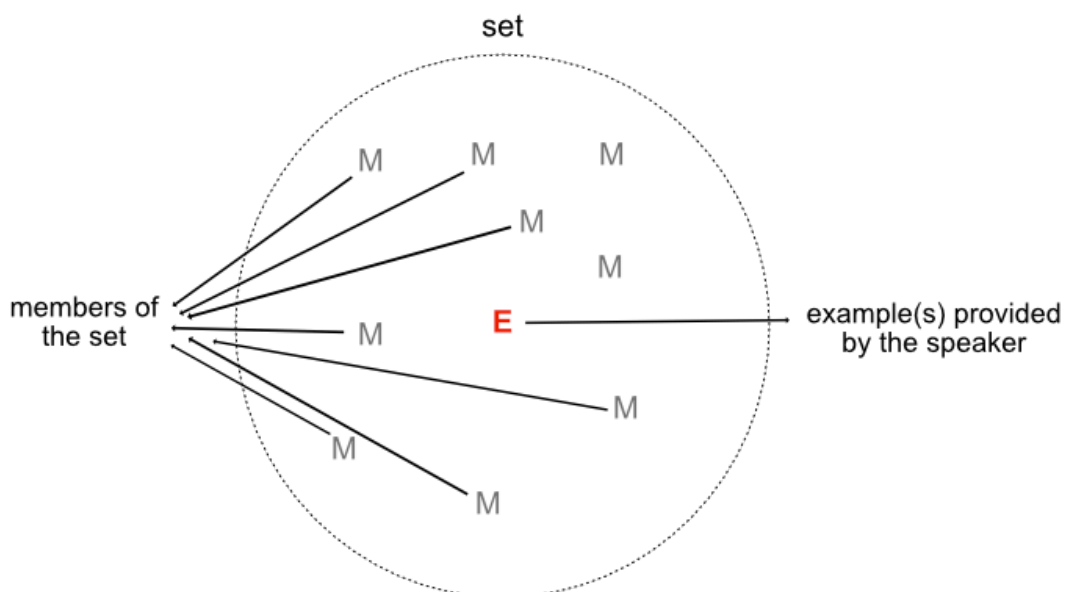
Figure 5.3: The roles of the example in the expression of pragmatic hedging

When exemplification is used to limit the assertiveness of the utterance, the real focus of the speaker is just the mentioned item/phrase marked as an example (in red in Figure 5.3),

while the other potentially available choices and the set they create (in black) serve the only purpose of creating a fuzzy background.

Therefore, to summarize, exemplification implies the presence of a set of items. This set can be used by the speaker to trigger inferential and abstraction processes, that lead to the identification and the online construction of a context-dependent category (cf. chapters 3 and 4). Otherwise, this set can also be used to establish a fuzzy background around the mentioned item (that is, the example) in order to achieve different hedging functions (cf. Kaltenböck et ali. 2010), such as semantic approximation like in (5.2) and pragmatic hedging like in (5.3). These two main cognitive processes can be schematized as follows:
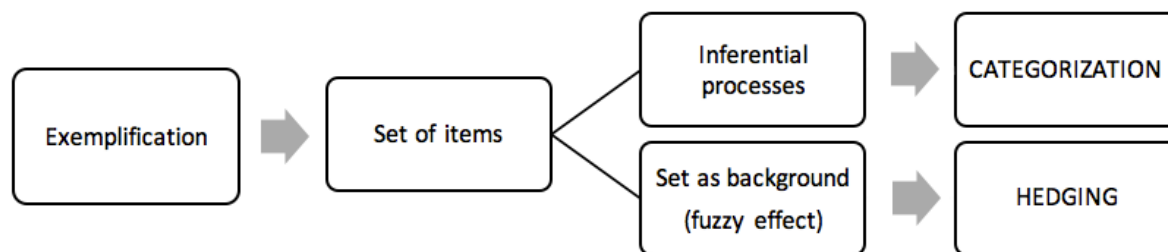
Figure 5.4: The process of exemplification and its communicative functions

Therefore, to a more cognitive function (i.e., categorization) we may add two other functions that operate at the semantic and pragmatic levels. In the following sections we will focus our analysis on the use of Japanese exemplifying constructions as hedging devices, that is, linguistic devices that can signal "a lack of commitment to either the full semantic membership of an expression or to the full commitment of the speech act being conveyed" (Fraser, 2010: 22).

5.2 EXEMPLIFYING CONSTRUCTIONS AS HEDGING STRATEGIES

As it has been underlined in previous section, exemplification can be used to stress the fact that the mentioned item should be conceived merely as an alternative chosen from a larger set of possibilities. Therefore, in these cases, the function of exemplifying strategies is not category implication, but hedging (cf. Fraser 2010, Kaltenböck *et al.* 2010). Speakers may deliberately introduce personal tenets, considerations, requests and other potential face-threatening acts as examples not in order to establish a category, but "as a discourse strategy that reduces the force of truth of an utterance and thus reduces the risk a speaker

runs when uttering a strong or firm assertion or other speech act" (Kaltenböck *et al.* 2010: 1). Consider the following exchange:

(5.4) A: *Kanreki   no oiwai ni wa, nani ga ii deshou ka?*
             60th birthday LK HON-gift DAT TOP what NOM good MOD PP
             'I wonder what would be good as a gift for a 60th birthday'

       B: (a)   *Udedokei toka dō desu ka?*
                   wrist-watch TOKA how COP PP
                   'How about a wrist watch?'

       B: (b)   *Udedokei wa dōdesu ka?*
                   wrist-watch TOP how COP PP
                   'How about a wrist watch?'

In (5.4a), the topic *udedokei* "wrist watch" is marked by *toka,* while in (5.4b), it is marked by *wa*. This difference in the topic marking exhibits consequences in the degree of specification of the mentioned item and thus in the way it should be conceived in the hearer's mind (cf. Suzuki 1998). While in (5.4b) *wa* specified one certain item (that is, the wrist watch), in (5.4a) *toka* indicates a lack of referentiality, because the item is introduced just as one possible alternative among others. By using *toka* instead of *wa*, the speaker wants to avoid the specification in order to attenuate the his or her commitment (that is, 'I am giving you my suggestion, I am not committed to it'). Despite both (5.4a) and (5.4b) being suggestions, (5.4a) results even more fuzzy and therefore less direct than (5.4b).

(5.5) A: *Kare, osoi   ne.*
             he   late   PP
             'He is late!'

       B: *Nebōshita   toka...*
             oversleep:DO:PAST   TOKA
             'Maybe he overslept...'

In the exchange above, the speaker B does not want to commit to the possibility that the person is actually late because of oversleeping, there may be other reasons. The speaker is not interested in creating a category, the purpose is to suggest probable possibilities or explanations in an indirect way. Therefore, the function performed by *toka* is not exemplifying to make reference to a category, but hedging to attenuate the commitment to the truth of the propositional content conveyed.

In the last decades, the concept of "hedging" has developed from narrowly semantic concept to a wider pragmatic one. The term was originally introduced by Lakoff (1973) while exploiting the developments of the prototype theory in cognitive psychology (cf. Rosch 1973) as a purely semantic concept, focusing on the capability of some linguistic elements (e.g., *sort of*) of signalling different degree of category membership. Lakoff loosely defined hedges as linguistic expressions "whose job is to make things fuzzier or less fuzzy" (1973:195). Since all the examples examined by Lakoff involved predicate adjectives or predicate nominals in declarative sentences, Fraser (1975) proposes to refer to this type of hedges as 'propositional hedging', "since it is the truth value of the whole proposition that is affected" (Fraser 2010: 17). For instance, in a sentence like *A penguin is sort of a bird*, the non-prototypical concept 'penguin' can be included in the category 'bird' due to the use of the hedge *sort of*. Therefore, hedges such as *sort of* affect the truth value of the propositional content.

Despite his focus being mainly semantic, Lakoff also addresses the pragmatic value of hedges, noting that their interpretation is context-dependent (1973: 213). This observation has been investigated further by Fraser (1975) and by Brown and Levinson (1978, 1987) to include expressions that modify the force of the speech act. This led to the identification of a second type of hedging, namely the "speech-act hedging" (Fraser 2010), which encompasses devices that attenuate the strength of the speech act and, more generally, the speaker's commitment towards the utterance. Brown and Levinson (1978: 169-176) considered hedges as linguistic expression that indicated primarily that the speaker is not adhering to one of the Grice's maxims (1975) and investigated their use mainly as means of negative politeness (Brown and Levinson 1987: 116).

Later studies (Prince *et al*. 1982, Hübler 1983) confirmed and further investigated – albeit changing the terminology – this distinction between two types of hedging or 'fuzziness', one type that affects the truth condition of the proposition and a second type that serves as an index of the speaker's commitment to the truth of the propositional content conveyed. For instance, Prince *et al*. (1982) refer to the former type as 'approximators', which is further subdivided into 'adaptors' (e.g., *sort of*) and 'rounders' (used to round measurement e.g., *about, around*); while the latter type is referred to as shields which "affect the pragmatics by inducing implicatures conveying markedness with respect to speaker commitment" (Prince *et al.* 1982: 86) and are further subdivided into 'plausibility shields' (expressions that relate doubt) and 'attribution shields' (attribute the responsibility of the message to someone other than the speaker). Caffi (2001, 2007) further developed this system while writing about the

notion of *mitigation*. She proposed a tripartite model identifying three different components of the utterance on which mitigation can operate and three corresponding classes of strategies, i.e. *bushes* (on the propositional content), *hedges* (on the illocution) and *shields* (on the deictic origin of the utterance).

While these classificatory models are necessary heuristic devices to better understand the phenomenon of hedging, it is important to note that "actual language use individual linguistic items may prove difficult to pigeon-hole, often as a result of their multifunctionality" (Kaltenböck *et al*. 2010: 6). As it will be shown in this sections, this clarification applies also to the strategies here discussed.

Nevertheless, the general picture that emerges is one in which we can distinguish two macro-types of hedging: 1) strategies that affect the truth value of the propositional content and 2) strategies that affect the degree of commitment of the speaker to the truth of the propositional content conveyed. The examples of *toka* in (5.4) and (5.5) are instances of the latter type.

Beyond few language-specific studies (e.g., Mihatsch 2010b, Ghezzi 2013, Voghera 2013), exemplifying constructions have not been particularly studied as hedges. Nevertheless, in those few studies, they have proven to be versatile strategies. For instance, in her study on Italian vagueness markers, Ghezzi (2013) highlights how some exemplifying strategies can function as approximating devices (e.g., the parenthetical expression *metti*) operating at the propositional level, while others operate at the illocutionary level (e.g., *per esempio* "for example) to weaken the speaker's commitment. Furthermore, others can even perform both functions (e.g., the taxonomic noun *tipo*).

Indeed, individual linguistic strategies may behave differently, performing some functions instead of others, particularly in relation to their diachronic paths. It follows that the picture regarding Japanese exemplifying strategies should probably be considered peculiar of these specific strategies. However, we hope that providing another insight into the usage of exemplifying constructions as hedging devices will pave the way for future investigation.

Before examining the Japanese exemplifying constructions here under study, it is important to clarify some methodological issues. In previous section, we have seen that the construction of categories and the hedging functions are closely related. For this reason, we may hypothesize that not all cases are so clear-cut. It follows that it is necessary to establish and monitor a number of parameters with the specific purpose of identifying instances of the former and instances of the latter.

Consider the followings:

(5.6) (a) *Kinenhin   nani-ga   hoshii? Udedokei   toka dō   kana?*
         souveir   what-NOM   DES   wrist-watch   TOKA how   MOD
         'Do you want something as a souvenir? What about a wrist watch?'

     (b) *Kinenhin nani-ga   hoshii?   Nekkuresu   toka   udedokei toka   dō kana?*
         souveir what-NOM DES   necktie   TOKA   wrist-watch TOKA how MOD
         'Do you want something as a souvenir? What about a necktie or a wrist watch (or
         something)?'

Reading the example in (5.6a), the most common interpretation would be that the speaker does not want to commit to the mentioned item by avoiding specification, therefore, the exemplifying strategy is used to hedge the speaker's commitment. On the contrary, in (5.6b), the speaker uses the exemplifying construction to make reference to a category of paradigmatic choices while suggesting possible gifts.

We argue that the different interpretations result from the different numbers of mentioned items. In fact, while in (5.6a) the speaker make reference to one single object, in (5.6b) there are two objects. In this regard, in instances like (5.6b), the identification of a common integrator (Lang 1984) or a shared Property P prevails and, along with the non-exhaustive tag *toka*, triggers the inferential processes that lead to the construction of a category 'everyday objects that are useful and thus may be good gifts'. In fact, while a reply such as 'what about sunglasses?' would be considered as not very appropriate in (5.6a), it would be much more appropriate in (5.6b), as the speaker is already making reference to a wider set of items.

Obviously, we may assume that categories can be created in discourse also for communicative purposes other than the communication of the category itself. For instances, providing a set of actual options may be considered just as a strategy to attenuate the commitment. Nevertheless, since there is always the mediation of a category, the cognitive process is not very different from the occurrences analysed in chapter 3 and chapter 4. Since in this chapter we are mainly focusing on instances of exemplification without categorization, examples like (5.6a) where it is still possible to identify an underlying categorization process have been already considered in the previous chapters of this book, and thus will not be considered again here.

While we propose that the number of examples is a parameter to identify instances of exemplification without categorization, still, as it was shown in the previous chapters, it is

indeed possible to make reference to a category using just a single exemplar as a starting point to make associative inferences, simply by drawing on the situational context. It follows that the number of examples is still not enough to distinguishes between instances of categorization and instances of hedging.

Other parameters that have been considered regard the presence of linguistic strategies employed to reduce speaker's commitment, such as modal verbs and epistemic modality devices (Fraser, 1980: 348, Pietrandrea 2005). More specifically, the potential status of contexts has been studied by monitoring the occurrence of (ir)realis devices such as directives, futures, epistemic predicates and lexical adverbs referring to (ir)realis contexts (Mauri 2008). In fact, epistemic and deontic contexts provide an ideal ground for instances of hedging, since the speakers is often compelled to attenuate his or her commitment towards the utterance.

At the discourse level, we also monitored the organization of texts and referential paths in conversation (Robert 2008), revealing the topic continuity (Givón 1983) of categories and examples. Specifically, we distinguished between cases in which the category is or becomes the topic of discourse and stays active through the subsequent text, and cases in which it is the specific example that is selected as topic. The latter case indicates that the example should not be interpreted just as a mere arrow to the category, but rather as bearing an independent (and discourse relevant) reference. It follows that the non-exhaustive tag that has been added to the example exhibits a different function than that of triggering categorization processes. In these situations, we can assume that the non-exhaustive tag actually works as a vague tag performing a hedging function. Consider the following:

(5.7) A: *Kinenhin nani-ga   hoshii? Udedokei   toka   dō   kana?*
           souveir what-NOM   DES   wrist-watch   TOKA   how   MOD
           'Do you want something as a souvenir? What about a wrist watch?'
       B: *Aaa   ii   ssu   ne,   kashio no   G-SHOCK   toka   saikō.*
           oh   good   COP   PP   Casio LK   G-SHOCK   TOKA   awesome
           'Oh, that's good! Casio G-SHOCK is awesome.'

In the exchange above, speaker B replies to the suggestion of buying a watch pointing out a popular wrist watch model, namely Casio G-SHOCK. This means that the other potential alternatives configured by the vague tag *toka* serve the only purpose of making the utterance less specific and less direct, mitigating the assertiveness of the utterance. In this

case, we identify the specific example (i.e., wrist watch) as the selected topic that stays active through the interaction.

Furthermore, the analysis of the co-text allows us to monitor also other constructs such as other politeness strategies (Brown and Levinson 1987), as they frequently occur in combination with hedging strategies. Nevertheless, on this point, it should be noticed that politeness is just one of many possible effects of a hedging act, an effect that is ultimately uncertain and not guaranteed, as there may be cases where the usage of hedging strategies can be perceived as impolite and vice versa (cf. Fraser 2010, Caffi 2007). Therefore, the relationship between politeness and hedging should not be considered as a direct implication, but more like a frequent correlation.

Finally, some occurrences have been verified by means of the questionnaire (cf. section 2.2.2 and see the questionnaire in Appendix B). Some of them were indicated as (less straightforward) instances of exemplification, therefore they have not been considered for this part of the analysis.

Now that we have established some parameters to identify instances of exemplifying constructions used as hedging strategies, a preliminary discussion about their general distribution is necessary.

The first point of interest is that in our corpus there is a very limited number of instances exemplifying constructions used as hedging strategies, compared to the number of instances of exemplifying constructions used to categorize. Table 5.1 indicates the distribution of the two functions:

Table 5.1: Distribution of discursive functions among exemplifying strategies

	Categorization	Hedging
ya	170	0
nado	194	3
tari	148	9
toka	152	32
Total	664	44

These numbers are determined, at least in part, by the nature of the data. As noted in section 2.1.1, the corpus that we used for our analysis, the Japanese section of the Leipzig Corpora Collection (LCC), consists mainly of articles from online newspapers and magazines. It follows that most of our occurrences are taken from journalistic texts where exemplification occurs mostly in the sense of providing concrete instances of an abstract

issue (cf. Zillmann and Brosius 2000), and there is also less need to use hedging strategies to attenuate the assertiveness, especially in the case of news reports describing objective facts.

To a lesser extent, some texts are 1) transcripts of interviews, 2) online message boards and forums with questions and answers from users, 3) newspapers editorials (which by design express the editor's opinions and, therefore, may need mitigation strategies to attenuate personal statements). In these latter types, exemplifying constructions performing hedging functions seems to be used with a higher frequency:

Table 5.2: Exemplifying strategies as hedging devices and types of texts

	Journalistic texts	Interviews (& similar)
nado	1	2
tari	2	7
toka	13	19
Total	16	28

This observation leads us to another factor that seems to greatly influence the distribution of functions and may also explain the limited number of exemplifying constructions as hedging devices in our corpus. Comparing our data to other quantitative studies regarding the same exemplifying strategies (cf. Taylor 2010), it clearly emerges that the frequency of hedging functions appears to be highest in (mainly informal) spoken conversation, rather than in written (mainly formal) language[48]. This observation is supported by studies on other languages, such as English (cf. DuBois 1992, Ediger 1995, Overstreet 1999) and Italian (Ghezzi 2013). Naturally, it follows that in a corpus based mainly on written newspaper articles, hedging functions may not occur very frequently. For this reason, this part of our study will be purely qualitative and not quantitative.

A second point of interest is that there are no attested occurrences of *ya* functioning as a hedge. This does to appear to be due to the overall low frequency of these functions in our corpus, since there are no studies or grammars that attest hedges functions (or more generally pragmatic functions) performed by *ya*. This lack is more likely linked to the nature of the strategy *ya* itself. In this regards, Suzuki (1998) suggests that the answer may lie in the fact that *ya* always requires at least two items. Contrary to the other strategies here

---

[48] On a wide range of features, spoken language has been shown to differ from written language (see Biber 1991, Brown and Yule 1983, Halliday 1989). In particular, remarks on the notion of intensity (Bally 1970 [1909]), which has been used to conceptualize the idea of mitigation, stress that expressive resources such as expressivity, subjectivity and affect are the distinctive features of spoken language.

analysed (i.e., *nado, toka, tari*) which can occur with one item functioning as general extender, *ya* cannot mark just one item (cf. section 1.3.2). For this reason, she speculates that "since *ya* requires the presence of more than one entity, the use of *ya* emphasizes the function of enumeration rather than the implication of lack of specification" (1998: 268). Since this lack of specification or referentiality is essential in evoking hedging functions, *ya* cannot perform any pragmatic function. This is consistent with our initial hypothesis that hedging is less likely to occur with two or more examples, as they act as a trigger to infer a common integrator or a shared property P, therefore naturally leading to abstraction processes.

As it was shown through the description of the various classification systems at the beginning of this section, not all hedging strategies are the same, as they may operate at different levels of the language. Without adopting a specific classification model, we still would like to point out that Japanese exemplifying strategies can perform a wide range of hedging operations (cf. Kaltenböck *et al.* 2010).

PROPOSITIONAL HEDGING: strategies that operate on the propositional level and affect the truth condition of the proposition conveyed. They correspond with Prince *et al.*'s (1982) approximators. In our corpus, the propositional hedging function is attested only for *toka,* as shown in (5.2), repeated here in (5.8).

(5.8)   *Kōkō*   *3 nen toka,   benkyō   shitenakatta   wa*
        high-school   3 year TOKA   study   do:STA:NEG:PAST   PP
        'Around my third year of high school, I wasn't study.'

In the utterance above, *toka* operates on the propositional content proper and contributes "to the interpretation [of the utterance] by indicating some markedness, that is, non-prototype, with respect to class membership of a particular item" (Fraser, 2010: 19).

Moreover, *toka* functions also specifically as a rounder (Prince *et al.* 1982) to approximate numbers of measurements, as shown in (5.9).

(5.9)   *100   MHz-toka-ni   kakudaishinaitoikenai.*
        100   MHz-TOKA-DAT   expansion:do:IMP
        'You have to expand to around 100 MHz.'

ILLOCUTIONARY FORCE INDICATION: strategies whose scope is to attenuate the strength of the illocutionary force of the speech act. They are comparable to Fraser's (1975) hedged performatives and partially correspond to Caffi's (2007) hedges. In our corpus, this function is attested for *nado, toka* and *tari*, as shown in the examples below:

(5.10) *Jiko-shōkai*   *nado* *dō*   *kana*
      self-introduction   NADO how   MOD
      'How about a self introduction?'

(5.11) *Koohii toka*   *nomitai.*
      coffee TOKA   drink:DES
      'I want to have a coffee (or something).'

(5.12) *Keshitari*   *shinaide*   *kudasai*
      turn.off:TARI   do:NEG:GRD   please
      'Please do not turn off the lights."

In the utterances above, exemplifying strategies are used to introduce potential face-threatening acts such as requests (5.9), desires (5.10), orders (5.11), while attenuating the illocutionary force of the utterance.

FELICITY CONDITIONS: strategies that indicate different degrees of uncertainty on the part of the speaker and operate by reducing the speaker's commitment. They are comparable to Prince *et al.*'s (1982) plausibility shields and partially correspond to Caffi's (2007) hedges. In our corpus, this function is attested for *toka* and *tari*, as shown in (5.5) repeated here in (5.13), and in (5.14).

(5.13) A: *Kare, osoi*   *ne.*
          he   late   PP
          'He is late!'
      B: *Nebōshita*   *toka...*
          oversleep:do:PAST   TOKA
          'Maybe he overslept...'

(5.14) *Kotchi*   *no*   *kata*   *ga*   *yokattari*   *shite...*
      this way   LK   way   NOM   good:TARI   do:GRD

'Maybe this way is better…'

In the utterances above, both *toka* and *tari* are used to imply some possibilities in a subtle way, that is, by implicating a level of uncertainty with respect to speaker's commitment.

Frequently, these last two types of hedging operations can co-occur and are not clearly distinguishable, to the point that, for example, in Caffi (2007) the notion of *hedges* covers both of them. For this reason, in order to calculate the frequency of all hedging strategies in our corpus avoiding arbitrary interpretations and useless distinctions, we decided to distinguish between two main macro-types of hedging (cf. Fraser 2010) depending on the level they operate: 1) propositional hedging and 2) speech-act hedging (i.e., covering both speaker commitment and indication of illocutionary force).

Table 5.3: Exemplifying strategies as propositional hedging and speech-act hedges

	Propositional hedging	**Speech-act hedging**
nado	0	3
tari	0	9
toka	21	11
Total	21	23

The high frequency of *toka* functioning as propositional hedging is not surprising, if we consider again the composition of our corpus. This is due to the fact that *toka* is frequently used as a *rounder* (cf. Prince *et al.* 1982) to approximate numbers or measurements. This is a type of approximation that can be found easily in newspaper articles, whenever the exact amount is unknown or is not completely relevant. On the contrary, speech-act hedging seems to be mainly a prerogative of spoken language (cf. Taylor 2010). Out of a total 18 occurrences of exemplifying strategies used as speech-act hedging, 12 occurrences appear in transcripts of interviews and online message boards. In this regard, the former represents a (naïve) transcription of spoken language, while the latter tends to imitate the style of spoken language. Because this type of texts (e.g., transcripts of interviews, online message boards, etc.) is not prevalent in our corpus, the overall low frequency of exemplifying construction as speech-act hedging is not surprising.

Once established that exemplifying constructions can function as hedging devices, we may wonder the reasons for using such devices in actual discursive situations. In other words, why do speakers feel compelled to signal a lack of commitment to either the full

semantic membership of an expression (propositional hedging) or to attenuate the full commitment to the force of the speech act being conveyed (speech-act hedging)? As Fraser (2010) points out, instances of hedging give rise to other discourse effects, such as vagueness, evasion, equivocation, and politeness, that can be essential for the success of communication in specific situations. In the next sections, we will investigate further this connection between hedging strategies and discourse effects, examining how Japanese exemplifying constructions may be used to achieve vagueness and politeness in discourse.

5.2.1 EXEMPLIFYING TO BE VAGUE

Vagueness is a perlocutionary effect (cf. Austin, 1962) that occurs "when the information you receive from a speaker lacks the expected precision" (Fraser 2010, 26). While vagueness is commonly considered a defect to avoid whenever possible, several scholars have pointed out that vague language can be as efficient as precise language is, praising its role in human language as "a desirable feature of natural languages" (Williamson 1994: 4869). In this regards, Channel (1994) argues that vagueness is not inherently good or bad, but its usage is judged based on appropriateness to the context. She points out that speakers tailor their language to make it suitable to the situation and the linguistic context, and one way to do this is by varying their precision and vagueness. It follows that the ability to "tune" the degree of precision according to the context and to interpret correctly vague expression are part of the speaker's communicative competence (Jucker et al., 2003).

Vagueness is thus often a deliberate choice by the speaker for a variety of reasons. For instance, the speaker does not know the precise details or does not feel it is necessary to provide all of them. Also, more precision may require more effort to retrieve information from the memory. In this regards, Channel (1994:32-33) suggests that vague expressions may be used to enable speakers to adhere to what Grice called the Co-operative Principle (1975). Grice (1975) posits that interlocutors operate by using four rules of conversation, which he calls maxims:

The maxim of Quality [be truthful according to evidence].
The maxim of Quantity [be informative but not over-informative].
The maxim of Relevance [be relevant to the conservation].
The maxim of Manner [the message should be clear, unambiguous and brief].

In this sense, vagueness should be seen as a device used by speakers to tailor their contributions "in such a way that they furnish the right amount of information for the purpose of the conversation" (Channel 1994: 173). For instance, if speakers do not know the exact details, vague strategies allow them to be truthful according to evidence available to them. Also, in casual conversation, providing the exact amount of information is not necessary and sometimes it is also considered misplaced by speakers (1994: 174). Moreover, tailoring the amount of information using vague strategies allows the speaker to indicate that the information is not very relevant, then focusing the attention towards what is considered most important in the utterance. Consider the following example:

(5.15) *Shōgakkō* *no* *koro* *toka* *wa* *hoka* *no* *kata* *mo*
elementary.school LK time TOKA TOP other LK person also
onaji *da* *to* *omou* *n* *desu* *ga,* *saishoni* *sawatta* *no* *wa*
same COP QT think NML COP but at first touch:PAST NML TOP
pokekon *desu* *ne.*
pocket computer COP:POL PP
'Around the time I was in elementary school, I think it is the same also for other people, but the thing that I touched at first was a pocket computer!'

The example above is part of an interview. The guest is a young web engineer who has been asked to recall his first encounter with the computer. He explains that his first experiences with computers happened around the time he was in elementary school, something that could be easily said for a lot of people of his age. The relevant information is that the first encounter was with a pocket computer, i.e., a small calculator-sized programmable computer available during the '80s. This information is relevant because it is less ordinary and also because he made his first attempts with computer programming by means of the pocket computer, as he describes in the next passage (*chīpuna pokekon de bēshikku o yatte*... "I did the program language BASIC with cheap pocket computer..."). Therefore, the use of the exemplifying strategy *toka* not only indicate that the period of time (i.e., when he was in elementary school) is vague, but also that it is not the most relevant part of his answer.

Generally speaking, as it was shown in the previous section, *toka* frequently functions as a prepositional hedge. More specifically, whenever *toka* is attached to a numeral, it indicates that the value is approximate. In other words, it functions as a rounder (Prince et al. 1982) or a vague additive (Channel 1994), that is, it designates "not precise numbers of quantities,

but rather intervals of numbers whose extent is apparently non exactly specified" (Channel 1994: 43). In our corpus, this usage is very frequent both in written and spoken (e.g., transcripts of interview) data.

(5.16) *Maikai,*   *1 Gbps-toka-o*   *jitsugensuru*   *niwa,* *hiroi* *taiiki-ga*
        Each time   1 Gbps-toka-ACC   implementation:do   for   wide   band-NOM

        hitsuyō-ni   *naru.*
        necessary-DAT   become

        'Each time, to implement around 1Gbps, a wide band becomes necessary.'

In the sentence above, the author of the article provides information about Docomo NTT wireless network. Here, *toka* is used to designate an interval of data transfer capacity of the wireless technology, which is around 1 gigabit per second (here abbreviated "Gbps").

(5.17) A: *Yahari, mēru* *wa*   *hinpanni*   *sōjushinsuru*   *no*   *deshou*   *ka?*
              Also,   emaiL TOP   frequently   send.and.receive:do NML   MOD   Q

              'Also, do you frequently send a receive emails?'

        B: *Heikinteki-ni*   *dato*   *10-tsū toka*   *kana.*
              average-ADV   if it is the case 10-CLF TOKA MOD.

              'On the average, I guess around 10.

On the other hand, the exchange above is part of an interview. Speaker B is asked about the average number of emails she sends and receives in a day. Since it is estimation and usually people do not keep track of the number of emails they send/receive per day, she is hesitant about the exact number. She uses *toka* to approximate the amount of emails (10 *tsū toka* "around 10") and the epistemic maker *kana* to express uncertainty. Specifically, *kana* is often used to provide uncertain information as a reply to a direct question (cf. Adachi 2002).

Other similar examples are provided in the utterance below.

(5.18) *Watashi*   *wa*   *denimu-ga*   *sutoretchi*   *nashide 26 ~ 27 kara 24* **toka**

I   TOP   jeans-NOM   stretch   without 26 - 27 from 24 TOKA

ni   *narimashita.*   *Kaigai*   *mono*   *da to*   *2*   *kara*   *0 saizu.*

DAT   become:POL:PAST   overseas   thing   if it is the case 2   from   0 size

23-Inchi **toka**   *desu.*

23 inch. TOKA COP:POL

My jeans (size) became about 24 from 26 - 27 without stretching. In case of things from overseas, from 2 to 0 size. It is about 23 inches.'

Here a girl comments her weight loss, saying that she went from 26-27 size to about 24 size[49], thus providing an approximate measurement by means of *toka*. As Taylor (2010: 131) notes regarding a similar example, although it is likely that she knows exactly how many sizes she has lost, she does not focus on the jeans size itself, probably because the relevant fact is that she has lost weight. The approximate measurement is also repeated in the next sentence regarding overseas (that is, western) sizes, i.e., *23-inchi toka desu* "it is about 23 inches".

In some cases, intentional vagueness is also the result of social norms preventing us from being too blunt. Therefore, varying the degree of vagueness contributes to sound less harsh and direct. It follows that, in some cases, vagueness and politeness simply co-occur and cannot be detached. For example:

(5.19) *Shitsureina*   *hanashi da*   *ga,*   *kono*   *rimokon*   *dake*   *miru*   *to,*   *980 en*

rude   talk   COP but   this   remote   only   look   when   980 yen

toka   *de*   *utte*   *sō*   *na…*

TOKA   STR   sell:GRD   look like   PP

'It is rude to say, but just when you look the remote, it looks like [it is] sold for about 980 yen.'

Here the author reviews a recorder whose market price is around 160 - 180 thousand yen. However, the product does not look like it is worth that much. He says that just by looking at the remote, it looks like it can be sold for about 980 yen. The author is aware that he is saying something that may sound offensive (*shitsureina hanashi da ga* "it is rude to say, but..."), therefore *toka* not only functions to provide an estimation of the value of the

---

[49] In Japan, pant sizes are generally labelled using a waist measurement, in centimetres or (as in this case) inches.

product, but also to mitigate the illocutionary force of his statement, sounding less direct than simply saying "it looks cheap".

Another communicative use of vague language identified by Channel (1994) is self-protection, that is "a safeguard against being later shown to be wrong" (1994: 188). Therefore, the speaker uses vague strategies in order to reduce the responsibility for what she is saying, which can entail the risk of being proved wrong. Consider the examples in (5.5) and (5.14), repeated here as (5.20) and (5.21):

(5.20) A: *Kare,   osoi   ne.*
            he   late   PP
            'He is late!'

      B: *Nebōshita   toka...*
            oversleep:do:PAST   TOKA
            'Maybe he overslept...'

(5.21) *Kotchi   no   kata-ga   yokattari   shite...*
       this way   LK   way-NOM   good:TARI   do:GRD
       'Maybe this way is better…'

In both cases, the speaker uses the exemplifying construction to imply some possibilities they are not sure about. Therefore, to avoid being proved wrong, they express the possibility they have in mind in a subtle and less direct way. More specifically, *toka* and *tari* signal the speaker's commitment to the truth of the proposition, by indicating different degrees of uncertainty on the part of the speaker (cf. the notion of plausibility shield in Prince et al. 1982). Moreover, in (5.21), the speaker is also providing a personal opinion, therefore he or she is even more compelled to use a hedging device not only as a safeguard, but also to avoid imposing her opinion on the hearer (cf. negative politeness strategy in Brown and Levinson 1987). Therefore, while in (5.20) the discourse effect of the hedging strategy is mainly vagueness, in (5.21), vagueness and politeness co-occur.

The use of vague language as a self-protection can apply for any type of vague expressions, but there are some situations where this operation is most needed, for example when we report the words of others. In this regard, we can include the use of *toka* attached to quoted utterances, like in *toka iu* "to say something like…" or *toka omou* "to think something like…", in order to avoid giving an exact quotation. Whenever the speaker is not able to reproduce the exact words (for example because of loss memory), he or she may

communicate the main point of the quoted utterance by attaching *toka* immediately before the verb of saying or thinking (cf. Taylor 2015, Ohori 2004). In this way, the speaker does not need to be concerned with being accurate. Consider the following example:

(5.22) *Yoku   mawarinohito-tachi-ga,   kenkō notameni   tabako-ya   sake-o*
      often   surranding.people-PL-NOM   health because of   tobacco-YA   sake-ACC

      yameta   toka   iimasu.
      quit   TOKA   say:POL

      'Often people around me say something like "I quit tobacco and sake because of health".'

Here, the speaker cannot reproduce the exact words that people around him say about quitting alcohol and cigarettes, therefore he expresses an approximation, that is, the main point of the quote by adding *toka*.

    Another example is provided below:

(5.23) *Danshi-ga   kore-o   tsuketeitara   "a-,   koitsu kanojo   iru   na"*
      guy-NOM   this-ACC   put:STA:COND ah   this guy girlfriend   AUX   PP

      toka   omotteshimai sō.
      TOKA   think:ASP   EVID

      'If a guy put this [on his phone], it is likely that [people] are going to think something like "ah, this guy has a girlfriend!".'

In the example above, the author describes a very girly and glittery cell phone strap, proposing that if people observe such an item attached to the mobile phone of a guy, they would presume that he has a girlfriend. The occurrence of *toka* conveys that the quoted utterance is just a conjecture, since it is impossible to presume what others may think exactly by looking at the strap. Moreover, at the end of the sentence, the speaker uses also the evidential marker *sō*(*da*) which indicates that the truth value of the statement is based on the reported evidence (hearsay evidence), to further display uncertainly.

    Approximation of numbers or measurements and approximation of quoted utterance seem to be peculiar of *toka*, and, in our corpus, have not been observed for others exemplifying markers (e.g., *tari* and *nado*). As for the use of *toka* after quotations, this is

probably due to the fact that this *toka* derives from the combination of the quotative marker *to* (which is used to introduce direct and indirect speech) and the indefinite marker *ka*[50].

More generally, in our corpus, *toka* is the only exemplifying construction that displays pure propositional vagueness as a possible discursive effect, probably due to the fact that it is the only strategy that can be used as a propositional hedge (i.e., it affects the truth value of the proposition).

5.2.2 EXEMPLIFYING TO BE POLITE

Politeness encompasses "a battery of social skills whose goal is to ensure everyone feels affirmed in a social interaction" (Foley 1997: 270). As a term used in the analysis of linguistic interaction (cf. Brown and Levinson 1987), it is based on the idea that each of us has a self-image or "face" that we want to preserve and to be respected by others (Goffman 1955). Brown and Levinson (1987) state that face can be positive or negative. The positive face can be defined as the individual's desire to be appreciated in social relationship, while negative face as the individual's desire for freedom of action and freedom from imposition. Although it is assumed that members of society generally cooperate, they are aware that the face is always at risk (see Goffman 1955). This results in a range of linguistic strategies developed in order to preserve and respect both positive and negative faces. More specifically, Brown and Levinson assume the existence of positive politeness strategies to preserve the former and negative politeness strategies to respect the latter. In this sense, positive politeness strategies are those that are directed toward positive face by communicating that one's own wants are in some respects similar to the address's wants. Whereas, negative politeness strategies are those employed by speakers to minimize a challenge to the hearer's negative face.

Brown and Levinson focus primarily on negative politeness strategies that includes the hedging of the illocutionary force to avoid imposing. For instance, consider again the example in (5.3), repeated here as in (5.24):

---

[50] As we have noted in the introduction (cf. section 1.3.2), Taylor (2015) speculates that there is not one single pragmatic marker *toka*, but two different markers: the first deriving from the combination of the comitative marker to and the indefinite marker *ka* (i.e., exemplifying *toka*), the second deriving from the combination of the quotative marker to and the indefinite marker *ka* (i.e., quotative *toka*). The use of *toka* to quote other people's words seems to derive directly from the latter.

(5.24) *Koohii toka   nomitai.*

      coffee TOKA   drink:DES

      'I want to have a coffee or something.'

The specific demand could cause tension between the interlocutors as the speaker imposes a personal wish on the hearer. Therefore, the speaker decides to make the statement less assertive by adding *toka*. The implication is that coffee is just an example of what the speaker would like, thus attenuating the speaker's commitment towards the demand.

It is in this sense that exemplifying constructions may function in a wide range of contexts making the utterance more polite, simply by presenting the potential threating item or utterance just as an example of different (potentially valid) alternatives. In most cases, these alternative are just a fuzzy background to weaken the assertiveness of the utterance, and not real choices available to the hearer.

The typical instances of negative politeness by hedging the illocutionary force regard the employment of exemplifying constructions to enunciate requests, advices, orders and other directive acts that represent a threat towards the freedom of action of the hearer, in a form of attenuation. Here are some instances of questions and requests.

(5.25) *Jiko-shōkai*   **nado** *dō*   *kana*

      self-introduction   NADO how   MOD

      'How about a self introduction?'

The sentence above is the title of a discussion in a forum where it is required a brief personal introduction to the users. Even if the implication of the exemplifying construction *nado* is that *jikoshōkai* "self introduction" is just an example of different other options, the underlying goal of the author is to focus exclusively on the mentioned item, but without sounding too imposing on the readers.

(5.26) *Ashita*   **toka** *hima* *desu* *ka?*   *Gorufu shimasen*   *ka?*

      Tomorrow   TOKA free   COP   Q   Golf   do:POL:NEG   Q

      'Are you free tomorrow? Would you like to golf?'

(5.27) *Kiraku-ni*   *o-shokuji*   **nado** *ikaga*   *desu* *ka?*

      confortable-ADJ HON-mean   NADO how about   COP   Q

'How about something like a dinner at your ease?'

Here above are two examples of suggestions or questions about future plans. The first example in (5.26) is part of an exchange between users in an online message board about golfing, while the author of (5.27) is a man in his mid-30s who is searching for women of his own age for dating purposes. In both cases, the exemplifying constructions (i.e., *toka, nado*) are used to make the request less direct and thus to attenuate the imposition over the reader.

(5.28) A: *Kanojo*   **toka**   *iru?*   *Ima.*
              Girlfriend   TOKA   AUX   now
              'Do you have a girlfriend? I mean now.'

        B: *Ima*   *wa*   *inai*   *desu*   *nee.*
              now   TOP   AUX:NEG   COP   PP
              'At the moment I don't.'

Here a girl is asking to a guy if he has a girlfriend. She uses *toka* to avoid sounding too harsh and direct. In this case, it is clear that the purpose of the exemplifying construction is not to create a concrete range of options (with the mentioned element merely as an example), but only to weaken the tension that might be caused by a specific request o suggestion.

Interestingly, exemplifying constructions as politeness strategies seems to operate regardless of the different levels of politeness and formality of the Japanese language. For instance, the exchange in (5.28) occurs between two young people. The girl formulates her question using informal casual speech: she does not use the polite form of the verb *iru* (i.e., *imasu*), nor the interrogative marker *ka* (which is dropped very frequently in casual Japanese), relying only on intonation. Nevertheless, she still chooses to use the exemplifying construction *toka* instead of the topic marker *wa* to avoid sounding too forward and direct. On the contrary, the question in (5.27) is more formal because of the polite form of the copula *desu* (instead of *da*) and the honorific prefix *o* before *shokuji* "meal". However, also in this case, an exemplifying construction (i.e., *nado*) is used to mitigate the assertiveness of the request.

Exemplifying construction can also be used to express direct orders or polite advices without sounding too harsh or direct. Here are two examples.

(5.29) *Keshitari   shinaide   kudasai*
      turn.off:TARI   do:NEG:GRD   please

      'Please do not turn off the lights."

(5.30) *Motomoto   mizu   wa   iranai   shokubutsu   na   node,*
      originally   water TOP   need:NEG   plant   COP   because

      dondon   mizu-o   agetari   shinai kata-ga   ii   yo.
      rapidly   water-ACC   give:TARI   do:NEG way-NOM   good   PP

      'Because it is a plant which do not need water, it is better not to give water rapidly.'

While directive acts can be considered a breeding ground for hedging strategies, it is still possible to find occurrences of exemplifying constructions as hedges also in declarative sentences, whenever it is needed to weaken the commitment. Consider the following example:

(5.31) *Nita   mono fūfu   to   iu   kotoba-ga   aru*
      resamble:PAST   person spouses   QT   say   expression-NOM   AUX

      yōni,   kattenagara   desu   ga,   yahari saisho   kara   umakuiku
      so that   taking the liberty   COP   but   still   beginning   from   go smoothly

      kappuru   niwa   iwakan-o   kanjinakattari suru   mono desu.
      couple   for   uncomfortable.feeling-ACC   feel:NEG:TARI do   NML   COP

      'So that there is an expression saying "like man like wife", I am taking the liberty to speak, but still for a couple that worked well from the beginning, it may be that they have not felt a sense of incompatibility.'

The sentence above is part of a gossip article. The author comments a famous couple who got divorced, saying that maybe because the couple worked well and smoothly at the beginning of the relationship, they did not realize that they were not well-matched. The implication is that they are not perceived as very suitable to each other, unlike what is suggested by the famous maxim about men and wives.

In this case, we argue that the usage of *tari* is not categorization, but hedging. First of all, it is difficult to even imagine a category of similar elements to the one explicitly mentioned. Moreover, the idea of not being/feeling compatible is actually the main point of the entire paragraph (e.g., the quotation of the proverb about spouses being similar). Secondly, the author is aware of commenting a private, and therefore delicate, issue. For instance, she uses the fixed expression *kattenagara* which, depending on the context, can be translated

as something like "doing without asking" or "taking the liberty to do". Sometimes it may also carry the feeling of being uncooperative. More specifically, here it means that she is speaking on her own initiative and therefore she may be wrong in her statement. Therefore, she deliberately weakens her commitment towards the problematic comment using *tari*, thus attenuating the assertiveness of the entire sentence.

In our corpus, this usage of *tari* to avoid being too forward or rude regarding a personal opinion is very frequent. Another example is the following:

(5.32) *Shashin o   seisasureba   tashō no   noizu-kan-ya   saibu no*
picture-ACC   examination:do:COND   a little LK   noise-feeling-and   detail LK
tsubure-ga   aru   monono   "kirei da   kara   okkē!!" to
collapse-NOM exist   although   pretty COP   because   okay   QT
omoetari   shite.
think:POT:TARI   do:GRD
'If you examine the picture, although there are a collapse of the details and a noise-feeling, you may think 'because it is pretty, it is ok!!'.'

In the sentence above, a blogger provides friendly advices and suggestions about how to use a photo camera in order to take good pictures. More specifically, he says that sometimes even if the picture is not perfect (e.g., the details are not tidy and there is noise in the picture), people may still think that the picture is fine because overall it is pretty to watch. Therefore, because the writer is interpreting other people's way of thinking, he chooses to end the sentence with *tari shite* implying a possibility in a more indirect form, and thus weakening the assertiveness of the utterance. The attenuate effect is not only the result of the usage of *tari* as a vague marker, but is also due to the choice of using the *te*-form at the end of the sentence. Despite not being grammatically correct to end a sentence with the *te*-form, it is often used when the speaker wants to trail off and leave the rest of the sentence implied. This contributes to the final indirect (and thus more polite) effect.

The interpretation above was confirmed by Japanese native speakers who were asked to fill out the questionnaire. According to the wide majority of the informants (72,4% in the Japanese version of the questionnaire, 100% in the English version), this sentence can be paraphrased as:

(5.33) *"Kireidakara okkē!!" to omoeru kamoshirenai.*
'You might think "because it is pretty, it is ok!!"'

Thus, they confirm that *tari* functions to mitigate the assertiveness of the utterance.

In some cases, exemplifying constructions can be even used in combination to further attenuate the illocutionary force of the utterance and show consideration for the hearer. In our corpus, two occurrences of the combination of *tari* and *toka* are attested. Consider the following example:

(5.34) A: *Raibu-o*   *yukikisuru*   *igaini,*   *gohan-o*   *tabe*   *ni*
          live-ACC   coming.and.going:do   in addition   meal-ACC   eat:GRD DAT
          ittari   *toka*   *nasaru*   *n*   *desu*   *ka?*
          go:TARI   TOKA   do:HON   NML   COP   Q
          'In addition to go back and forth live, do you (go to) eat?'

     B: *Sore*   *ga*   *desu*   *ne.*
          that   NOM   COP   PP
          'Yes, we do! '

This exchange is part of an interview with a famous singer (therefore, *raibu* "live" should be interpreted as "live performances"). The addition of *tari* to the verb *iku* "to go" makes the question less direct and therefore more polite. The attenuation effect is further stress by adding *toka* after *tari*. Taylor (2010) notes that the string *tari toka* is very frequent in the spoken discourse "to have a successful and smooth conversation" (2010: 179). In such cases, these two exemplifying markers provide no contribution to the propositional content, but their usage in combination further weakens the illocutionary force of the utterance.

Exemplifying constructions function also as devices employed to protect the positive face of the speaker and to signal solidarity with the hearer. Consider the following exchange:

(5.35) A: *Watashi*   *mo*   *fudan*   *kyōshitsu*   *kara*   *derarenai*   *kara,*
          I   also   often   classroom   from   exit:POT:NEG   because
          jibun-no   *o-bentō*   *tottari*   *toka*   *(emi)*
          self-DET   HON-bento   take.picture:TARI   TOKA   (laugh)
          'Me too, because usually I cannot get out from the classroom, I take pictures
          of my lunch box (laugh).'

     B: *Boku*   *mo*   *hirugohan*   *totteru (emi).*
          I   also   lunch   take.picture:STA (laugh)
          'I take pictures of lunch too (laugh)'.

The topic of the conversation above regards photography. Speaker B is explaining that since he is always very busy with his job in a company, he always carries his camera with him so that he can take pictures at any time he can, and that he prefers pictures portraying everyday objects and activities. Speaker A intervenes laughing to say that she as well is often so busy that the only pictures that she can take are those of her lunch box. Finally, speaker B laughs saying that he does the same. Like in (5.34), here the combination of *tari* and *toka* at the end of the utterance *jibun no obentō toru* "(I) take picture of my lunch box" is for hedging and not for exemplifying. The whole interview regards the hobby of photography, so it is very unlikely that speaker A wants to refer to other similar actions beyond taking pictures. Also, if speaker A had wanted to emphasize "bento box" as an example of objects to photograph in that specific situation, *toka* would have been used instead of the object marker (here elided), thus immediately after "bento box". On the contrary *tari toka* are added at the end of the utterance and their scope is indeed the entire utterance. Moreover, "taking picture of my lunch box" is indeed the topic and stays active through the subsequent interaction ("I take pictures of my lunch too").

In this exchange, we argue that the actual purpose of the combination of *tari toka* is twofold. On one hand, the positive face of the speaker may be threatened by exposing a personal habit which might seem weird (i.e., taking picture of her lunch box) even by the speaker herself (it is indicated that she is laughing), but that helps her to mark solidarity with the addressee. On the other hand, while signalling solidarity, she does not want to threat the negative face of the hearer by implying that her taking picture of the lunch can be really comparable to the addressee's work as a photographer. In order to find a compromise between these two forces, she chooses to attenuate the strength of the statement by using two exemplifying markers, i.e., *tari* and *toka* for the sake of interpersonal communication. At the end, the communication is successful: speaker B replies accepting and remarking the offered solidarity with speaker A (i.e., "I take picture too").

The picture emerging from this section is one in which exemplifying strategies are versatile tools to express politeness. This is true even in a language such as Japanese which is definitely not lacking in dedicated politeness markers. Nevertheless, exemplifying strategies do not seem to be a prerogative of polite (i.e., *keitai*) or formal (i.e., *keigo*) levels of politeness, but, on the contrary, they are widespread in the entire spectrum of speech styles and registers. For example, in this section, we have seen instances of exemplifying strategies used with the informal style, such as in (5.28), with a polite style, such as in (5.27) and (5.29), and even with a mix of polite and formal style, such as in (5.34). Even more,

sometimes, they can be used by speakers to search for the right level of politeness: when honorific forms are considered too formal (or too personal), exemplifying strategies can be used instead to softening the force of the utterance without exaggerating (cf. Taylor 2010: 183).

6. EXEMPLIFYING CONSTRUCTIONS BEYOND EXEMPLIFICATION

6.1 EXEMPLIFYING CONNECTIVES IN EXHAUSTIVE CONTEXTS

As we explained in the introduction (cf. section 1.1.2), one of the reason for choosing Japanese was the existence of synthetic dedicated strategies used with the main purpose of expressing non-exhaustive lists. More specifically, Japanese has an extremely rich system of connectives which specifically signal that the linked items are examples taken from a larger set (Chino 2001, Tanimori 1994: 121-122, 265), for instance *ya*, *tari* and *toka*, as shown in the following examples:

(6.1) *Biiru-ya   sake-o   takusan   nomimashita.*
      beer-YA   sake-ACC   lots of   drink:POL:PAST
      '[I] drank lots of beer and sake and stuff like that.' (Kuno 1973: 115):

(6.2) *Osaka-de   kaimono-o   shitari kankoku-ryoori-o   tabetari   shimasu.*
      Osaka-loc   shopping-acc   do:TARI Korean-meal-ACC   eat:TARI   do:POL
      'In Osaka, I will do such things as shopping, eating Korean food and so on.' (Banno 2000:215)

(6.3) *Koohii-toka   koocha-toka iroirona mono-ga   arimashita*
      coffee-TOKA   tea-TOKA   various   thing-NOM   exist:POL:PAST
      'There were various things such as coffee, tea and so on.' (Maynard 1990: 106)

Due to the ability of encoding open-ended lists, these connectives are often described by grammars (e.g., Kuno 1973) as dedicated non-exhaustive connectives, thus implying that they are incompatible with any kind of exhaustive contexts. Actually, to encode exhaustive lists, Japanese exhibits a different set of connectives. For instance, to join nouns and noun phrases, the exhaustive combination marker *to* (6.4) and the exhaustive alternative marker *ka* (6.5) are attested:

(6.4) *Shiroi yuri-to   akai bara no   hana-o   kaimashou.*
      white   lily-and   red   rose LK   flower-ACC   buy:POL:VOC
      'Let's buy some white lilies and red roses.' (Chino 2001: 33)

(6.5) *Kohi   ka   kocha ka   nomitai   desu ne.*
       coffee or   tea   or   drink:DES   COP   PP

       'I would like to drink a coffee or a tea.' (Chino 2001: 47)

To join verbs and verbal phrases, the converb –*te* (or *te*-form of the verb) functions as the exhaustive combination marker (6.6)[51]. Moreover, the exhaustive alternative marker *ka* is used also to link verbs as in (6.7).

(6.6) *Ojii.san-ga   yama-de   hatairaite,*   *obaasan-ga*   *mise-no*   *ban-o*
       old.man-NOM   mountain-LOC work:and   old.woman-NOM   store-GEN   sitting-ACC

       shita.
       do:PAST

       'The old man worked at the mountain, and the old woman tended the store.' (Yuasa and Sadock 2002: 92, quoted in Haspelmath 2004: 34)

(6.7) *Ryoko-ni   iku   ka   ikanai   ka,   mada kimeteimasen.*
       trip-LOC   go   or   go:NEG   or   still   decide:STA:POL:NEG

       'I still haven't decided whether I am going to take a trip or not.' (Chino 2001: 47)

This set of connectives can be schematized as follows:

Table 6.1: Japanese logical connectives

		Exhaustive	**Non-exhaustive**
Nouns	Combination	*to*	*ya, toka*
	Alternative	*ka*	
Verbs	Combination	*-te*	*tari, toka*
	Alternative	*ka*	

While the schema above is not completely incorrect, linguistic studies focusing on the actual usage of these connectives and evidences attested in our corpus, portray a more varied picture.

---

[51] In our analysis, we decided to focus only on the *te*-form of the verb, because it is the most widespread strategy. However, there are other strategies to mark combination relations among verbs in Japanese. For instance, both the *i*-form (or stem form) and the marker *shi* can be used to conjunct verbal phrases. They will not be considered in our analysis, because their usage is less widespread and more constrained, also at the functional level. For example, the *i*-form is mainly limited to the written formal style (cf. Ohori 2004) and usually marks only atemporal relations (cf. Mauri 2008). The marker *shi* is frequently used to give reasons, motivations, excuses, etc. For this reason, it often conveys a cause-effect meaning (Kaisen et *ali*. 2001: 90).

For example, as briefly mentioned while describing the parameters of analysis (cf. section 2.3.3.1), it is well acknowledged that the converb –*tari* can perform other functions beyond non-exhaustive representation (cf. Martin 1924: 566-573). In particular, –*tari* has been linked to notions such as iterative aspect and distributive aspect, since it can be used to express the repeated performance of one or more actions. For example, Alpatov (1997) notes that the usage of *tari* is "connected with the semantic zone of the multiplicative and distributive" (1997: 392). More specifically, *tari* can be linked to the meaning of "intermittent" (Alpatov 1997: 392), indicating the discontinue repetition of an action or similar actions repeating and alternating with one another.

(6.8)   *Maru-de*   *inukoro*   *mitai-ni,*   *fuzakeattari*   *kenka-o*   *shitari.*
        Quite   puppy   like-ADV   play:TARI   quarrel-ACC   do:TARI
        (lit.) 'They were like puppies, they played, they quarrelled.'
        (id.) 'They were like puppies. At times they played, at other quarrelled.' (Alpatov 1997: 393)

In the sentence above, even without explicit adverbs to indicate the repetition and alternation of the actions, the usage of *tari* to mark opposing events gives rise to an iterative interpretation. Similar instances are also highlighted by Narrog (2012: 145) while analysing the expression of modality in Japanese.

Moreover, the fact that *tari* can indicate repetition of contrasting actions is often reported even by grammars targeting L2 learners (e.g., Chino 2001).

(6.9)   *Konshū,*   *kabu* *wa* *agattari*   *sagattari*   *shiteimasu.*
        this week   stock   TOP   go.up:TARI   go.down:TARI   do:STA:POL
        'This week, stocks are going up and down.' (Chino 2001: 109)

Interestingly, in all these instances, the lists of events encoded by *tari* are exhaustive, in the sense that the mentioned events are not representative of a larger group. This has been noted also by Alpatov, that clearly distinguishes instances of *tari* with an "intermittent" meaning and those with a "representative" meaning (1997: 392).

Being *tari* a converb originated from an aspectual auxiliary, it is possible to argue that, at least in some contexts, it may still bear some aspectual values. Nevertheless, this does not seem the case. In Old Japanese, the original aspectual auxiliary -*tari* was a resultative perfect, denoting completion of an event and the state of result of this event. Indeed, its Modern Japanese counterpart is the past/perfective marker –*ta*, which is also used to denote

perfect aspect (Iwasaki 2013: 132-133). Thus, connecting a hypothetical imperfective aspect of modern *tari* with the original resultative one would be problematic. In this respect, Narrog (2012) notes that *tari* in Modern Japanese has lost its aspectual meaning. He also claims that the possibility of using *tari* to mark iterative events relates to a lower factuality, "since through *-tari* the event becomes marked as unspecific, i.e. not bound to a specific time, and ambiguous between a singular and a plural reading" (2012: 147). This means that *tari* has emancipated itself from the originally aspectual meaning to become a marker of generic events (then assuming also the exemplifying function). Therefore, we can say that the "intermittent" meaning does not originate from *tari* being an aspectual marker.

Moreover, in other instances, the iterative meaning is much subtler, as shown in following example from our corpus:

(6.10) | *Jiyū-undai-no* | *bōru-bu-o* | *koteishitari* | *furī-ni* | *shitari* | *suru* |
|---|---|---|---|---|---|
| universal-tripod-GEN | ball-component-ACC | fix:do:TARI | free-DAT | do:TARI | do |
| *tameno shimetsuke* | *nobu.* | | | | |
| for   tightening | knob | | | | |

'The tightening knob for fixing and/or freeing the ball part of the universal tripod.'

Another point of interest concerns the fact that *tari* assumed this exhaustive "intermittent" meaning only when it is used as a connective of at least two items. In other words, whenever *tari* is used as a general extender attached to only one event, it is typically used to exemplify (cf. chapters 3 and 4) or to perform pragmatic functions (cf. chapter 5), but it does not indicate the iteration of a single event.

Therefore, another possible interpretation is that these exhaustive meanings are actually encoded by the usage of *tari* as a connective. In other words, we may argue that *tari* codifies a specific type of semantic relation occurring among the connected items. For example, in (6.8) and in (6.9), *tari* indicates that both events have happened several times, but never at the same time.

This perspective is backed up also by the fact that the usage of "non-exhaustive" connectives in exhaustive contexts is not a prerogative of *tari*. Indeed, *ya* exhibits similar functions.

While investigating the acquisition of Japanese connectives, Ichikawa (1991) notes that L2 learners show some difficulties in understanding the different modalities of use of *to* (i.e., the exhaustive connective for nouns) and *ya* (i.e., the non-exhaustive connective for nouns).

Specifically, in order to describe the Japanese custom to go to temples and shrines during New Year's Day, one of the learners expresses the concept in the following way:

(6.11) *Nihonjin*   *wa o-shōgatsu-ni*   *jinja-to*   *o-tera-ni*   *iku.*

Japanese-person   TOP HON-New.Year-LOC   Shinto shrine-and   HON-temple-LOC go

'On New Year's Day, Japanese people go to Shinto shrines and Buddhist temples.'

In other words, the L2 learner chooses to use the connective *to* instead of *ya* to link *jinja* 'Shinto shrines' and *otera* 'Buddhist temples'. The reason behind this choice is merely based on exhaustivity: Japanese people do not go to other types of religious buildings during New Years' Day (e.g., churches), *jinja* and *otera* are the only elements of the list, which is therefore exhaustive. Ichikawa notes that many grammars and textbooks targeting L2 learners explain the difference between *to* and *ya* only referring to the parameter of (non-)exhaustivity: *to* encodes the exhaustive combination relation, while *ya* is usually described as the non-exhaustive connective to encode open-ended lists. Therefore, the reasoning underlying the choice of the right connective can be schematized as follows:

(6.12) *jinja to otera* = exhaustive

On New Year's Day, Japanese people go to Shinto shrines and to the Buddhist temples.

(6.13) *jinja ya otera* = non-exhaustive

On New Year's Day, Japanese people go to Shinto shrines, to the Buddhist temples, and so on. (e.g., there are other options, such as churches)

In this sense, the choice of the L2 learner to use *to* instead of *ya* in this context is perfectly understandable: *jinja* and *otera* are indeed the only two members of the list.

Nevertheless, this formulation does not sound natural to native speakers. Japanese native speakers would express the same concept as:

(6.14)   *Nihonjin*   *wa*   *o-shōgatsu-ni*   *wa*   *jinja-ya*

        Japanese.people   TOP   HON-New.Year-LOC   TOP   Shinto.shrine-YA

        o-tera-ni   *iku.*

        HON-temple-LOC   go

        ''On New Year's Day, Japanese people go to Shinto shrines and/or[52] Buddhist temples.'

Therefore, for native speakers, in this context, the use of *ya* is more natural than the use of *to* (i.e., *jinja ya otera*), implying that *ya* is indeed compatible with at least some exhaustive contexts.

This fact is confirmed by evidences in our corpus, where occurrences of *ya* used to create exhaustive lists are attested. Consider the following example:

(6.15)   *Youchien-ya*   *hoikusho-ni*   *kayou* *ko-o*   *motsu*   *hogosha*

        kindergarden-YA   nursing-LOC   attend   kid-ACC   possess   guardian

        yaku   6500-nin   *no*   *yaku 4-wari-ga*   *"san-nin*   *nori-o*

        about 6500-CLF   LK   about 4-ten.percent-NOM   "three-CLF   ride-ACC

        shitatameru   beki   da"   *to*   *kaitou.*

        approve   IMP   COP"   QT   answer

        'About the 40% of the approximately 6,500 parents who have children attending kindergartens or nursery schools answered: "(Riding with) three people on the same vehicle should be allowed".'

Here, *ya* is used to joint *youchien* "kindergarden" and *hoikusho* "nursery school". In this context, it is unlikely that *ya* encodes an open-ended list mainly for two reasons: 1) since the context refers to small children, the only other potential option beyond those mentioned would be elementary school, thus it would have been easier just to add it, instead of using a non-exhaustive list; 2) being a survey, it is unlikely that the sample is treated as a non-defined set, especially because there are not many potential elements to justify an omission. The more likely interpretation is that the list is indeed exhaustive and that there are other reasons to use *ya* instead of *to*.

At this point, an obvious question is: knowing that those in (6.13) and (6.14) are indeed exhaustive lists, why do native speakers use a connective typically employed to express non-exhaustive lists? The first part of the answer is simple: the potential of being used in

---

[52] Interestingly, in this context, both English connectives are possible (cf. Ariel and Mauri, forthcoming).

exhaustive/non-exhaustive contexts is not the only difference between *to* and *ya*. However, this also means that there are other parameters that we need to investigate further.

To better understand the differences between *to* and *ya*, it is crucial to consider the reasons for choosing *ya* in (6.14). While it is true that there are no other elements in the list of places where Japanese people go during New Year's Day, we should also consider the fact that people do not go in both places (i.e., Shinto shrines and Buddhist temples), but that depending on the family's habits and traditions, they decide for one of the two over the other. In other words, some people go to the Shinto shrine, other people go to the Buddhist temple and these sets tend not to overlap. Thus, we may say that both options occur, yet they do not occur at the same time for the same people.

This is the reason as to why native speakers do not use the connective *to*. In fact, even if the list is exhaustive, the use of *to* would also imply that Japanese people go to both places, that is, to the Shinto shrine and also to the Buddhist temple (i.e., 'all of X and Y'). On the contrary, the connective *ya* allows to codify this different semantic relation between the elements of the list, which can be well paraphrased as 'any of X and Y'.

This example reveals that *ya* does not exclusively encode non-exhaustivity, but on the contrary, it can be also used in exhaustive contexts to express types of semantic relations that seem to be incompatible with other coordinating markers.

In the light of this, we may also explain the usage of *ya* in (6.15). In that sentence, the author refers to the parents who have children attending kindergartens or nursery schools. In other words, there are two different groups of people (i.e., parents of children attending kindergarten, parents of children attending nursery school) and what follows holds for both. Following the interpretation suggested by Ichikawa (1991), it is likely that the usage of *to* instead of *ya* would have implied that only parents having children attending kindergartens and (also) nursery schools were considered. In other words, two or more children attending to both types of school.

Consequently, the author decides to use *ya* to highlight that 1) there are two different groups and that 2) both were considered. In a sense, we may argue that *ya* seems to convey a distributive "nuance".

At this point, we may wonder how close these exhaustive uses of *tari* and *ya* are. Consider the following example of *tari* from our corpus, while confronting it with the example in (6.15).

(6.16) *[Shashin-o   shumi to   shitari,   satsuei- ga   shigoto dattari suru]*
picture-ACC   hobby as   do:TARI   photographing-NOM   job   cop:TARI do

hito   nitotte, kamera ya   renzu wa   mochiron,   sankyaku-ya undai
people for   camera-YA lens   TOP   of course   tripod-YA   platform

mo   hijō-ni   jūyōna   kizai.
also   extremely   important   gear

'For those who take picture as a hobby and/or for those who take picture as a job, camera and lens, of course [are important gears], also tripod and camera platform are very important gears.'

In the sentence above, we can find a very similar situation to (6.14), despite the conjuncts being verbs. The author refers to two groups of people (i.e., those who take photos as a hobby and those who do the same as a job) that mainly do not coincide and points out that what follows holds for both (i.e., both groups need some specific gears to take good pictures).

The fact that we can find similar exhaustive meanings in *ya* and *tari* despite having different scopes, is a good cue that we are facing similar issues, and that the answer should be investigated in the semantic relation that these connectives can codify.

Furthermore, similar meanings are exhibited also by the last "non-exhaustive" connective under analysis, *toka*. The usage of *toka* in exhaustive contexts has not been investigated in depths (unlike the usage of *ya* and *tari*), yet, in our corpus, it does exhibit some exhaustive meanings, albeit exhibiting a smaller range compared to that of *ya* and *tari*. Consider the following examples from our corpus.

(6.17) *"Shinkon ryokō"   wa   "honeymoon trip"   toka   "honeymoon" to   iimasu.*
Honeymoon trip   TOP   honeymoon   trip   TOKA   honeymoon   QT   call:POL
'[In English] "Honeymoon trip" is called "honeymoon trip" or "honeymoon".'

(6.18) *Hoka no   yūzā-o   "dōshi"   toka   "raibaru"   toshite kuwaeru   koto*
other LK   user-ACC   comrade   TOKA   rival   as   add up   NML

mo   kanō.
also   possible
'It is also possible to add up other users as "rivals" or "comrades".'

In both cases, *toka* seems to codify a semantic relation that can be paraphrased as "both are valid, but not at the same time". In (6.17), the author claims that in order to refer to the concept 'honeymoon trip' in English, speakers can use the expression *honeymoon trip* or

the expression *honeymoon*. Both expressions are valid, but they do not occur at the same time. In (6.18), the author says that in the game it is possible to add other users as "rivals" (against whom to play) or as "comrades" (with whom to play). Both options are possible, but of course not at the same time: a user can be added as a rival or as a comrade. Again, as in (6.15) and in (6.16), it is possible to identify a distributive "nuance".

In our corpus, occurrences like the ones described in this sections are relatively scarce, as shown in Figure 6.1.
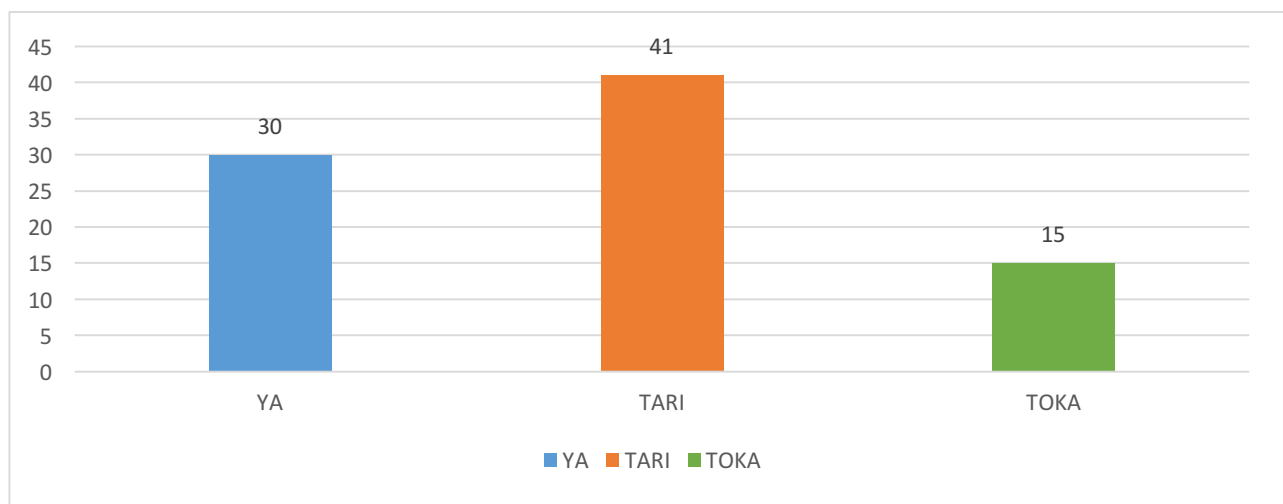
Figure 6.1: Frequency of 'non-exhaustive' connectives in exhaustive contexts.

To identify this type of occurrences (thus checking the exhaustivity of the encoded lists), we have strongly relied on the interpretation of native speakers, for example by means of the questionnaire. As it was explained in section 2.2.2, we asked to the informants to value the paraphrases of some occurrences from our corpus. More specifically, we often proposed the exhaustive and the non-exhaustive versions of the same paraphrase (see the questionnaire in Appendix B) for the native speakers to evaluate the naturalness. For example, consider the following occurrence.

(6.19) *Anata-no*   *konpyūta-ni*   *wa*   *aki*   *ryōiki*   *wa*   *jūbunni arimasu ka?*
        you-GEN   computer-LOC TOP   free   space   TOP   enough exist:POL Q
        Sōdenainara, *suwappu*   *saizu-o*   *ōkikushitari,*   *butsuri*   *memori-o*
        If not   swap   size-ACC   increase:do:TARI   physical memory- ACC
        fuyasu   *koto*   *de*   *taisho*   *dekimasu.*
        increase   NML   STR   deal with   POT:POL
        'Is there enough free space on your computer? If not, you can deal with it by increasing the swap size or increasing the physical memory.'

According to the wide majority of the informants (96,6% in the Japanese version of the questionnaire, 100% in the English version), this sentence can be paraphrased as:

(6.20) *Suwappu saizu o ōkikusuru, matawa, butsuri memori o fuyasu. Kono futatsu no taisho hō ga ari, kono ryouhou o shitemo, dochiraka dake demo yoi.*
'There are two possible solutions: you can do both or just one of them. You can increase the swap size or you can increase the physical memory.'

Therefore, they confirm the exhaustive interpretation, despite the presence of a "non-exhaustive" connective, namely *tari*.

Interestingly, also in (6.19) it is possible to infer an underlying distributive nuance. For this reason, we have also monitored the aspectual value of the occurrences (cf. section 2.3.3.1)

While all the connectives under analysis have been attested in exhaustive contexts, they do not behave uniformly, exhibiting different ranges of contexts in which they can be used.

For example, it is noteworthy that in our corpus all the instances of *toka* conveying exhaustive meanings are identical to (6.17) and (6.18). In other words, exhaustive *toka* is always used to connect parts of direct speech such as expressions or labels. This fact is interesting since *toka* is often conceived as a quotative marker (cf. section 1.3.2.4 and section 5.2.1) to introduce direct and indirect speech while expressing non-exhaustivity or vagueness. Furthermore, almost all the occurrences follow the same pattern of (6.16), that is, *toka* is used do point out that a certain expression can be translated (usually in English) as X or as Y. This means that the relation between the use of *toka* as quotative marker and that as a connective is very strong, to the point that in some cases they are indistinguishable.

To sum up, at least in our corpus, the range of exhaustive contexts in which *toka* can be used is very specific and strictly linked to its usage as a quotative marker. On the contrary, the ranges of exhaustive contexts in which *ya* and *tari* can be used appear to be more various. Moreover, it is noteworthy that all the connectives under analysis, that is, connectives that have been described as "non-exhaustive", can indeed exhibit exhaustive meanings, and more importantly, these meanings seem to be related to each other.

The picture thus portrayed casts serious doubts on the schema of the Japanese connectives proposed in Table 6.1. Moreover, in this regard, labels such as 'combination' and 'alternation' (see Mauri 2008) start getting too loose.

6.1.1 BETWEEN 'AND' AND 'OR': CONJUNCTION IN ALTERNATION

In her cross-linguistic study on coordination, Mauri (2008) highlights how several European languages (included for instance English) fall within the so-called 'And-But-Or' language type. Among other parameters, the 'And-But-Or' languages are characterized by a general marker used for both temporal and atemporal combination ('And'), a general marker used for both choice-aimed and simple alternatives ('Or'). Therefore, for instance, English exhibits the connectives *and* and *or*, Italian exhibits *e* and *o*, French exhibits *et* and *ou*, etcetera. In other words, these languages create a bipartite division of the cognitive space (cf. Croft 2001) of coordination. As Mauri notes (2008: 289-293), this behaviour is very peculiar, because languages around the world exhibit different systems. For example, Japanese divides the same space in a tripartite system, where some meanings of 'And' are covered by *to* and *ya*, and some meanings of 'Or' are covered by *ka* and *ya* at the noun phrase level; and where some meanings of 'And' are covered by the converbs *te* and *tari*, and some meanings of 'Or' are covered by *ka* and *tari* at the verbal phrase level.

As noted in the previous section, the impression given by many descriptive grammars (e.g., Kuno 1973, Martin 1975, Chino 2001) is that exhaustivity is the only parameter that differentiates between the meanings covered by *ya/tari* and those covered by *ka* and *to/te*. However, we have seen that this is not always the case. Thus, a further investigation is required to better understand the actual meanings covered by Japanese connectives.
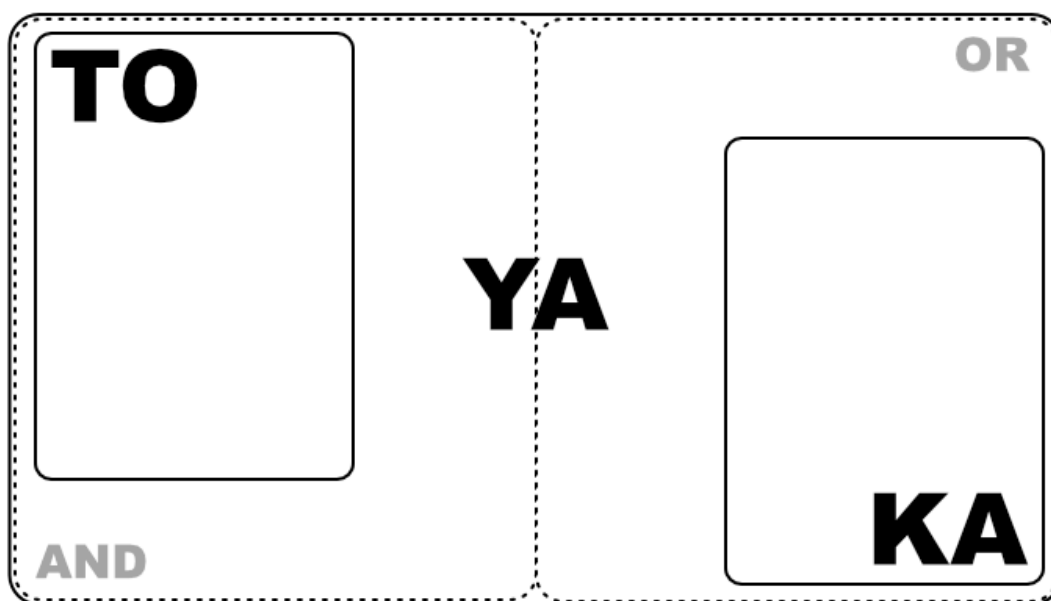
Figure 6.2: Coordination in English and Japanese (noun phrase level)

To achieve this investigation, we should abandon labels such as 'combination' and 'alternative', which does not seem helpful, and try to focus in depth on the different nuances of meaning behind them.

A different approach that can be applied to our analysis is to consider that the interpretation of utterances (and therefore also of the utterances formed around truth-functional connectives) very often combines the compositional meaning with pragmatic inferences (Carston 2002), with the latter bridging the gap between what is linguistically expressed (cf. the concept of *logical form*, Carston 2002: 57), and what is said, that is, the full propositional representation.

This type of approach has already been employed, for instance, regarding the English connective *or*. While it is true that *or* constructions always introduce a set of alternatives, this alternativity should be conceived only as a starting point for expressing a range of different meanings (cf. Ariel and Mauri, forthcoming). In this regard, ever since Grice (1989) and Horn (1972), it has been assumed that *or* has an inclusive lexical meaning, but a predominantly exclusive use. Recently, Ariel and Mauri (forthcoming) propose a more fine-grained distinction starting from a usage-based model. In order to map all the available meanings of *or* in context, they propose Relevance-theoretic *explicature* (cf. Sperber and Wilson 1986/1995) as the appropriate meaning level with which to characterize *or* readings. In this regard, explicatures stand for the messages directly expressed by the speaker, thus, without considering *i*) pragmatic inference which are not part of the speaker-intended message or *ii*) special ad-hoc interpretations. In this sense, they define an *or* reading as "any distinct type of explicature recurrently developed out of an utterance containing *or*" (Ariel and Mauri, forthcoming). Through this type of analysis, it has been possible to identify a far richer variety of readings than has been previously recognized.

For instance, they identify a reading labelled as *Higher-level category Or* (or HLC), where speakers introduce alternatives only to make reference to a higher-level category. This reading is what we have discussed in depth in chapters 3 and 4 regarding *ya/tari/toka*.

Moreover, following the insights provided by Jennings (1994), Ariel and Mauri recognise a type of reading in which – despite the alternativity – the speaker commits to both the conjuncts (at least potentially) being the case. This reading has been labelled as *Conjunctive Or*. According to Ariel and Mauri, it encompasses instances where 1) the speaker intends to actually convey multiple alternatives, 2) the speaker is taken to commit to at least one of the alternatives, and 3) the speaker does not intend to convey a mutually exclusive 'not both' interpretation, as shown in (6.21).

(6.21) *All the sections are kind of self-sufficient, having kitchen units **and/or** bathrooms* (LSAC).

The *Conjunctive Or* reading can be further subdivided into two conjunctive cases: *Separative conjunction* and *Free alternative*.

Separative Conjunction reading can be defined as one where the speaker necessarily commits to both alternatives being the case, although each is separate from the other. In most cases, this means that they occur in separate situations. This reading is therefore motivated by the speaker's intent to highlight the "distinctness rather than the cumulativeness of the possibilities mooted" (Jennings 1994: 299). An example of this reading is provided in example (6.22).

(6.22) JIM: *But for mathematics **or** for science, ((1 LINE OMITTED))*
               it's an opportunity for them ((3 LINES OMITTED))
               to get closer to,
               (H) to the chaos, (SBC: 017)

Interestingly, instances of Separative conjunction reading can be paraphrased by *and* as shown in (6.23), since the speaker is indeed committing that possibly all disjuncts are the case.

(6.23) JIM: *But for mathematics **and** for science, ((1 LINE OMITTED))*
               it's an opportunity for them ((3 LINES OMITTED))

This fact metaphorically places the separative conjunction reading in between what canonically is considered 'and' and what canonically is considered 'or'.

Free alternative reading is similar to the separative conjunction one, except for the fact that it carries an additional "free choice" flavour (cf. Kamp 1973), where the speaker seems open to any of the alternatives, as shown in example (6.24). This 'no matter' flavour is often (but not exclusively) given by epistemic or deontic irrealis contexts.

(6.24) RANDY: *... But you can wrap em around,*
                    *to one-thirty **or** one-sixty,* (SBC: 022)

To sum up, in both cases the speaker proposes multiple alternatives. In the Separative Conjunction reading, the speaker also commits to all the conjuncts being the case. On the contrary, in the Free Alternative reading, the speaker only commits to the possibility of each of the alternatives to be realized. Moreover, he or she expects only one alternative to actually be realized, although he or she is impartial as to which alternative is actually realized.

In order to precisely identify this types of readings, Ariel and Mauri propose the linguistic formula "both X and Y" (for separative conjunction) and "any of X and Y" (for free alternative) to use in the 'That is (to say)' test. For instance,

(6.25) Jim said that for mathematics or for science, **that is (to say)** for both of mathematics and science it's an opportunity to get closer to the chaos. (Separative conjunction)

(6.26) Randy said that you can wrap em around to one-thirty or one-sixty, **that is (to say)** to any one of one-thirty and one-sixty. (Free alternative)

While this analysis is meant to describe the English connective *or*, without any ambition to provide a cross-linguistic model, we can still imagine that these readings should be available in any natural language. In other words, regardless of their language, speakers need some linguistic means to address situations where there are multiple alternatives, while at the same time committing to all of them being the case or potentially being the case. Therefore, speakers will identify and use markers that can encode (or at least are compatible with) these readings in their own language.

Interestingly, if we consider the instances provided in the preceding section and test them by means of the formula indicated by Ariel and Mauri (forthcoming), we discover that "non-exhaustive" connectives *ya, tari* and *toka* can also express these exhaustive conjunctive alternative readings. For instance, consider (6.10), (6.14), (6.15), (6.16) and (6.17), here repeated and tested as (6.27), (6.28), (6.29) (6.30) and (6.31).

(6.27) *Jiyū-undai-no*   *bōru-bu-o*   *koteishitari*   *furī-ni*   *shitari*   *suru*
        universal-tripod-GEN   ball-component-ACC   fix:do:TARI   free-DAT   do:TARI   do
        tameno shimetsuke   nobu.
        for   tightening   knob
        'The tightening knob for fixing or freeing the ball part of the universal tripod.'
        TEST: He said that the tightening knob is for fixing or freeing the ball part, **that is (to say)**, it is for **both** (of) fixing **and** freeing the ball part. = Separative conjunction

(6.28) *Nihonjin*   *wa*   *o-shōgatsu-ni*   *wa*   *jinja-ya*

Japanese.people   TOP   HON-New.Year-LOC   TOP   Shinto.shrine-YA

o-tera-ni   *iku.*

HON-temple-LOC   go

''On New Year's Day, Japanese people go to Shinto shrines and/or[53] Buddhist temples.'

TEST: He said that Japanese people go to Shinto shrines or Buddhist temples, **that is (to say)**, to **any of** Shinto shrine **and** Buddhist temples. = Free alternative reading (due to the habitual aspect)

(6.29) *Youchien-ya*   *hoikusho-ni*   *kayou ko-o*   *motsu*   *hogosha*

kindergarden-YA   nursing-LOC   attend kid-ACC   possess   guardian

yaku   6500-nin   *no*   *yaku 4-wari-ga*   *"san-nin*   *nori-o*

about 6500-CLF   LK   about 4-ten.percent-NOM   "three-CLF   ride-ACC

shitatameru   beki   da"   *to*   *kaitou.*

approve   IMP   COP"   QT   answer

'About the 40% of the approximately 6,500 parents who have children attending kindergartens or nursery schools answered "(Riding with) three people on the same vehicle should be allowed.'

TEST: He said that they interviewed parents who have children attending kindergartens or nursery schools, **that is (to say)**, attending **any of** kindergartens **and** nursery schools. = Free alternative reading (due to the habitual aspect)

---

[53] Interestingly, in this context, both connectives are possible, even if *or* sounds more natural (cf. Ariel and Mauri, forthcoming).

(6.30) *[Shashin-o   shumi to   shitari,   satsuei- ga   shigoto dattari suru]*

picture-ACC   hobby as   do:TARI   photographing-NOM job   cop:TARI do

hito   nitotte, kamera ya   renzu wa   mochiron,   sankyaku-ya undai

people for   camera-YA   lens   TOP   of course   tripod-YA   platform

mo   hijō-ni   jūyōna   kizai.

also   extremely   important   gear

'For those who take picture as a hobby and/or for those who take picture as a job, camera and lens, of course [are important gears], also tripod and camera platform are very important gears.'

TEST: He said that the above holds for those who take picture as a hobby or for those who take picture as a job, **that is (to say)**, for **both of** those who take picture as a hobby **and** those who take picture as a job. = Separative conjunction reading

(6.31) *"Shinkon ryokō"   wa   "honeymoon trip"   toka   "honeymoon" to   iimasu.*

Honeymoon trip   TOP   honeymoon   trip   TOKA   honeymoon   QT   call:POL

'[In English] "Honeymoon trip" is called "honeymoon trip" or "honeymoon".'

TEST: He said that "Honey moon trips" is called "honeymoon trip" or "honeymoon", **that is (to say)**, it is called **any of** "honeymoon trip" **and** "honeymoon". = Free alternative reading (due to the habitual aspect)

In the light of this, we can now say that *ya/tari/toka* do not simply convey non-exhaustivity or, to follow the model designed by Ariel and Mauri, the HLC reading, but also exhaustive readings, namely, separative conjunction and free alternative readings, depending on the modality (e.g., irrealis vs. realis) or aspectual value (e.g., perfective vs. imperfective) of the utterance.

Using the test indicated by Ariel and Mauri, we can now calculate and schematize the overall distribution of both readings in our corpus as follows:
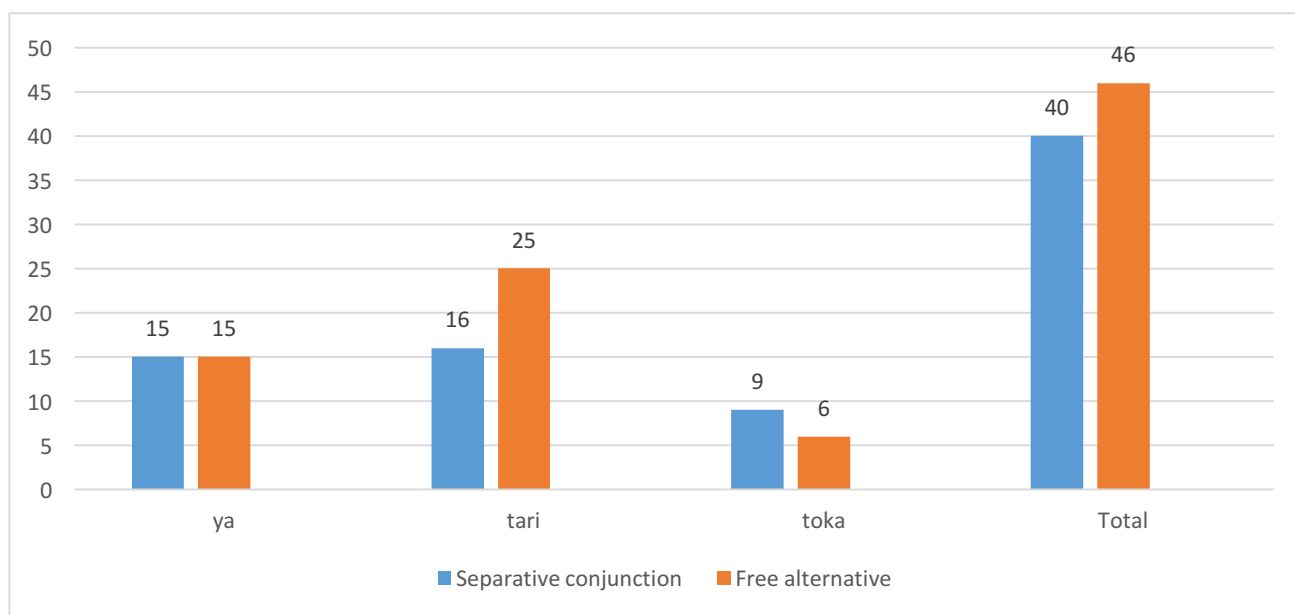
Figure 6.3: Distribution of Separative conjunction and Free alternative readings

At this point, some further questions arise: 1) why do speakers choose to convey these readings by means of *ya/tari/toka* and not through other available exhaustive connectives? 2) are these readings also available for other Japanese connectives?

In the previous sections, we have seen that *to* is incompatible with these readings. The reason stated by Ichikawa (1991: 62) is that the usage of *to* conveys the cumulativeness of the items introduced. Therefore, *to* would imply that all Japanese go to the Shinto Shrine but also to the Buddhist temple in (6.14) and that the interview regarded parents with children attending kindergarten and nursery school (that is, with two or more children) in (6.15)[54]. In other words, *to* cannot convey (and thus it is incompatible with) the distinctness at the core of conjunctive alternative readings. This fact is not surprising if we consider that *to* is firstly a comitative marker (Chino 2001: 33-34), also used to encode a specific type of combination relation. It is thus possible that in its functional extension, *to* has ended up covering the type of combination relation which is more similar to the comitative function, that is, a combination relation in which the conjuncts occur together, in the same place and/or at the same time. To test this assumption, we examined 50 occurrences of *to* taken from random samples in the Japanese section of the LCC (cf. section 2.2.1). While this corpus is not enough to conduct a concrete investigation regarding the functional space of

---

[54] Considering how the sentence is formulated, it is possible that native speaker would convey this interpretation using a different construction, instead of simply replacing *ya* with *to*. Moreover, the interpretation itself (that is, having two children attending both) seem less natural, therefore it would be likely emphasized by means of expressions such as *ryouhou* "both".

to (which is quite beyond the scope of our analysis), it still reveals some interesting insights. For instance, *to* tends to express a type of co-occurrence which can be often paraphrased as 'X alongside Y'. This is true not only for those utterances with a precise location in time and space (54%), but also for atemporal utterances in which the same nuance of meaning persists (20%), as shown in (6.32) while describing the beautiful things in Okinawa islands.

(6.32) *Aoku   tōmeina   umi-**to** odayakana   kikō,   rekishi-ga   tsuchikatta*
        blue:ADV   transparent:adj sea-TO calm:ADJ   climate history-NOM   cultivate:PAST
        dokutokuno   bunka.
        uniqueness:ADJ   culture
        'Blue transparent sea, calm climate and a unique culture cultivated by history.'

Moreover, in 13 occurrences (26%), there are linguistic expressions which stress this co-occurring reading. For instance, *kyoudou de* "jointly" (which also means "doing together something as equals"), *onaji ni* "at the same time", *taiou* "coping with"; but also verbs which commonly take a comitative complement such as *kyouryoku suru* "to cooperate/collaborate", *daku* "to embrace", *kousa suru* "to intersect".

While we do not pretend our analysis to be comprehensive, this potential tendency seems to confirm that *to* as a connective conveys cumulativeness, which stands in opposition to the distinctness needed to express separative conjunction and free alternative readings. Moreover, this strong connection with the comitative function can also explain why *to* is incompatible with non-exhaustive readings.

The above is likely to hold also for the *te*-form of the verb, albeit with the obvious differences due to the different scopes. In this regard, it is noteworthy that the *te*-form has also been referred to as the "gerund marker" or the "gerundive form" (Martin 1975: 475) of Japanese, since it can be combined with other verbs to create aspectual distinctions (e.g., progressive or resultative, *-te iru/aru*) or complex expressions (*-te miru* "to try to do something").

When it is used alone, it mainly functions as a conjunctive connective with different nuances of meaning. For instance, it can mark 1) sequential combinations, where the events are located along the same time axis at successive points, so that they are interconnected as part of the same overall sequence of events; and 2) simultaneous combinations, where the events overlap, as they are located at the same point along the time axis (cf. Mauri 2008: 84). The fact that *te* codifies events that are parts of the same overall sequence (or, we may

say, of the same overall frame) is supported also by its usage to express the manner as shown in (6.33) or the cause as shown in (6.34).

(6.33) *Arui**te** iku.*
  walk:te go
  '[I] go (by) walking.'

(6.34) *Tenki-ga   waruku**te**,   Fujisan-ga   miemasen*
  weather-NOM be.bad:te   Fuji-NOM   be.seen:POL:NEG
  'Because the weather is bad, Mount Fuji cannot be seen.'

While it is attested that the *te*-form can also mark atemporal relations (that is, in which the location of the events along the time axis is simply not relevant to the combination, cf. Mauri 2008: 85), in our corpus of 50 occurrences, there is only one occurrence of this type and it still conveys a nuance of cumulativeness, as both events cannot be valid separately:

(6.35) *"Tifanī-no   tēburumanā" toiu   hon-ga   at**te**,   ehon   mitaina katachi de*
  Tiffany-GEN   table manner QT   book-NOM   exist:TE book picture like   form   STR
  tēburumanā-o   oshiete   kureru n   desu.
  table-manner-ACC   teach:GRD   give   NML   COP:POL
  'There is a book called "Tiffany's table manners" and it teaches table manners through (forms like) book pictures and the likes.'

However, in the wide majority of occurrences, the *te*-form seems to encode events which are parts of the same narrative frame, and therefore strongly interconnected to each other, like in the following example:

(6.36) *Aite-o   sagashite,   tsukiat**te**,   uwakisarete,   wakareru.*
  partner-ACC   search:te   date:te   cheating:do:PASS:te separate.
  '[You] look for a partner, you date, you get cheated, you separate.'

While we believe that our corpus is not sufficient for a comprehensive analysis, it still reveals a tendency which can explain why the *te*-form is incompatible with separative conjunction and free alternative readings.

We have sketched some reasons as to why the canonical combination markers of Japanese are incompatible with the conjunctive alternative readings, unlike the English

combination marker *and*. However, since we are considering readings that firstly presuppose the mention of some alternatives, it is even more essential to investigate the compatibleness with the canonical alternative marker of Japanese, namely *ka*.

First, it is noteworthy that beyond expressing logical disjunction, *ka* is also (and foremost) an interrogative marker. It is used to mark questions at the end of the sentence and to form indefinite pronouns (suffixed to wh-words). Moreover, it is also commonly used to express doubt and uncertainty (cf. Kaiser 2001, Chino 2001, Ohori 2004).

For our investigation, we have tried to collect at least 50 occurrences of *ka* used specifically as a disjunctive marker in constructions which can be paraphrased as 'X or Y' (regardless whether X and Y are nouns or verbs). However, we have managed to collect only 20 occurrences (out of nearly 400 occurrences checked). The clear majority of *ka* occurrences in the Japanese section of the Leipzig Corpora Collection is used to mark questions or indefinite pronouns. Moreover, a quite frequent pattern is 'X *ka dou ka*', that is "whether X or not", where the addressee is asked to make a choice between X and non-X.

A similar figure of low frequency was obtained checking the Japanese 2011 Web corpus on Sketch Engine (out of a total of around 600 occurrences of *ka* checked).

Beyond its lower frequency, the actual usage of *ka* as a disjunctive marker appears to be extremely restricted also in the types of disjunctive readings it can convey. Using the schema provided by Ariel and Mauri, *ka* appears to encode frequently would-be (exhaustive) exclusive alternatives (pp. 22-27). In other words, it marks the mutual incompatibility between the explicit alternatives, as shown in (6.37) and (6.38).

(6.37) *Kōho*   *wa*   *Ohashi **ka***   *Taniguchi to*   *omotteiru*.
Candidate   TOP   Ohashi KA   Taniguchi QT   think:STA
"(we) are thinking about Ohashi **or** Taniguchi as a candidate."

(6.38) *Mamonaku kono hotondo-ga*   *Facebook Connect*   ***ka***   *Google no*   *Friend*
before long this   mostly-NOM   Facebook Connect   KA   Google gen   Friend
Connect-ni   *akewatasareru*   *koto*   *ni*   *naru*   *darou*.
Connect-DAT   surrender:do:PASS   NML   DAT   become   MOD
'I think it won't be long before most of this will be surrendered to Facebook Connect **or** Google's Friend Connect.'

From our brief investigation, it appears that *ka* is always the preferable choice whenever the speaker commits only to one option. There can be doubts and uncertainty (see the

Choice *or* reading, Ariel and Mauri forthcoming), but at the end of the process, only one alternative is the case (see also Ohori 2004: 57-58).

Nevertheless, we have managed to find one occurrence of *ka* used to express a conjunctive alternative reading:

(6.39) 6   *Gurūpu-ga*   *geneki*   ***ka***   *akaji*   *de[su].*

     6   group-NOM   decrease in profit   KA   deficit   COP:POL

    'Six groups are decreasing their profit or they are in deficit.'

Here, the author describes six companies facing economic difficulties: some of them register a decrease in their profit, others are in deficit. Therefore, both alternatives are the case, but in separate situations (that is, for different companies).

The fact that *ka* is compatible with these conjunctive readings is not an issue for our analysis regarding *ya/tari/toka*, since the same reading can be compatible with two or even more markers. For example, we have seen that the separative conjunction reading in English is compatible both with *or* and with *and*.

We argue that the reason for this functional overlapping lies in the diachronic developments of *ka*[55] and *ya*. In Old Japanese, *ka* and *ya* were interrogative markers in competition (cf. Greenberg 1987, Frellesvig 2010). Over time, especially through Middle Japanese, they have changed repeatedly their functional spaces to avoid overlapping. For example, in Old Japanese, *ya* was used in rhetorical questions or exclamations, such as in *ari ya nasi ya* "is there or is there not?". By the end of Late Middle Japanese, *ya* as sentence interrogative has declined in favor of sentence final *ka*, developing at the same time the uses as a connective (Frellesvig 2010: 359). Nevertheless, as Martin (1975) notes, some traces of the past functional overlapping between *ka* and *ya* persist in Modern Japanese in fixed expressions, as shown in (6.40) and (6.41).

(6.40) *Are*   *ka*   *kore*   *(ka)*

    that   KA   this   (KA)

    'That and this' (Martin 1975: 157)

---

[55] Consequently, this affects also the functional space of *toka*, since it is a combination of the marker *to* and the interrogative/disjunctive marker *ka*, as we have seen in the introduction (cf. section 1.3.3.4).

(6.41) *Are   ya   kore   (ya)*
          that   YA   this   (YA)
          'That and this' (Martin 1975: 157)

The expressions in (6.40) and (6.41) are perfectly equivalent and they are also instances of conjunctive alternative readings. This interpretation is even more evident by the fact that Martin translates both using the English connective *and*, instead of *or*, since they are both compatible (*this and that*, *this or that*).

Therefore, while *ya* and *ka* have separated their functional spaces over time, with *ya* losing most of its interrogative value, it appears that conjunctive alternative readings are the only points of the functional spaces where they still overlap with each other, a clear trace of their common past as irrealis markers.

This intrinsic irrealis value likely explains also the compatibility of *tari* with the conjunctive alternative readings, since, as Narrog notes, "Modern Japanese *-Tari* is not specifically an irrealis subordinate mood, but it does lead to a lower factuality of the event portrayed" (2012: 147).

It thus emerges a connection between conjunctive alternative readings and markers that can convey irrealis values. Moreover, this also explains why markers such as *to* and the *te-form* are incompatible with these readings: as we have seen above, both express clear realis (and often strictly temporal) values and a high factuality of the items/events portrayed. This is not true for the English *and*, which is much more general and less binding in its semantic implications.

To describe the findings of this section, we use the Semantic Map method (cf. Croft 2001, Haspelmath 2003). As Haspelmath (2003) explains, "A semantic map is a geometrical representation of functions in 'conceptual/semantic space' that are linked by connecting lines and thus constitute a network." (2003: 213). In other words, it consists in displaying the different functions of a marker (or different markers at the same time, as in our case) on a geometrical space, to describe the relations occurring among them. More specifically, two functions are connected only if there exists at least one language that has the same formal expression for the two functions and at least one language that has different formal expressions for the same two functions (which means that they should be considered two different but linked functions, cf. Haspelmath 2003: 216-217).

Figure 6.4 and Figure 6.5 represent the semantic maps of the combination/alternative space in English and Japanese.
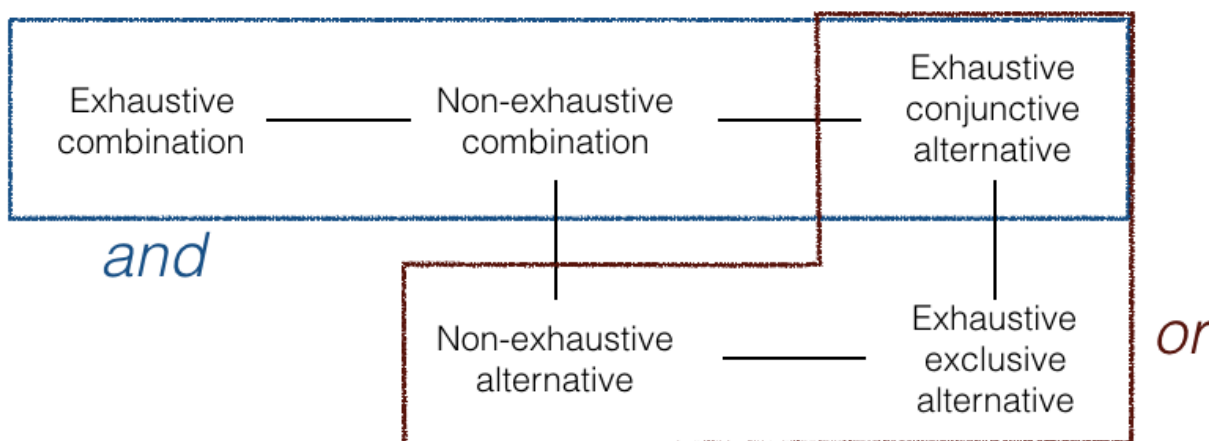
Figure 6.4: Semantic map combination/alternative in English
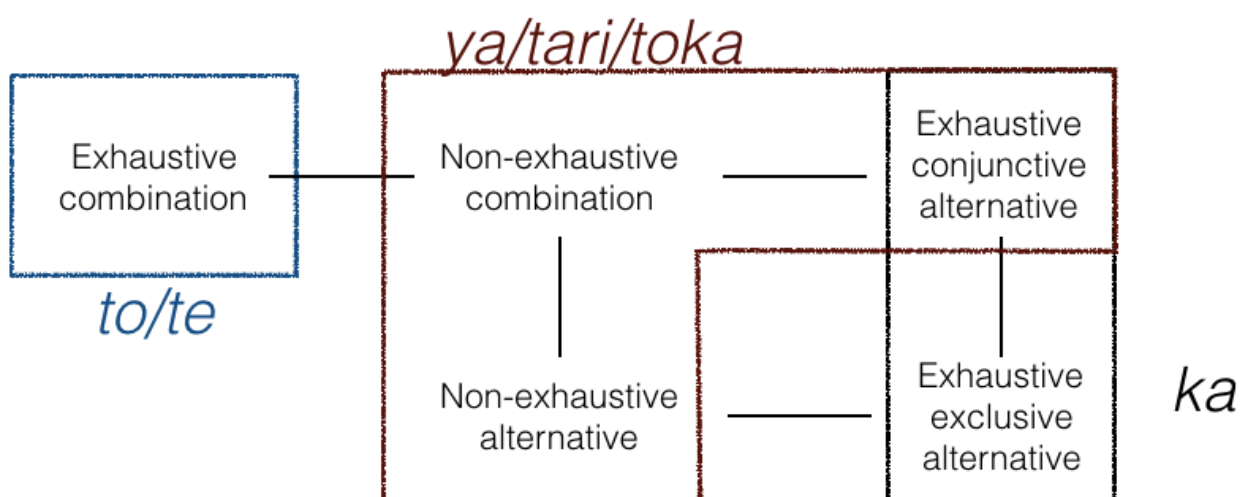
Figure 6.5: Semantic map combination/alternative in Japanese

Despite the overlapping in conjunctive alternative readings, the usage of *ka* as a logical connective still appears very restricted in frequency compared to that as an interrogative marker. This likely explains why speakers decide to use other strategies to convey these readings.

More generally, considering all the issues described above, it emerges that Japanese exhibits a wide range of connective makers, but only one, namely *ya,* is a dedicated connective maker[56]. Most of these connectives exhibit other functions (e.g., comitative

---

[56] As we have already seen, not even *tari* and *toka* are dedicated connectives, since they are also used as general extenders and therefore, can also perform functions that do not seem compatible with connectives

marker, interrogative marker, gerundive form, etc.), which are often the predominant and/or the more frequent functions. Moreover, in some cases, these functions seem also to influence the kind of readings the markers can express as connectives (e.g., *to*). This means that even when these markers are used as connectives, their functional space is very restricted compared to those of their English counterparts.

6.1.2 THE OCCURRENCE OF SEPARATE EVENTS

In the first section of this chapter we have seen that, particularly regarding *tari*, some of these separative conjunction and free alternative readings have been reported as instances of iterative or distributive sentences, thus referring to aspectual distinctions. Considering this fact, we should now question how these alternative conjunctive readings may give rise to aspectual interpretations.

Coordination relations can be described along two dimensions: cooccurrence and non-cooccurrence (Mauri 2008: 80). Combination is located on a cooccurrence dimensions, where the linked elements are declared as co-occurring. On the contrary, alternative is located on a non-cooccurrence dimension, where the elements constitute replaceable non-cooccurring possibilities. We schematize this conceptual opposition in Figure 6.6. Co-occurring elements are located along a horizontal dimension, one after the other. Conversely, non-co-occurring elements are located along a vertical dimension, one above the other, since they are assumed as replaceable possibilities.

(e.g., hedging functions as described in chapter 5, intensification of the negative polarity as described in section 6.4).

he ran out
of gas

OR

he got lost
along the way

| he took the car | AND | he got stuck in the traffic | AND | he arrived late for the date |
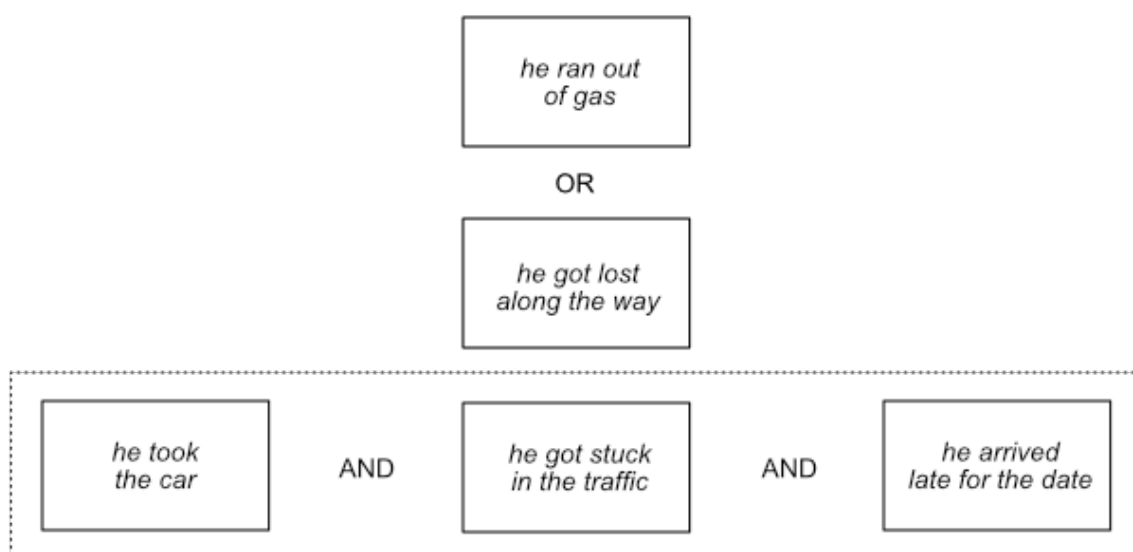
Figure 6.6: Co-occurrence vs. non-co-occurrence opposition (cf. Mauri 2008: 81)

Let us focus on alternative readings in which the speaker commits to only one alternative being the case. Since the alternatives are conceptualized as equivalent possibilities, this also means that given a slot 'X' in a possible world, it can be occupied by only one of the mentioned alternatives. As noted in section 2.3.3.2, it follows that until a choice is made or the speaker comes to know which hypothesis is actually realized, either alternative could be the non-occurring one and therefore, for instance, both can be conceptualized as irrealis. This is exactly what happens in Japanese, where *ka*, an irrealis marker, highlights that the alternatives are only a set of paradigmatic possibilities, and only one of them will be situated on the temporal axis.

For example, consider again the example (6.37) in the previous section, repeated here as (6.42).

(6.42) *Kōho*   *wa*   *Ohashi **ka***   *Taniguchi to*   *omotteiru*.
Candidate   TOP   Ohashi KA   Taniguchi QT   think:STA
"(we) are thinking about Ohashi **or** Taniguchi as a candidate."

In this case, only one between Ohashi and Taniguchi will be the candidate. The other alternative will not be realized. Therefore, for now, they are only possibilities located in the paradigmatic axis, and not in the temporal axis of realis occurrences.

In the previous sections, we have seen that in some alternative readings, the speaker commits to both the mentioned alternatives being the case, in the sense that they are both realized (i.e., separative conjunction reading) or at least potentially realizable (i.e., free alternative readings). In this cases, the alternatives must be construed as (at least potentially) occurring while at the same time stressing the independence of each of the alternatives. To accommodate these readings, we need a further dimension, namely, the occurrence dimension, which concerns events which are (or may be) realized, but in separate situations (i.e., without co-occurring). In other words, we are now investigating the actual occurrence of non-cooccurring events.

This is also the reason why these readings are metaphorically situated between 'And' and 'Or' (that is, between combination and alternative): they represent the case in which the alternatives get all to be situated in the dimension of the (potential) occurrence (thus like in combination relations), despite being not co-occurring at the same time or in the same situation (thus like in alternative relations). Considering this, it is not surprising that some languages (e.g., English) can convey these readings by means of both the alternative marker and the combination marker. Moreover, it is also interesting that other languages use connectives which (metaphorically) stand in between this division, covering some readings of combination domain and some readings of alternative domain, such as *ya/tari/toka* in Japanese.

Nevertheless, the process of locating all the alternatives in the occurrence dimension can give rise to different situations depending on the parameter of temporality. Based on this parameter, we can distinguish three situations: 1) irrealis utterances, 2) realis atemporal utterances, 3) realis temporal utterances.

In the first situation, the utterance is not a fact (at least yet) and therefore still belongs to the irrealis domain. This means that even if the speaker commits to the possibility of each of the alternatives to be realized, in the moment of the utterance, none of them has been realized yet. In other words, they are still in the non-occurrence dimension, as shown in (6.43).

(6.43) *Hyōjisareru*   *shashin-no*   *setto-o*   *susumetari*   *modoshitari*   *dekiru.*
       display:do:PASS   picture-GEN   set-ACC   advance:TARI put.back:TARI POT
       'You can put forward or put back the set of pictures that is displayed.'

Consequently, the process of locating them in occurrence dimension is not at issue here.

In the second situation, the entire utterance is a fact (i.e., realis mood), but it does not exhibit a clear collocation in the temporal axis (i.e., atemporal utterances). Consequently, the linked events do not need to be precisely located either, with respect to each other. For example, consider the following sentence from our corpus.

(6.44) *Shikashi konoyō-na rīdāshippu   wa,   yokkyū-fuman-o   mitasu shudan-o*
        but   such-ADJ   leadership   TOP   frustration-ACC   satisfy means-ACC
        ushinattari, aruiwa   esukarēto   shisugita   toki-ni, hōkai   e   no
        lose:TARI   or   escalate   overdo:PAST when collapse to   NML
        michi-o   tadorimasu.
        path-ACC   pursue:POL
        'However, when it has lost the means to satisfy the frustration or it has escalated too much, such a leadership pursues the path towards the collapse.'

By means of the questionnaire, we have asked to native speakers to select the best paraphrases for this sentence. According to the majority of the informants (93,1% in the Japanese version of the questionnaire, 100% in the English version), this sentence can be well paraphrased as:

(6.45) *Ika no futatsu no baai ga aru. Yokkyū fuman o mitasu shudan o ushinatta toki, matawa, esukarēto shi sugita toki.*
        'There are two cases as follows. When it has lost the means to satisfy the frustration, or when it has escalated too much.'

Therefore, the utterance describes two events construed as alternatives. Since the events are not co-occurring, they might happen in the same country but at different times (which means that the events may alternate), or at the same time but in different countries/situations (which means that the events are scattered). Both interpretations are possible, but understanding precisely which one is the valid interpretation is not required, since the entire utterance is atemporal. The only important point is to stress that both case scenarios are valid, albeit they are not co-occurring. In this regard, the first two situations (i.e., irrealis utterances and realis atemporal utterances) are very similar.

Finally, in the third situation, all alternatives are realized and located along the temporal axis. This process is schematized by Figure 6.7.
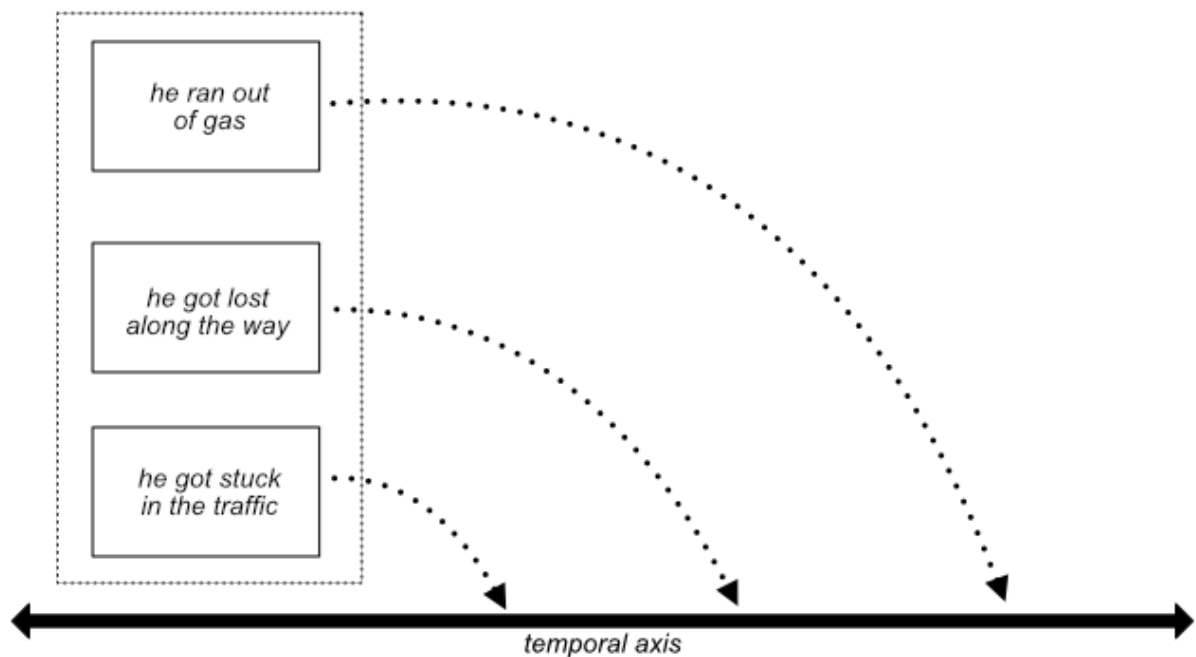
Figure 6.7: Alternative conjunctive readings and the temporal axis

In this case, it becomes pivotal to understand how the events relate to each other with respect to their position on the temporal axis, as it is requested by the utterance itself. In fact, the linked alternatives need to be construed as both real, but not jointly real (therefore unlike combination relations). Therefore, if we need to consider the temporality parameter while stressing the independence of the alternatives, the previous statement also means that at the time in which one of the alternatives is real, the other is not, and vice versa. For example, in the sentence "the tightening knob for fixing or freeing", fixing and freeing are both real, but when fixing is real, freeing is not, and vice versa.

This process of locating the 'both real but not jointly real' elements in the temporal axis may give rise to further interpretations. Let us consider this example from my corpus.

(6.46) *Asou   shushō   wa,   yuruyakana   sakugen-haba-o   shuchōsuru*
       Aso   Prime Minister TOP   gradual   reduction-ACC   request:do
       keizaikai-ya   rōdōkumiai   no   daihyō   aitsuide
       business-world-YA   labour.union-GEN   representative subsequently
       kaidan shi,   iken-o   kiita.
       conversation:do:GRD opinion-ACC   hear:PAST
       'Prime Minister Aso listened to the opinions and met in succession the representatives of
       labour unions and of the business community who request a gradual reduction.'

Through the questionnaire, we have asked to native speakers to select the best paraphrases for this sentence. According to the majority of the informants (88,5% in the Japanese version of the questionnaire, 66,7% in the English version), this sentence can be paraphrased as:

(6.47) *Shushō wa keizaikai no daihyō to kaidanshi, sorekara rōdō kumiai no daihyō to kaidan shita.*
'The prime minister met with representatives from the business circle, and then met with representatives of labour unions.'

In this case, the key word is *aitsuide* "subsequently", which means that the Prime Minister did not meet them together, but separately, one after the other. This means that in order to be 'both real but not jointly real' in the temporal axis, the event is iterated at least twice, with different participants. Therefore, the final interpretation of the utterance can be one of iteration.

This iterative interpretation is even more strong when the events stand in opposition semantically, like in the example (6.10), where the addressee can construe the two events (i.e., fixing and freeing) as iterating in alternation; or when the speaker provides some explicit aspectual adverbs.

In other cases, a distributive interpretation prevails. Here, the distinctness is given not by alternation through time (i.e., iteration), but by the elements being scattered or separated in different groups. For example, consider the following occurrence from our corpus:

(6.48) *Kashiwazaki-shi de wa 59 tōhyōsho no uchi 18 tōhyōsho-ga hinansho*
Kashiwazaki-city LOC TOP 59 polling place out of 18 polling place-NOM shelter
to "dōkyo".
with coexistence.
Tsuitate-de hinansho-no ikkaku-ni tōhyō-basho-o kakuhoshitari,
screen-STR shelter-GEN corner-LOC voting place-ACC guarantee:do:TARI
betsu-no heya-ni setsueishitari shita.
different-ADJ room-LOC construction:do:TARI do:PAST
'In Kashiwazaki city, out of 59 polling places, 18 of them "coexisted" with the shelters. They guaranteed voting places in a corner of the shelters through screens or they set up (the voting places in) different rooms.'

In the sentence above, the author refers to an election taking place after an earthquake. In several locations, people were forced to set up polling places within the shelters. For instance, regarding 18 polling places out of 59 in Kashiwazaki, in some of them, they set up the polling place in a corner of the shelter using a screen, in others they set up a different room to vote. Both solutions have been activated, but in different polling places.

(6.49) *Shutsudai   misu   wa   izuremo kokugo-no   mondai   de,*
  question   mistake TOP   all   national language-GEN   problem   COP:GRD
  seitō-ga   sentakushi-chū-ni   nakattari,   fukusū   attari   shita.
  correct.answer-NOM   choice-among-LOC   exist:NEG:TARI several exist:TARI   do:PAST
  'All question errors were problems of language; the correct answer was not available among the alternative options or there were several correct answers.'

Here, the author refers to the fact that a multiple-choice test was poorly written, with several mistakes. For instance, in some questions, the correct answer was not available and all the options were actually wrong, while in other questions, there were several correct answers.

Similarly, we can also analyse the example provided by Alpatov (1997: 393) as instance of distributive aspect:

(6.50) *Soko-ni   wa   mata danjo-no   komodo-tachi-ga   nanninmo*
  There-LOC   TOP   again male.and.female-GEN child-PL-NOM   as many as come
  oyoidari   moguttari   shiteita.
  swim:TARI   dive:TARI   do:STA:PAST
  'Again as many boys and girls were swimming and/or diving there.' Alpatov (1997: 393)

The actual meaning of this sentence is that some boys and girls were diving and some other boys and girls were swimming. This interpretation arises from the fact that the acts of swimming and diving need to be construed as 'both real, but not jointly real'. It follows, that the two actions need to be scattered between two different groups of people. In this sense, *tari* is not an aspectual marker, but only a connective that stresses the independence of the mentioned events.

To conclude, we propose that the aspectual readings that have been reported regarding *tari*, are not the result of *tari* functioning as an aspectual marker. On the contrary, as pointed out by Narrog (2012), *tari* does not hold any intrinsic aspectual value anymore, to the point

of simply referring to events in a very general and undetermined way (which probably facilitate readings such as Higher-level category and Free alternatives).

Nevertheless, separative conjunction and free alternative readings may give rise to further aspectual readings whenever the linked events or elements need to be situated on the temporal axis (that is, realis temporal utterances). In these cases, the necessity to stress the elements as occurring albeit independent entails the construal of the events as alternating along the temporal axis (i.e., iterative readings) or scattered in different situations or groups (i.e., distributive readings). Moreover, this fact is true not only for *tari*, but also for the other connectives that convey these conjunctive alternative readings, such as *ya* and *toka*.

6.2 EXEMPLIFYING CONSTRUCTIONS IN NEGATIVE EXHAUSTIVE CONTEXTS

Japanese exemplifying connectives are not the only type of exemplifying constructions that can be used in exhaustive contexts. More specifically, it appears that some Japanese general extenders perform an intensifying function in negative utterances.

Consider the following exchange from our corpus:

(6.51) A: *Hitoritabisuru   saini   anzen'na kuni   wa doko   deshou   ka?*
           travelling.alone:do   in case of   safe   country TOP where COP:MOD Q
           'Where is a safe country to travel alone?'

     B: *Zettaini   anzen'na   kuni   nado   arimasen.*
           Absolutely   safe   country NADO exist:POL:NEG
           'There is no absolutely safe country.'

     B: *Dokodemo   anata-no   kōdō   shidai   desu.*
           anywhere   you-GEN   action dependent upon   COP:POL
           'Anywhere it depends upon your actions.'.

Here, author A raises a question in a message board regarding a safe country where to travel alone. Author B replies that there is no such a thing as an absolutely safe country in the world, but it always (*dokodemo* "anywhere") depends on the traveller's actions.

Interestingly, in the exchange above, the function of *nado* in the second line is not exemplification. In other words, it cannot be paraphrased as 'safe country and things like that', with "safe country" as an example of the wider inferred category. On the contrary, here *nado* is used to emphasize the fact that there are no safe countries at all, that is, to

emphasize the negative polarity of the utterance. In our corpus, there are three such instances.

Under this regard, several studies (cf. Morita 1980, Lee 2004, Chen 2005, Sawada 2016) and some grammars targeting L2 learners (e.g., Chino 2001: 44) have acknowledged that, beyond exemplification (which is its main function), *nado* can also be used to give emphasis to negative evaluations. However, among these studies, only Lee (2004) tries to explain the relationship between the different functions of *nado*, pointing out that they depend on the speaker's modality (positive/neutral vs. negative).

Moreover, some of them (e.g., Sawada 2016) briefly focus also on the relation between this function of *nado* and the usage of emphatic negative-polarity items such as *totemo* "very", which denotes a high scalar value. In fact, negative polarity items (e.g., *zettaini* "absolutely", *totemo* "very", *zenzen* "not at all" with negative verbs) seem to frequently co-occur with *nado* whenever the latter is used to intensify the negation of the utterance. This fact is also confirmed by the attested occurrences in our corpus, as shown in (6.51) with *zettani*, (6.52) with *totemo* and (6.53) with *chittomo* "not at all".

(6.52) *Ruiseki-akaji*   *wa*   *4-oku-en.*   *Akaji* *da*   *kara,*   *atarashī*

accumulated deficit TOP   400 million yen.   deficit COP   because   new

iryōkki   *nado,* *totemo kaenai.*   *Son'na*   *jōtaidakara,* *da*

medical equipment NADO very   buy:POT:NEG such   situation   COP

kara   *wakai ishi* *wa*   *kita*   *garanai.*

because   young doctor TOP   come:PAST feel:NEG

'The accumulated deficit is 400 million yen. Because of the deficit, I couldn't (ever) buy medical equipment. Because of such situation, young doctors do not want to come.'

(6.53) *Keiki*   *nado* *chittomo*   *yoku* *nai!*

economy   NADO not at all   good AUX:NEG

'The economy is not good at all!'

In this regard, Sawada (2016) notes that it is natural to use *totemo* with the particle *nado* instead of other case markers, because *nado* "signals that the given proposition/event is currently under discussion and that the speaker construes it negatively" (2016: 2).

Therefore, following Sawada's insights, it appears that, in such instances, *nado* functions like a focus marker, thus similarly to the topic marker *wa*, which often replaces the

nominative marker to mark the subject in negative utterances to create a contrastive effect (cf. Iwasaki 2013: 243-247).

Moreover, in some cases, the usage of *nado* to intensify the negative polarity of a statement seems to give rise also to further readings. For example, in (6.52) the combination of *nado*, *totemo* and the negative modal statement (i.e., *kaenai* "I cannot buy") emphasises the impossibility of the situation (cf. Sawada 2016, Chino 2001: 44-45).

In other cases, the usage of *nado* gives rise to a pejorative reading, that is, it denotes the speaker's contempt toward the entity that is marked by it (cf. Martin 1975, Chino 2001, Suzuki 1998), as shown in the following example:

(6.54) *Tanaka-san   nado   wa,   totemo shachou-ni   wa   narenai.*
       Tanaka-hon   NADO   TOP   very   president-DAT TOP   become:POT:NEG
       (lit.) 'Mr Tanaka absolutely could not become the president of the company,'
       (id.) 'There is no way that anyone like Tanaka [that the likes of Tanaka] could become president of the company.' (Chino 2001: 44)

(6.55) *Aitsu   kara nado hanagami   ichimai   demo moraitakunai.*
       that guy from NADO tissue-paper one-piece   even   receive:DES:NEG
       'From the likes of him I wouldn't even accept a Kleenex.' (Martin 1975: 162)

Such instances are not attested in our corpus. Nevertheless, they are useful to understand the overall functional space of *nado*. In fact, Suzuki (1998) notes that the etymology of *nado* does not suggest any relationship with the concept of devaluation. On the contrary, *nado* originated from the combination of *nani* "what" and the comitative/conjunctive marker *to* (cf. Yamaguchi 1988). While its etymology can explain why *nado* is mainly used to denote a lack of referentiality (like when it is used to exemplify or as a hedging strategy), at first glance, it does not seem to explain the pejorative reading, and not even the function of intensifying the negative polarity. Nevertheless, Suzuki suggests that it is the ability to denote a lack of referentiality that may give rise to the pejorative effect. In fact, it suggests the lack of the speaker's willingness to commit to the attached item, which, in some context, may be interpreted as the speaker's contempt toward the item itself.

Therefore, the true functional core of *nado* is indeed the lack of referentiality, which also explains the strong connection between the belittling effect and hedging functions noted by Suzuki, since they are both evoked from the suggestion of a lack of referentiality (cf. 1998: 261).

At this point, we may therefore wonder if this statement is true also for the overall function of intensifying the negative polarity of utterance. In other words, is this function triggered by the ability of denoting a lack of referentiality?

There are two facts that seem to confirm an affirmative answer to this question.

First, in his grammar targeting L2 learners, Chino (2001: 44-45) tries to translate literally from Japanese to English the sentences with *nado* used to intensify the negation. To achieve this, he uses constructions like 'anyone like X', 'such a X', 'anything like X', which means that *nado* still preserves its functional core. Moreover, the connection between these constructions in English (e.g., 'the likes of X') and the use of *nado* has been highlighted also by Suzuki (1998: 267). The acknowledged link between these English constructions and the usage of *nado* denote the actual role of the latter in such instances, as well as a strong connection with the ability of denoting a lack of referentiality.

Secondly, Suzuki's claim arises also from the fact that other exemplifying strategies denoting a lack of referentiality can perform the same function, for example *tari* and *toka*. This fact is confirmed also by our corpus, where two such occurrences of *tari* and one of *toka* are attested.

(6.56) *Zettaini   anata-o   wasuretari   shimasen!!*
       absolutely   you-ACC   forget:TARI   do:POL:NEG
       'I will never forget you!!'

(6.57) *Mata, uketottemo, zettaini   henji-o   okuttari   shitewaikemasen.*
       also   receive:COND absolutely reply-ACC send:TARI   do:IMP:NEG
       'Also, even if you receive the mail, absolutely do not replay.'

(6.58) *Aitsu,   zenzen   kanojo   toka   inai   n   da   yo.*
       that guy   absolutely girlfriend TOKA   AUX:NEG   NML   COP   PP
       'That guy does not have a girlfriend at all!'

In all these utterances, *tari* and *toka* are used to intensify the negative polarity of the utterance. The only exemplifying marker that does not perform this function (among those here analysed) is *ya*. This is not surprising: as it was noted in chapter 5, *ya* does not perform hedging functions either. This lack was explained by pointing to the nature of *ya* as a connective, which cannot be used with less than two items. On the contrary, *nado*, *tari* and *toka* can be perform both functions, probably because they can be used with just one item.

Thus, as we pointed out in chapter 5, it is possible that both the hedging and the intensifying functions cannot be performed with more than one item.

The fact that almost all the exemplifying strategies here studied can be used to intensify the negative polarity suggests that there may be indeed some sort of connection between these two functions. In chapter 5, we have seen that the configuration of a set is the link between the exemplifying function and hedging functions. While in the first case the set is the actual focus of the discourse, in the second case, the set works as a fuzzy background to create the attenuate effect desired by the speaker (cf. section 5.1). Now, we argue that also in the case of intensifying the negative polarity, exemplifying markers behave similarly, configuring a set of similar items to the one mentioned which will be ultimately refuted in its entirety.
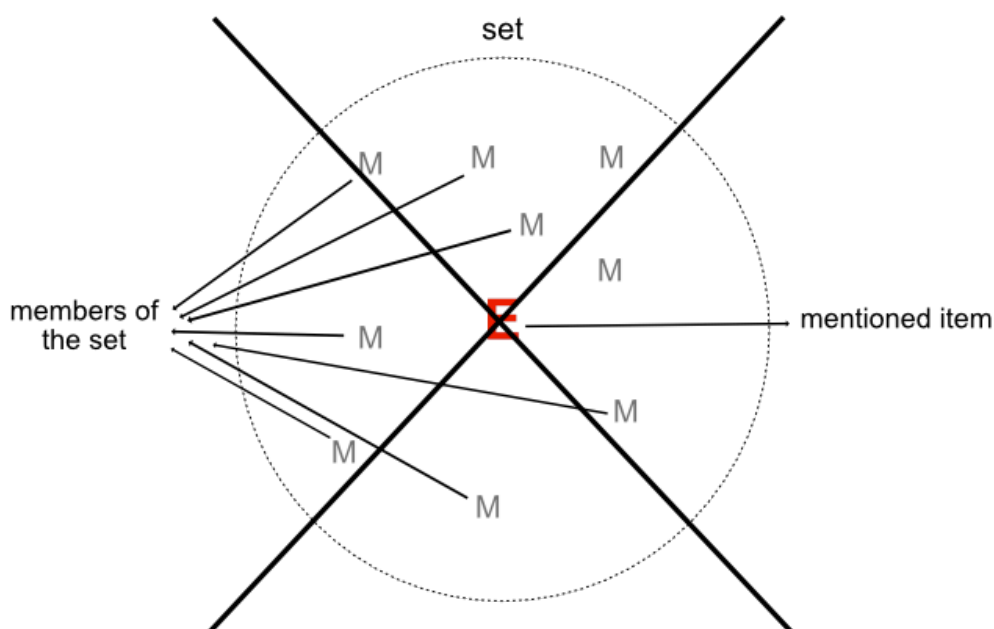
Figure 6.8: The roles of the example in intensifying the negative polarity of the utterance

In other words, not only the speaker is refuting the mentioned item, but also any other possible items similar to the one mentioned. The ultimately effect is an emphasis of the negation, because its scope has been extended to an entire category of similar elements. Therefore, for example, in (6.51), the speaker is not only negating the presence of a "safe country", but also the existence of something similar to "an absolutely safe country", broadening the scope of the negation. In (6.56), the speaker says that he or she won't do such a thing as to forget the addressee, that is, nothing of the likes. Etcetera.

This may seem at odd with the initial claim that these functions are related to exhaustive context. Nevertheless, it is not. The context is indeed exhaustive, because the speaker does

not intend to actually make reference to items other than the one mentioned (differently from some hedging functions, in which other possible options are indeed taken into consideration). Like in the cases of pragmatic hedging (cf. section 5.1), the set is simply a cognitive tool to obtain a certain communicative effect.

7. THE FUNCTIONAL SPACE OF EXEMPLIFICATION

7.1 INTRODUCTION

We started our investigation selecting four strategies used in Japanese to specifically encode exemplification (chapter 1). Then, we overturned our perspective, from a top-down to a bottom-up approach, thus assuming the possibility for these same markers to cover other functions beyond exemplification (chapter 2). The reason for this choice was *i*) to monitor the functional behaviour of the markers in the actual language, *ii*) to investigate potential connections among functions, and *iii*) to improve our understanding of exemplification at the linguistic level. Finally, the actual analysis (chapter 3, 4, 5, 6) has revealed a varied picture, where these markers expand their functional spaces sometimes even in interesting and unpredictable ways (e.g., the "exhaustive" functions examined in chapter 6).

At this point, we need to tie the loose ends together, trying to straighten up all the tendencies that have emerged. Therefore, in this final chapter, the functional domains of the examined exemplifying constructions are compared, described and schematized, in order to point out interesting correlations and ultimately to highlight potential functional and structural patterns regarding exemplification.

The comparison develops in two phases. First, in section 7.2, we will consider solely the synchronic level, describing the correlations and tendencies that have emerged during our investigation. Second, in section 7.3, we will briefly examine the diachronic level, to determine if the correlations attested at the synchronic level can be explained looking at the diachronic developments of the markers here under study.

7.2 THE FUNCTIONAL SPACE OF EXEMPLIFYING CONSTRUCTIONS

As noted in the introduction of this chapter, the analysis of Japanese exemplifying constructions has revealed recurring patterns of usage and tendencies.

The first main result of our investigation concerns the attested multifunctionality of these markers. In fact, we have seen that exemplifying makers can perform different types of functions: cognitive functions (i.e., the communication of contextually relevant categories),

pragmatic functions (i.e., hedging), semantic functions (i.e., conjunctive alternative connectives, the intensification of the negative polarity of the utterance).

To describe and understand this multifunctional situation, we use the Semantic Map method (cf. Croft 2001, Haspelmath 2003). As already noted in section 6.1.1, it consists in displaying the different functions of a marker (or different markers at the same time, as in our case) on a geometrical space, to describe the relations occurring among them. In this regard, Croft (2001: 93) proposes a further important distinction between a Semantic Map and a Conceptual Space: while the latter is meant to represent the network of functions of a specific domain and it is argued to be universally valid, the former is a language-specific representation of the Conceptual Space.

In our analysis, we prefer the term 'functional', as in 'Functional Map', instead of 'Semantic Map', since we are considering also functions that are pragmatic, and thus do not properly pertain to the semantic domain.

The Functional Map in Figure 7.1 schematizes the different functions covered by the Japanese exemplifying constructions and the relations occurring among them.
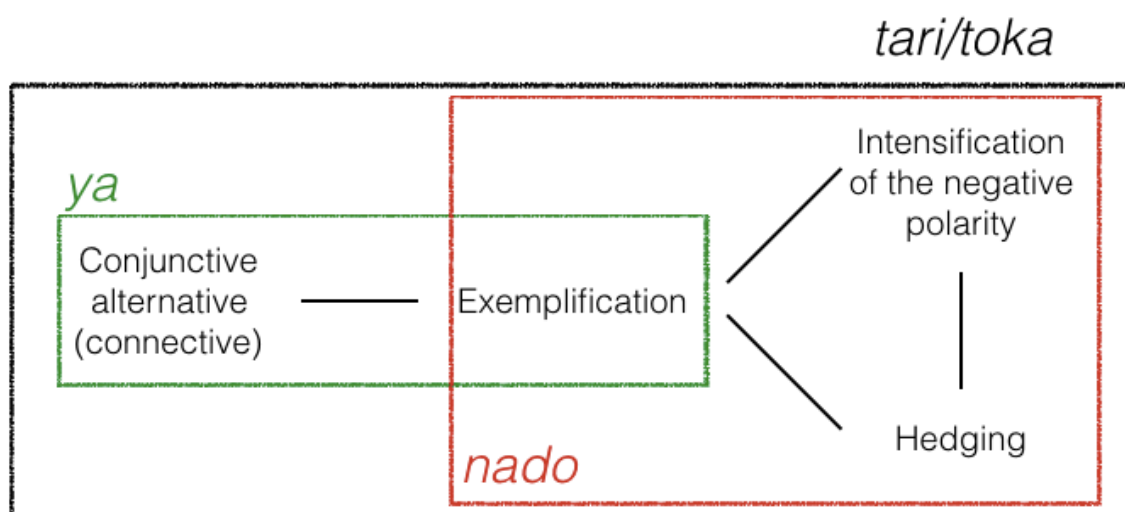
Figure 7.1: Functional Map of Japanese exemplifying constructions

The more significant pattern emerging from this Functional Map is that the functional domain strongly depends on the type of construction the markers belong to, that is, on the modality of usage of the markers. More specifically, non-exhaustive connectives (e.g., *ya, tari, toka*) cover some functions, while general extender (e.g., *nado, tari, toka*) cover other

functions, with the only common point being exemplification (and therefore, the coding of context-relevant categories).

In other words, the analysis has revealed a strong correlation between the structural level and the functional level. Additionally, the functional extension depends on a structural extension: non-exhaustive connectives that can be used also as general extender (e.g., *tari* and *toka*) can perform also the functions that are typically associated with general extenders (e.g., hedging, intensification of the negative polarity). It also appears that markers belonging to the same type of constructions perform the same functions.

This correlation is ultimately motivated by the fact that general extender can codify open-ended list with just one single example. As seen in chapter 5 and chapter 6, when only one example is mentioned, it is possible to use the profiled set as a fuzzy background, while two or more examples always trigger the deduction of a common integrator (Lang 1984), that is, the common property of the category (cf. chapter 4). The usage of the set as a fuzzy background appears to be a prerogative of general extenders and thus also the functions linked to this configuration (e.g., hedging, intensification of the negative polarity of the utterance) are not available for non-exhaustive connectives.

In this regard, the investigation has also revealed that some of these functions are related to each other (e.g., the use of the set as a fuzzy background, cf. chapter 5 regarding the correlation between exemplification and hedging, and chapter 6 for the correlation between the intensification of the negative polarity with hedging and exemplification).

Finally, focusing on the exemplifying function proper, some minor tendencies have arisen as well with regards to the presence or the absence of category labels, the type of categories that can be encoded (i.e., similarity-based or frame-based categories), the syntactic and semantic types of examples that are usually used (cf. chapter 3 and chapter 4).

In the next sections, these tendencies will be synthesized into two functional spaces, one dedicated to non-exhaustive connectives (cf. section 7.2.1) and the other to general extenders (cf. section 7.2.2).

These functional spaces are organized along two axes, which symbolize an increasing in the number of examples expressed by verbal phrases, to the detriment of the number of examples expressed by noun phrases. Therefore, moving along the axes, the continuum goes from markers that can be used only with noun phrases (i.e., *ya*) to markers that can be used only with verbal phrases (i.e., *tari*).

The reason for choosing the syntactic nature of the examples as the parameter along which the tendencies may develop, is empirical-based but, at the same time, also cognitively

motivated. In the first sense, our data have confirmed that the syntactic nature of the examples is the parameter that trigger most of the attested correlations (cf. section 3.3 and 4.3). Moreover, at the cognitive level, we can assume that the syntactic nature of the mentioned items is likely the first parameter determined by the speaker, when he or she formulates the message, and that all the others follow consequently. This is likely due to two reasons: 1) the markers under examination exhibit different behaviour regarding this parameter (cf. section 1.3.2) which thus need to be determined as soon as possible to choose the correct marker, and 2) the syntactic properties deeply influence the way the examples are elaborated and communicated (cf. Langacker 1987a, 1991b, Givón 2001 and section 2.3.2.2).

7.2.1 FUNCTIONAL SPACE OF NON-EXHAUSTIVE CONNECTIVES

Figure 7.2 represents the functional space of non-exhaustive connectives.
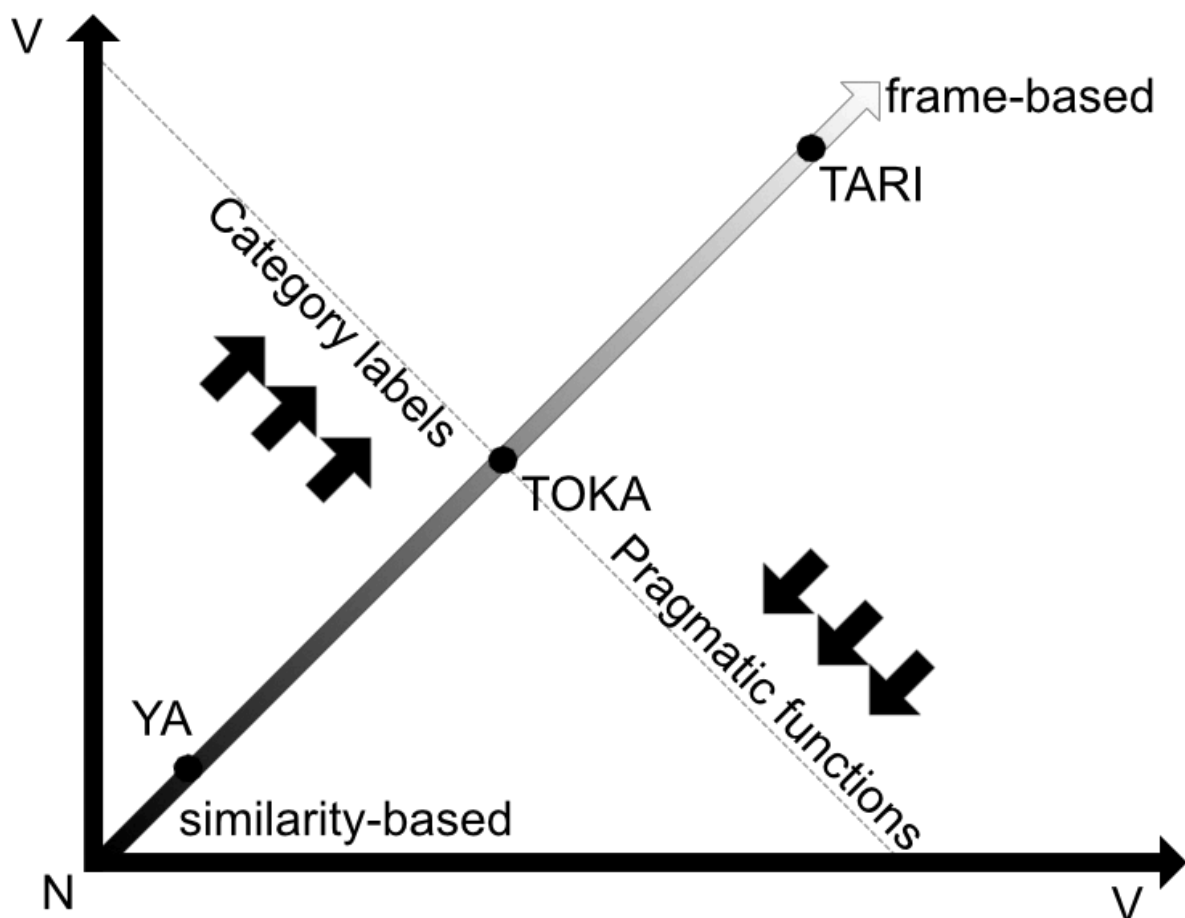
Figure 7.2: Functional space of non-exhaustive connectives

Let us discuss thoroughly each strategy.

Toka is the non-exhaustive connective that occurs more frequently in combination with category labels (46%). In addition, it is also the only connective that can be used as connector to joint examples and labels. Finally, it is also the non-exhaustive connective that can be used more extensively as general extender (63 occurrences) and therefore, that can potentially convey pragmatic functions.

Ya is the non-exhaustive connective that more frequently encodes similarity-based categories (73%). Even though it joints only nouns or noun phrases, the frequency of lexicalization is quite low (37%). Unlike *toka*, it cannot be used as connector, and unlike both *toka* and *tari*, it cannot be used as general extender (that is, with a single example).

Tari is the connective that more frequently encodes frame-based categories (44%). Moreover, it is the connective that occurs less frequently in combination with category labels (25%). Like *ya*, *tari* cannot be used as connector to joint labels and examples. However, unlike *ya*, *tari* can be used as general extender, even though it is less widespread (18 occurrences) than the usage of *toka* as general extender (63 occurrences).

The tendency regarding the frequency of similarity-based and frame-based categories is easily explained by looking at the syntactic nature of the examples (cf. chapter 3).

The tendency is formulated as follows:

(7.1) (a) Non-exhaustive connectives that are typically (or solely) used with nouns or noun phrases are more likely to encode similarity-based categories.

(b) Non-exhaustive connectives that are typically (or solely) used with verbs or verbal phrases are more likely to encode frame-based categories.

Nevertheless, the syntactic nature of the examples is not enough to explain the other tendencies. We argue that tendencies regarding the frequency of category labels and the frequency of pragmatic functions (which passes through the extension to general extender) are due to structural motivations.

Connectives can be used following two structural patterns: 1) the connective is placed between the jointed items as shown in (7.2a), 2) the connective follows each item as shown in (7.2b).

(7.2) (a) X TOKA Y
       (b) X TOKA Y TOKA

Pattern (7.2a) is mandatory for *ya* in Contemporary Japanese (i.e., X YA Y), while it is optional for *toka*. Pattern (7.2b) is mandatory for *tari* (X TARI Y TARI), while it is optional for *toka*.

It is noteworthy that, in the latter case, the non-exhaustive connective is also placed after the last item of the list, which is 1) the closest position to the category labels in Japanese whenever the label is directly linked to the examples (see chapter 3), and 2) the typical position of general extenders. We argue that this second pattern is what enables *toka* to be used as connector and what enables both *toka* and *tari* to be used also as general extenders.

Let us consider the first case, that is, the usage as connector to joint examples and labels. These considerations will be language-specific, given the peculiarities of Japanese words order. First, we can exclude the usage of *ya* as connector, since *ya* cannot occur after the last item of the list (i.e., the pattern in (7.2a) is mandatory).

As for *tari*, at first glance, we may assume that the usage as a connector should be available, since *tari* must occur after the last verb of the list. However, being *tari* a converb (cf. section 1.3.3.3), the predicates suffixed by it are in a non-finite form, and consequently, the last *tari* should be followed by a conclusive *suru* "to do", as shown in (7.3).

(7.3)   X TARI Y TARI SURU [CONNECTOR] [CATEGORY LABEL]

Therefore, it is likely that the final *suru* acts as "obstacle" preventing *tari* to be directly linked to the category label, and thus blocking its usage as connector[57].

On the contrary, toka is the only connective that can be directly connected to the category labels without any "obstacle", as shown in (7.4).

(7.4)   X TOKA Y TOKA [CATEGORY LABEL]

Umi	*toka*	*yama*	*toka*	*iroirona*	*tokoro-ni*	*ikitai.*
sea	TOKA	mountain	TOKA	various	place-LOC	go:DES

'I want to go to different places, such as the sea and the mountains.'

We have seen in chapter 3 that labels tend to occur more frequently 1) when they are directly connected to the examples and 2) when the examples are expressed by nouns or noun phrases. Considering this, it is likely that the usage of *toka* as a connector alongside the high frequency of its occurrences with examples expressed by nouns, explain also the

---

[57] Even though in some cases, final *suru* is omitted, the overall rule seems to persists quite firmly.

high frequency of lexicalization. As we will see in the next section, a similar pattern is found also with the general extender *nado*, which is the most lexicalized strategies in our investigation.

To sum up, we may assume that:

(7.5) (a)   Non-exhaustive connectives that are typically (or solely) used with nouns or noun phrases are more likely to be used in combination with category labels.

(b)   Non-exhaustive connectives that can be used as connector to joint examples and labels are more likely to be used in combination with category labels.

Therefore, if a connective complies with both conditions (e.g., *toka*), it will likely be used in combination with category labels very frequently. If a connective complies with only one condition (e.g., ya), the lexicalization will be still well attested, but less widespread. If a connective does not comply with either the conditions above (e.g., *tari*), the lexicalization will be very sporadic.

Structural patterns also explain the usage as general extender and the further functional extension toward pragmatic functions:

(7.6) (a)   Non-exhaustive connectives that can be placed after the last element of the list, are more likely to be used also as general extender (and thus to convey pragmatic functions).

(b)   Non-exhaustive connectives that cannot be placed after the last element of the list, are hardly used as general extender.

Regarding *tari*, in this case, the presence of *suru* after the last *tari* does not act as an "obstacle" for its usage as a general extender. Nevertheless, the fact remains that *tari* as general extender is less widespread (at least in our corpus) than *toka* as general extender. We do not have a clear answer to this fact, but we may still formulate some hypothesis.

Tari developed first from a perfective auxiliary (see chapter 6), and then into a connective. Moreover, unlike *toka* which is part of the colloquial speech and does not exhibit structural constrains, *tari* is also widespread in formal written language and follows precise structural rules. In this regard, the overall impression is that while the usage of *tari* as a connective is considered part of the norm prescribed by the standard language, the usage as general

extender is perceived more as a violation of the norm. In fact, for instance, according to some linguists during the '50s and the '60s, the usage of a single *tari* was considered 'awkward' (Nagano 1958: 280, Iwabuchi 1961: 180), despite its usage has been attested since the later Edo Era (i.e., Nineteenth Century, cf. Kaneda 1962). This probably means that some speakers may still feel reluctant regarding its usage as general extender.

This reluctance does not occur with *toka*, which is used by speakers with a high degree of freedom, without any specific structural or functional constrain.

To conclude this section, we would like to point out that this functional extension is not peculiar of Japanese, but that similar patterns are attested also in other languages. Specifically, it is noteworthy the functional extension of the Italian connective *piuttosto che* "rather than" (cf. Bazzanella and Cristofoli 1998, Brucale 2012, Mauri and Giacalone 2015). This originally preferential construction is nowadays attested both with its source value and a disjunctive 'or' meaning, albeit the latter is still restricted to the colloquial variety. More specifically, with the latter meaning, it can only be used when the speaker's aim is to name some potential exemplars of a non-exhaustive list, as in (7.7).

(7.7)	*c'è*		*il*	*vantaggio*	*che*	*ti*	*puoi*		*customizzare*	*le*
	there.is		DEF	advantage	that	CLIT	can. 2SG		customize	DEF
	macchina	*come*	*vuoi,*		*in*	*relazione*	*alle*	*tue*		*esigenze*
	machine	as	want.2SG		in	relation	to.DEF	your:PL		need:PL
	(grafica,	*piuttosto che*			sviluppo,		*piuttosto che*	*giochi).*		
	graphics	PIUTTOSTO CHE			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).' (Mauri 2014)

When it is used as disjunctive non-exhaustive connective, *piuttosto che* can also be attached to the last element of the list, following a structural pattern like that of *toka*.

(7.8)	*Abbiamo*	*il*	*galletto*	*con*	*le*	*patate*	*piuttosto che*	*la*
	have:PRS.1PL	DEF	cockerel	with	DEF:PL	potatoes	PIUTTOSTO CHE	DEF:PL
	grigliata	*di*	*salsiccia*	*e*	*maiale piuttosto che*		*le*	*polpette*
	grilled	of	sausage	and	pork	PIUTTOSTO CHE	DEF:PL	meatball
	piuttosto che		*i*	*wurstel*	*piuttosto che…*			
	PIUTTOSTO CHE		DEF:PL	wurstel	PIUTTOSTO CHE			

'We have the cockerel with potatoes, grilled sausages and pork, meatballs, wurstel and so on…' (Brucale 2010)

Interestingly, *piuttosto che* can be also used as a general extender added after a list of items or after a single element as shown in (7.9). As a general extender with a single item, it can also perform pragmatic functions (e.g., pragmatic hedging).

(7.9) *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* *perché*   *non* *le*   *ho*
        ex   kalo   piuttosto che   but   maybe   because   NEG   CLIT   have.1SG
        mai   *provate.*
        never   tried

[talking about diets] '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' (Mauri 2014)

This brief cross-linguistic excursus reveals that the patterns and tendencies here attested for Japanese are not meant to be language-specific, but on the contrary, they are potentially valid for any language that exhibits similar constructions.

7.2.2 THE FUNCTIONAL SPACE OF GENERAL EXTENDERS

Figure 7.3 represents the functional space of general extenders.
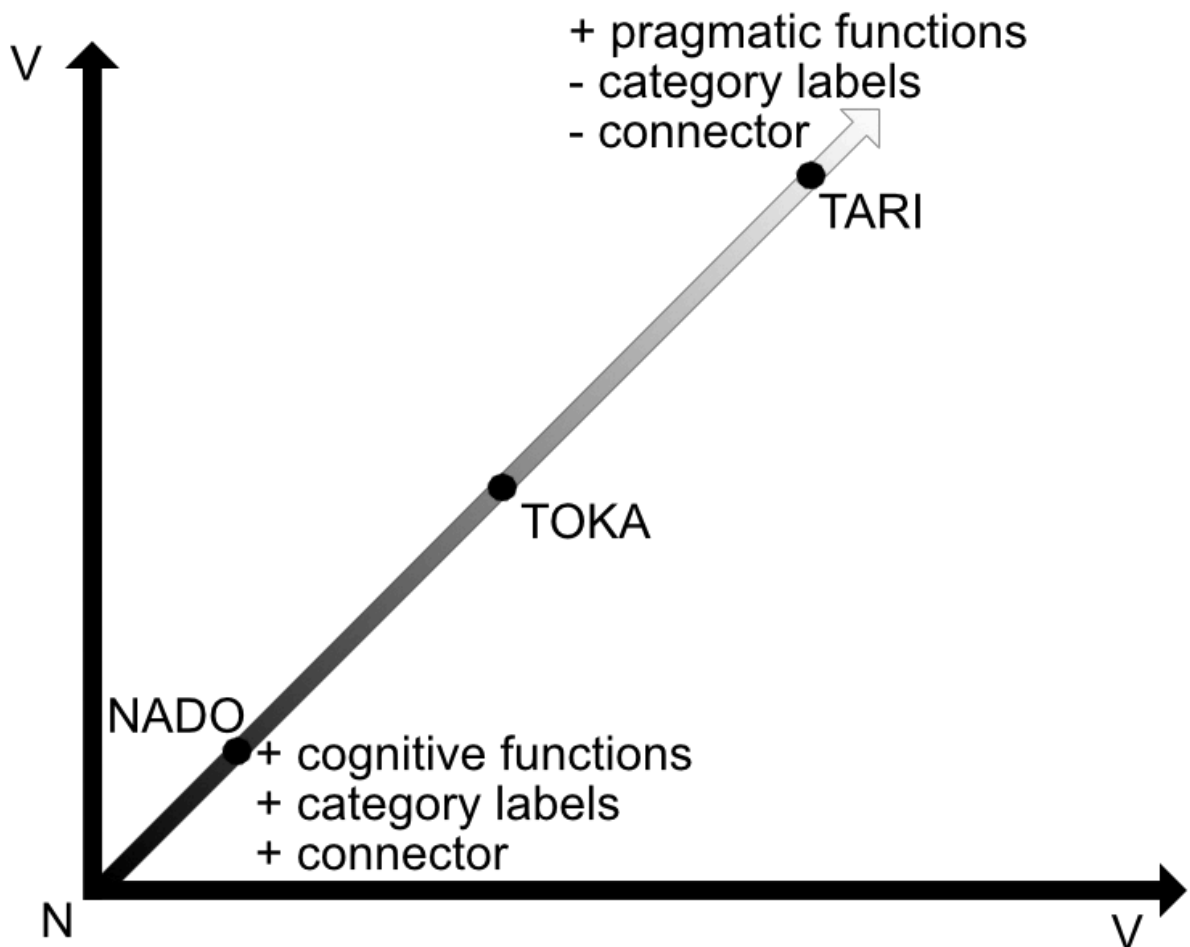
Figure 7.3: Functional space of general extenders

Let us discuss thoroughly each strategy.

Nado very frequently occurs in combination with category labels (52%). Moreover, *nado* is the general extender that most frequently is used as linguistic connector to joint labels and examples. Finally, it is also the most "cognitive" general extender, rarely used to convey pragmatic functions (e.g., hedging): out of 200 occurrences, only 3 perform pragmatic functions.

Toka, as a general extender, represents a middle ground for all these parameters. Out of 63 occurrences of *toka* as a general extender, 38 (60%) are instances of categorization, 25 (40%) are instances of pragmatic functions, such as pragmatic hedging or semantic approximation. Regarding the lexicalization process, 40% of occurrences are combined with

a category label[58]. Finally, *toka* can be used as a connector between labels and examples, but the frequency is lower than that of *nado*[59].

Tari is not frequently used as general extender in our corpus (18 occurrences). Nevertheless, when it actually occurs, typically it performs pragmatic functions (61%). It follows that the overall frequency of lexicalization is very low (just 1 occurrence where the category label is expressed by a simple noun). Moreover, *tari* cannot be used as linguistic connector. As previously noted regarding *tari* as a connective, this can be due to two reasons: 1) the role of the final verb *suru* "to do" which acts as an "obstacle" between tari and the category label, 2) the fact that generally, examples expressed by means of verbs tend to be non-lexicalized (cf. chapter 3).

We argue that these tendencies are due to the syntactic properties of the scope of the general extender.

As for the frequency of category labels, in chapter 3, we have seen that examples expressed by means of nouns or noun phrases tend to be lexicalized more frequently because the process requires a lower cognitive effort. Therefore, we can schematize these tendencies as follows:

(7.10) (a)   General extenders that are typically used with nouns or noun phrases are more likely to be used in combination with category labels. Moreover, considering their position at the end of the list, it is also more likely that this type of general extender will end up covering also the function of connector.

(b)   General extenders that are typically (or solely) used with verbs or verbal phrases, are hardly used in combination with category labels.

As for the types of functions (i.e., cognitive vs. pragmatic), we argue that when the scope of the general extender is extended to the whole utterance (e.g., at the end of the sentence), it is more likely that the function of the general extender is not to exemplify, but to mitigate the assertiveness of the entire utterance. Furthermore, we have seen that examples tend to be syntactically easier to facilitate their elaboration (e.g., nouns are preferred over verbs). In some cases, speakers even express actions by means of verbal nouns instead of verbs.

---

[58] This percentage is not far from the overall frequency of *toka* used in combination with category labels (therefore, as a connective and as a general extender), that is, 44%.

[59] Indeed, in some cases, *toka* is used as a non-exhaustive connective to join the examples and *nado* occurs at the end of the same list as connector.

Therefore, it is very unusual that entire utterances are used as examples. This fact is confirmed by our corpus (0%).

Considering the above, we can identify a continuum regarding the scope of the general extenders that follows the degree of syntactic complexity[60], as schematized by Figure 7.4.

Figure 7.4: Continuum regarding the scope of general extenders

In other words, nouns are more likely to be elaborated as examples, utterances are more likely to be elaborated as requiring pragmatic hedging. Consequently:

(7.11) (a)   General extenders that are typically used with nouns or noun phrases are more likely to perform cognitive functions (i.e., exemplification to categorize).

     (b)   General extenders that are typically attached to verbs and therefore tend to occur at the end of the utterance, are more likely to perform pragmatic functions (i.e., pragmatic hedging).

The second tendency explains the case of *tari*. In fact, when it is used as general extender, *tari* often behaves as a final-utterance marker, extending its scope to the entire utterance. This justifies the high frequency of pragmatic functions whenever *tari* is used as general extender.

---

[60] We place noun as the left edge of the continuum, that is, as the less complex syntactic structure (while at the opposite side we place the entire utterance). We decide to use noun instead of noun phrase because noun phrases can profile further connections (i.e., complements and other linguistic adjuncts), while the noun profiles only the bare element, without any further connection (cf. Langacker 1987a, 1991b, Givón 2001).

7.3 A DIACHRONIC GLANCE ON JAPANESE EXEMPLIFYING CONSTRUCTIONS

Synchronic investigation can reveal instances of constructions that have more than one use and consequent tendencies and correlations, but it cannot formulate strong hypotheses about the relationships and implications among these functions.

Recent investigations on Semantic Maps and thus on methods of representing polyfunctionality (see van der Auwera 2013, Narrog 2009, Narrog and van der Auwera 2011), have highlighted the important role of the diachronic dimension, both to explain and predict processes of functional extension.

In this regard, van der Auwera (2013: 165) notes that a possible reason why a marker can have more than one function is that one function may have developed out of the other one. In this regard, diachronic maps can offer more information but also makes the mapping more restrictive and predictive.

Our case is even more complex, since we are investigating four markers which are very dissimilar to each other at different levels (cf. section 1.3.2), but that exhibit strongly homogeneous functional spaces. It is therefore natural to wonder if the diachronic dimension can help us to understand this functional homogeneity. Our investigation has revealed that all these markers share a connection with the irrealis value.

In section 1.3.3.4, we have seen that *toka* is originated by the combination of the marker *to* (which can be interpreted both as the comitative and conjunctive marker *to* or as the quotative marker *to*) and the interrogative/indefinite marker *ka*. Interestingly, in the grammar written during the '70s by Martin (1975) *toka* is described as "*to ka*", with a space between the two markers, indicating that the process of combining is still ongoing. On the contrary, in recent studies (e.g., Suzuki 1998, Ohori 2004, Taylor 2010, 2015), *toka* is described as a single, all one word marker.

Figure 7.5: Diachronic development of *toka*

In section 6.1, we have seen that the non-exhaustive connective *tari* originated by the auxiliary *-tari* which was a resultative perfect, denoting completion of an event and the state

of result of this event in Old Japanese. However, over time, *tari* has lost any aspectual value, developing towards that lack of referentiality noted by Suzuki (1998) and Narrog (2012). In this regard, Narrog (2012) notes that "Modern Japanese -*Tari* is not specifically an irrealis subordinate mood, but it does lead to a lower factuality of the event portrayed, since through -*Tari* the event becomes marked as unspecific" (2012: 147). According to him, this lower factuality is what enables the actual functions of *tari* to emerge (e.g., exemplification, iterative and distributive meanings, hedging). The emergence of the usage of *tari* as a connective is generally attributed to Middle Japanese (Narrog 2012).
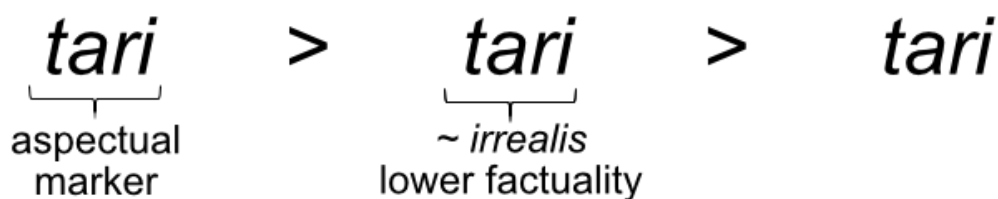
Figure 7.6: Diachronic development of *tari*

The non-exhaustive connective *ya* can be traced back to the Altaic **j-* interrogative root[61]. This root is well attested in all branches of Altaic to form different interrogative markers and constructions (cf. Greenberg 1987). For example, in Sakhalin Ainu –*ya* means "or", and the Hokkaido dialect Batchelor (north of Japan) gives *ya X ya Y* as "whether X or Y". Kanazawa (1910: 38) connects the Ainu marker *ya* to the Japanese homonym. As noted in section 6.1.1, in Old Japanese, *ya* and *ka*[62] were interrogative markers in competition. Over time, they have changed repeatedly their functional spaces to avoid overlapping with each other. For example, in Old Japanese, *ya* was used in rhetorical questions or exclamations, but hardly for pure interrogative uses (e.g., *ari ya nasi ya* "is there or is there not?"). According to Frellesvig (2010: 359), through Early Middle Japanese *ya* largely replaced *ka* inside yes/no questions, then it has gradually been reinterpreted as expression of uncertainty 'I wonder', often used in combination with various modal forms[63].

---

[61] Regarding the classification of Japanese and its relation with Altaic, see Shibatani 1990.

[62] Interestingly, also interrogative marker *ka* is originated by an Altaic interrogative root, namely **k-* (see Greenberg 1987).

[63] From the combination of *ya* and one modal form of *ar-* originated the marker *yara*, which is another connective to indicate non-exhaustivity (see Frellesvig 2010: 335). This means that also *yara* is linked to the irrealis mood, further confirming our discussion.

The use of *ya* as sentence interrogative in non-rhetorical sentences has declined since Late Middle Japanese in favor of sentence final *ka* (see Lewin 1959, Frellesvig 2010). By the end of Late Middle Japanese, *ya* has developed uses as a connective 'and, or, or the like' (Frellesvig 2010: 359).

Some traces of its past usage as interrogative remains in Modern Japanese in fixed expressions such as *are ya kore (ya),* equivalent of *are ka kore (ka)* "that and/or this" (see Martin 1975: 157).
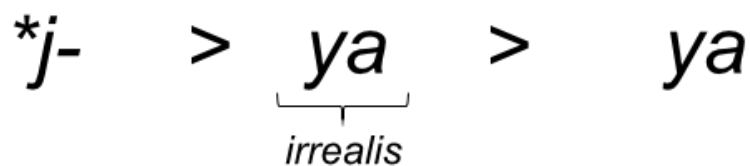
Figure 7.7: Diachronic development of *ya*

The morpheme *nado* can be traced back to ancient texts such as the Manyoshu (cf. Yamaguchi 1988). Yamaguchi (1988) explains the generally accepted theory of the development of *nado*, as in *nanito > nando > nado*. He also claims that the meaning of *nani* (i.e., indefinite interrogative, 'what') in the original word *nanito* remains in *nado* as it is used in Modern Japanese to create indeterminacy (i.e., the lack of referentiality) regarding the attached item(s). In this regard, Frellesvig (2010) emphasises the role of *to* in *nanito*. He notes that *to* and its variant *tote* functioned as purposive conjunctional marker, undergoing a grammaticalization process from its primary copula function. In this regard, *nani-to* (and its variant *nani-tote*) can be paraphrased as 'being what?' (2010: 245).

Moreover, Vovin (2003) attests the use of *nado* in Classical Japanese as a bound morpheme, having the function of a representative plural (Vovin 2003), opposed to the additive plural form –*domo* (e.g., *hito-domo*, "persons"). For instance:

(7.12) *Tani   no   soko-nado   ni   wa*
       Valley gen   bottom-repr   loc   top
       'At the bottoms of valleys and other places like that' (Vovin 2003: 40)

Therefore, we may argue that *nado* has shifted from morphology to lexicon, but also from grammar to pragmatics, along a path that has been called *degrammaticalization* (Ramat 1992). *Nado* has shifted from being a bound morpheme that was part of the number

paradigm (in Classic Japanese), to become an independent morpheme characterized by syntactic flexibility and pragmatic functions (in Modern Japanese).

$$nani + to \; > \; nando \; > \; nado$$
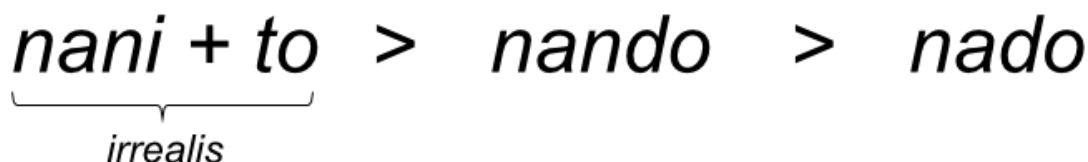
irrealis

Figure 7.8: Diachronic development of *nado*

To sum up, this brief outline shows that, despite the different diachronic developments, at some point of their existence, all these markers have been related to the notion of irrealis. In this regard, we argue that it is indeed this intrinsic irrealis value they still carry that allows them to perform all the functions now available in Contemporary Japanese. More specifically, we propose that irrealis markers can develop towards the encoding of non-referentiality and/or the lower factuality of the events described.

This development allows exemplifying markers to cover different functions depending on the type of structural patterns they have developed into. Therefore, those markers that have developed uses as connectives (e.g., *ya, tari, toka*), have ended up covering 1) exemplification and 2) conjunctive alternative readings. Those markers that have developed uses as general extenders (e.g., *nado, tari, toka*) have ended up covering 1) exemplification, 2) hedging, 3) intensification of the negative polarity.

Irrealis > lower factuality / lack of referentiality

connective
- exemplification
- conjunctive alternative

general extender
- exemplification
- hedging
- intensification (negation)
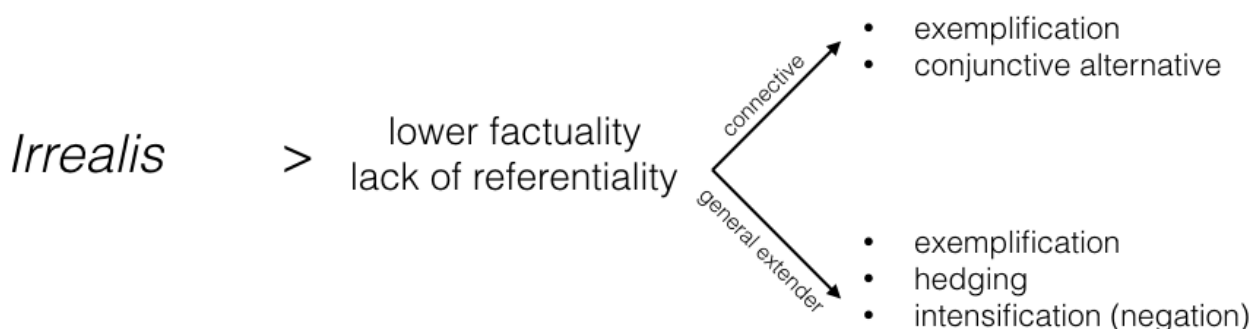
Figure 7.9: Diachronic development of Japanese exemplifying constructions

The obvious question regards the connection between this non-referentiality/lower factuality and all the attested functions. In chapter 5, we have already seen the connection

with hedging functions and, in chapter 6, the connection with the intensification of the negative polarity of the utterance. Moreover, in chapter 6, we have also highlighted the possible incompatibility between conjunctive alternative readings and connectives that convey a higher factuality such as *to* and the *te*-form.

Yet, what about exemplification? Actually, the relationship between these two semantic core (i.e., lack of referentiality and lower factuality) and the role of the example is rather straightforward. In fact, in order to be examples, in the sense of being representative of a larger group, items must lack their own referentiality. In other words, whenever they are examples, they are never referent of themselves (i.e., they do not bear any independent reference), but only cognitive arrows towards the larger group, that is, the category. At the linguistic level, this lack of referentiality is then translated in non-exhaustivity, as showed in section 4.2.

The diachronic pattern from irrealis to exemplifying markers described in Figure 7.9, is not specific of Japanese or limited to connectives and general extender. For example, we have already examined the case of the disjunctive (and therefore linked to the notion of irrealis, cf. Mauri 2008) non-exhaustive connective *piuttosto che* "rather than" in Italian, and its functional extension to general extender (cf. section 7.2.1).

Similar functional extensions have been attested also in Italian, regarding the epistemic markers *non so* "I don't know" (and analogues, *che so* "what do I know", *che ne so* "what do I know about it") as shown in (7.13), the epistemic marker *magari* "maybe" (cf. Mazontti 1998, Masini and Pietrandrea 2010, Ghezzi 2013) as shown in (7.14), parenthetical expressions *metti*, *mettiamo* and *poniamo* "let us suppose" (cf. Schneider 2007, Mihatsch 2010b, Ghezzi 2013) often used to introduce hypothetic facts as shown in (7.15).

(7.13) *Prendiamo*   *non so*   *il*   *cinese.*
        Take:SBJV.1PL   NON SO   DEF   Chinese
        'Let us consider I don't know Chinese.'

(7.14) *Tu*   *hai*   *un*   *UNICO*   *filmato*   *di,*   *poniamo,*
        you.SG   have:PRS.2SG   INDEF   unique   clip   of   PONIAMO
        mezz'ora [...]
        half hour
        'Imagine to have a UNIQUE clip extending, let us suppose, half an hour [...]' (Nunc Corpus)

(7.15)   *Assistiamo*   *sempre*   *alle*   *stesse immagini*   *cambiano*   *soltanto*

        assist:PRS.1PL   always   to:DEF same images   change:PRS.3PL   only

        magari i   *nomi*   *delle*   *battaglie*   *o*   *il*   *numero*   *dei*   *feriti*

        MAGARI DEF   names of:DEF   fight:PL   or   DEF   number   of:DEF injured

        eccetera […]

        etcetera

        'We always see the same images; the only different things are maybe the names of fights or the numbers of injured people [...]' (Lip Corpus)

All these markers can perform the exemplifying function and some hedging functions (cf. Ghezzi 2013, Manzotti 1998).

Given the distance between Italian and Japanese, but also among the strategies here discussed, it appears that the functional extension irrealis > exemplification is likely universal, and probably also cognitively motivated.

In fact, beyond the semantic core of lacking referentiality and factuality, it is also noteworthy the parallelism between the relationship among examples and that among irrealis elements. Indeed, as already noted in chapter 5, one of the components of the notion of exemplification is the interchangeability of particular instances (see Manzotti 1998): examples are fundamentally arbitrary choices, since the selected item is just one of a number of possible other examples. This fundamentally means that there are always other perfectly available elements that can be added to the list or that can replace the mentioned one without any loss. In this sense, they do share a significant amount of similarities with irrealis elements.

8. CONCLUSION AND PROSPECTS

As indicated in the introduction to this study, the aim of this research was to investigate the cognitive and linguistic mechanisms underlying the coding of exemplification, which has been defined in functional terms, as a process in which an instance is profiled and construed as representative of an abstract formulation. The status of examples is thus signalled by suggesting the presence of further elements. The functional definition of exemplification has made it possible to identify a set of Japanese exemplifying constructions, without relying on structural parameters (i.e., analytic constructions such as *for example*). The investigation on the attested patterns of usage reveals several tendencies, showing the mechanisms and principles at work in the expression of exemplification.

The attested patterns demonstrate that category labels and examples are not merely competing strategies to communicate conceptual categories, but instead they frequently occur together, in the same construction. Both provide a specific semantic contribution to the inferential process and closely collaborate by covering any lacks in the reference provided by the other element. Labels allow to make direct reference to the defining property of the category but, being still anchored to the abstract dimension, they require a greater effort to be interpreted in a specific context. On the other hand, examples allow the contextualization and actualization of the category, but they do not clearly specify the shared property, which needs to be identified by means of a comparison, not only among the examples, but also between examples and the broader context, requiring an extra cognitive effort.

This functional cooperation is mirrored by their linguistic coding. Indeed, the syntactic and semantic properties of labels strongly correlate with how examples are selected and how they are encoded. Specifically, the process of lexicalization can be schematized by means of a continuum. The more the examples resemble to the prototypical nouns, the easier is for the speaker to create specific labels to designate the target category. On the contrary, the more the examples resemble to the prototypical verbs, the more lexicalization relies on syntactically complex labels or, alternatively, unspecific syntactically simple nouns. More generally, labels and examples are specifically chosen and encoded to ease the cognitive effort of the speaker and of the hearer to ultimately elaborate the abstract formulation, namely the category.

Beyond examples, the analysis has also revealed the important role played by two other elements in the process of constructing and communicating conceptual categories: non-

exhaustivity and context. Non-exhaustivity (encoded by means of non-exhaustive tags) is a crucial linguistic tool to trigger inferential processes. Non-exhaustive tags work as indexical items to activate the presupposition of an unspecified open variable to be saturated through access to the context and ultimately to make the target category coincides with the "what-is-said" part of the utterance meaning. As for the context, the analysis has indicated that it should not be considered merely as some inert background to the cognitive processes, but as an active participant. More specifically, the co-text can provide semantic clues towards identification of the defining property of the category, directing the inference. Therefore, categorization is actively driven by the context, signifying a deeply dynamic process.

Exemplifying constructions can perform other functions beyond exemplification proper. They can be used to stress the fact that the mentioned item should be conceived merely as an option chosen from a larger set of possibilities, giving rise to hedging functions, such as semantic approximation or pragmatic hedging. At the communicative level, they can be used to achieve discourse effects such as vagueness and politeness, with the ultimate purpose of having a successful and smooth conversation.

Starting from the same lack of referentiality or lower factuality, exemplifying constructions can also perform an intensification of the negative polarity of the utterance. In this case, the underlying set configured by exemplification is used to widen the scope of negation, to the point of refuting not only the mentioned item, but also any other similar element. This function can be further used to achieve discourse effects, such as emphasizing the impossibility of the situation or suggesting the speaker's contempt toward the item itself.

The lower factuality is also at the basis of the usage of some exemplifying connectives to convey conjunctive alternative readings (i.e., separative conjunction reading and free alternative reading). More precisely, because of their functional core, these markers allow to profile events that are 'both real but not jointly real'. The further collocation of these events in the temporal axis may give rise to aspectual interpretation, such as iterative and distributive values.

The three major results of this study can thus be described as follows. First of all, tendencies on the linguistic coding of examples show interesting correlation with the way language users create and process conceptual categories. While categories may be driven from driven from exemplars of any kind, speakers do not choose them randomly but to better represent concrete experiences (i.e., the preference for concrete entities) and to be easily processed by the hearer (i.e., the preference for examples encoded by noun phrases, since they profile less complex set of interconnections). These tendencies are evident regardless

of the presence or absence of a category label. Moreover, they are likely cognitively motivated by the basic function of exemplification, that is to elaborate complex (and often abstract) information starting from more concrete and straightforward material. This does not affect only the type of examples selected, but, more interestingly, also the way they are encoded and thus processed.

Secondly, the analysis has revealed a strong correlation between the structural level and the functional level. In this sense, the functional domain strongly depends on the type of construction the markers belong. In addition, the functional extension depends on a structural extension. For instance, non-exhaustive connectives that can be used also as general extender (e.g., *tari* and *toka*) perform also the functions associated with general extenders (e.g., hedging, intensification of the negative polarity).

Thirdly, all the functions attested are related to each other by means of a functional core denoting a lack of referentiality or lower factuality of the events profiled. In the case of Japanese exemplifying constructions, this correlation is motivated at the diachronic level, since all the markers investigated shows similar developments with regards to the irrealis value. In other words, all these markers have been related to the notion of irrealis in some way. The functional extension irrealis > exemplification is likely universal, and probably also cognitively motivated. Examples are fundamentally profiled as arbitrary interchangeable choices, since the selected item is just one of several possible other examples. This ontological component of exemplification paves the way for a connection with the notion of irreality.

There are at least two directions along which this research could be continued. First, the attested tendencies may be verified on other languages, even applying a cross-linguistic approach. More specifically, it would be interesting to test some of them on languages which exhibit a different word orders than Japanese. If those tendencies that do not strongly rely on the word order (e.g., the frequency of lexicalization and all the relative relations between labels and examples, the preference for examples expressed by noun phrases, the development of non-exhaustive connectives into general extenders, the patterns and frequency of pragmaticalization) were confirmed by the investigation on other languages, the research on exemplification would be improved especially with regards its validity at the cognitive level. With greater and more varied evidences, the attested tendencies on the linguistic coding of exemplification might be better supported, thus making the resulting generalizations more powerful. This would ultimately confirm the role of exemplification as a universal cognitive mechanism, with evident and investigable impacts on the modalities of

linguistic coding.

The second direction in which this study may be continued is diachronic. As pointed out at various points of this study, exemplifying markers seem to bear underlying links with irreality dimension. Some of the makers in the present work have been irrealis makers at some point of their existences (i.e., *ya, nado, toka*), while others have showed a strong relation with irrealis mood (i.e., *tari*). This does not seem to be a language-specific diachronic patter, since similar situations have been attested also in other languages (e.g., the usage of parenthetical expressions and epistemic markers in Italian to introduce exemplification, cf. Manzotti 1998, Ghezzi 2013). Moreover, it also appears that the irreality dimensions is the driving force underlying some functional extension patterns. A diachronic exam of the evolution of exemplifying markers (with a specific focus on the development from hypothetical or interrogative constructions) would be crucial also to confirm functional patterns identified synchronically.

REFENRECES

Primary Sources

Japanese section (jpn_news_2005-2008) of the Leipzig Corpora Collection (LCC)

= Written corpus of Japanese (Internet corpus)

http://corpora.uni-leipzig.de/en

Secondary Sources

Japanese Web 2011 (jpTenTen11), Sketch engine = Written corpus of Japanese

https://www.sketchengine.co.uk/

English Web 2013 (enTenTen13), Sketch engine = Written corpus of English

https://www.sketchengine.co.uk/

Ahmad, Khurshid and Heather Fulford. 1992. Knowledge Processing: 4. Semantic Relations and their Use in Elaborating Terminology. In *Computing Sciences Report* CS-92-07. Guildford: University of Surrey.

Alpatov, Vladimir M. 1997. Expression of the meanings of situational plurality in Modern Japanese. In Xrakovskij, Victor S. (eds.), *Typology of Iterative Constructions*. 388-399.

Ariel, Mira and Caterina Mauri. forthcoming. Why use *or*?. To appear in *Linguistics*.

Aristotle. 1984. Rhetoric. In Jonathan Barnes (eds.), *The Complete Works of Aristotle: The Revised Oxford Translation*, volume LXXI, tome I of Bollingen Series. Princeton, NJ: Princeton University Press.

Austin, John L. 1940. *The meaning of a word*. Reprinted in Philosophical Papers. Oxford: Oxford University Press, 1961.

Austin, John L. 1962. *How to do things with words*. Boston: Harvard University Press.

Baesler, James E. and Judee K. Burgoon. 1994. The temporal effects of story and statistical evidence on belief change. *Communication Research* 21. 582-602.

Bally, Charles. 1970 [1909]. *Traité de stylistique française*. Genève: Librairie de l'Université Georg & Cie S.A.

Banno, Eri. (2000). *An integrated course in elementary Japanese*. Tokyo: The Japan Times.

Bar-Hillel, Maya. 1980. The base-rate fallacy in probability judgments. *Acta Psychologica* 44(3). 211- 233.

Barotto, Alessandra and Caterina Mauri. 2016. *Constructing lists for constructing categories*. Paper presented at the 50th Società Linguistica Italiana (SLI) Conference, Milan, September 22-24, 2016.

Barsalou, Lawrence W. 1982. Context-independent and context-dependent information in concepts. In *Memory & Cognition* 10. 82-93.

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

Barsalou, Lawrence W. 1991. Deriving categories to achieve goals. In G.H. Bower (eds.), *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. Cambridge, MA: MIT Press/Bradford Books. 121-176.]

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

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

Bazzanella, Carla and Mirella Cristofoli. 1998. Piuttosto che e le alternative non preferenziali: un mutamento in atto? *Cuadernos de filologia italiana*, 5. 267-278.

Bernini, Giuliano and Paolo Ramat. 1996. *Negative sentences in the languages of Europe*. Berlin; New York: Mouton de Gruyter.

Biber, Douglas. 1991. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Billig, Michael. 1987. *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge: Cambridge University Press.

Birk, Elisabeth. 2007. The Use of Examples. In *Beyond Mimesis and Nominalism: Representation in Art and Science.* London. 22-23.

Bisang, Walter. 2009. On the evolution of complexity—sometimes less is more in East and mainland Southeast Asia. In G. Sampson, D. Gil and P. Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press. 34-49

Bisang, Walter. 2013. Language contact between geographic and mental space. In Auer, P., M. Hilpert, A. Stukenbrock and B. Szmrecsanyi (eds.), *Linguistic Perspectives on Space: Geography, Interaction, and Cognition*. Berlin: Mouton de Gruyter. 61-100

Blanche-Benveniste, Claire, Mireille Bilger, Christine Rouget and Karel Van den Eyende 1990. *Le français parlé. Etudes grammaticales*. Paris: Editions du Centre National de la Recherche Scientifique.

Bonvino, Elisabetta, Francesca Masini, & Paola Pietrandrea. 2009. List constructions: A semantic network. Paper presented at the 3rd International AFLiCo Conference – Grammars in Construction(s), Paris, May 27-29, 2009.

Borillo, Andrée. 1996. La relation partie-tout et la structure [NI à N2] en francais. *Faits de Langue*, 7. 111-120

Brosius, Hans B. & Anke Bathelt. 1994. The Utility of Exemplars in Persuasive Communications. *Communication Research* 21(1). 48–78.

Brosius, Hans B. 2003. Exemplars in the News: A Theory of the Effects of Political Communication. In Jennings Bryant, David R. Roskos-Ewoldsen & Joanne Canto (eds.), *Communication and emotion: Essays in honor of Dolf Zillmann*. New York: Routledge. 179-197.

Brown, Gillian, and George Yule. 1983a. *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, Penelope & Levinson, Stephen C. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Brown, Penelope and Levinson, Steven C. 1978. Universals in language usage: Politeness phenomena. In Goody, E. (eds.), *Questions and politeness. Strategies in social interactions*. Cambridge: Cambridge University Press. 56-311.

Brucale, Luisa. 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.

Bruner, Jerome S., Jacqueline J. Goodnow, & George A. Austin. 1956. *A study of thinking*. New York: Wiley

Bruner, Jerome S. 2002. *Making Stories: Law, Literature, Life*. New York: Farrar, Straus and Giroux.

Caffi, Claudia. 2001. *La mitigazione. Un approccio pragmatico alla comunicazione nei contesti terapeutici*. Münster: LIT Verlag.

Caffi, Claudia. 2007. *Mitigation*. Amsterdam: Elsevier.

Carston, Robyn. 2002. *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.

Carston, Robyn. 2010. Metaphor: ad hoc concepts, literal meaning and mental images. Proceedings of the Aristotelian Society 110.

Channell, Joanna. 1994. *Vague Language*. Oxford: Oxford University Press.

Chater, Nick and Ulrike Hahn. 1997. Representational Distortion, Similarity, and the Universal Law of Generalization. In Proceedings of the Interdisciplinary Workshop on Similarity and Categorization, SimCat97, University of Edinburgh. Dept. of Artificial Intelligence, University of Edinburgh

Chen, Lian D. 2005. "Nazo" to "nanzo" no imi, kinoo. *Sekai no nihongo kyoiku 15*. Tokyo: Japan Foundation. (The meaning of "nazo" and "nanzo").

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

Comrie, Bernard. 1976. *Aspect.* Cambridge: Cambridge University Press.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Oxford: Blackwell. (2nd edition)

Corbett, Greville G. 2000. *Number*. Cambridge University Press, Cambridge.

Croft, William. 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: University of Chicago Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Cambridge: Cambridge University Press.

Croft, William. 2003. *Typology and universals*. 2nd ed. Cambridge: Cambridge University Press.

Croft, William and Alan D. Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cruse, Alan D. 1986. *Lexical semantics*. Cambridge, England: University Press.

Cruse, Alan. 2000. *Meaning in language. An introduction to semantics and pragmatics*. Oxford: Oxford press.

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

Daniel, Michael and Edith Moravcsik. 2005. Associative plurals. In Matthew S. Dryer, Martin Haspelmath,

De Haan, Ferdinand. 2010. Building a Semantic Map: Top-Down versus Bottom-Up Approaches, *Linguistic Discovery* 8(1). 102-117.

Dines, Elizabeth R. 1980. Variation in discourse – 'and stuff like that'. *Language in Society* 9. 13-31.

Dubois, Sylvie. 1992. Extension particles, etc. *Language Variation and Change*, 4. 179-203.

Eckart Thomas, Uwe Quasthoff. 2013. Statistical corpus and language comparison on comparable corpora. In *BUCC - Building and using comparable corpora*. Berlin: Springer.

Ediger, Anne. 1995. *An Analysis of Set-marking Tags in the English Language*. Dissertation, University of California, Los Angeles.

Edwards, Derek. 1991. Categories are for talking: On the cognitive and discursive bases of categorization. In *Theory and Psychology* 1 (4). 515-542.

Edwards, Derek. 1997. *Discourse and Cognition*. London: Sage.

Elliot, Jennifer R. 2000. Realis and irrealis: forms and concepts of the grammaticalization of reality. *Linguistic Typology*, 4-1. 55-90.

Enfield, Nick. 2007. *A grammar of Lao*. Berlin, New York: Mouton de Gruyter.

Erman, Britt. 2001. Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics,* 33. 1337-1359.

Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2). 222–254.

Fillmore, Charles J. and Paul Kay. 1995. *Construction Grammar*. Manuscript, University of California, Berkeley.

Foley, William A. and Robert Van Valin. 1984. *Functional syntax and universal grammar.* Cambridge University Press.

Foley, William 1997 *Anthropological Linguistics: An Introduction*. Blackwell.

Fraser, Bruce. 1975. Hedged perfomatives. In Cole, P. and Morgan, J. L. (eds.), *Syntax and semantics vol.3. Speech Acts.* New York: Academic Press. 187-210.

Fraser, Bruce. 1980. Conversational mitigation. *Journal of Pragmatics*, 4. 341–350.

Fraser, Bruce. 2010. Pragmatic Competence: The Case of Hedging. In Kaltenböck, G., Mihatsch, W. and Schneider, S. (Eds.), *New Approaches to Hedging*. Bingley: Emerald. 15-34.

Frellesvig, Bjarke. 2010. *A History of the Japanese Language*. Cambridge: Cambridge University Press.

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

Gerdes, Kim and Sylvain Kahane. 2009. *Speaking in piles: Paradigmatic annotation of French spoken corpus.* Processing of the 5th Corpus Linguistics Conference, Liverpool.

Ghezzi, Chiara. 2013. *Vagueness Markers in Contemporary Italian: Intergenerational Variaton and Pragmatic Change*. Ph.D. dissertation.

Gibson, Rhonda. and Dolf Zillmann. 1994. Exaggerated versus Representative Exemplification in News Reports: Perception of Issues and Personal Consequences. *Communication Research* 21(5). 603–24.

Goffman, Erving. 1955. On Face-work: An Analysis of Ritual Elements of Social Interaction. *Psychiatry: Journal for the Study of Interpersonal Processes,* 18. 213-231.

Givón, Talmy. 2001. *Syntax. An introduction (2◦ edition), volume I e II*. Amsterdam; Philadelphia: John Benjamins.

Givón, Talmy. 1983. Topic Continuity in Discourse: A Quantitative Cross-Language Study. Amsterdam: John Benjamins.

Gluck Mark A. and Gordon H. Bower. 1988. From Conditioning to Category Learning: An Adaptive Network Model. *Journal of Experimental Psychology: General*, 117. 227-247.

Glucksberg, Sam. 2001. Understanding Figurative Language. Oxford: Oxford Univerity Press.

Göksel, Aslı and Celia Kerslake. 2005. [Third reprint 2010], *Turkish, A Comprehensive Grammar*. London, Routledge.

Goldhahn, Dirk, Thomas Eckart, and Uwe Quasthoff. 2012. *Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages*. In Proceedings of LREC 2012. 759-765.

Goldstone, Robert L., Alan Kersten and Paulo F. Carvalho. 2013. Concepts and Categorization. In Weiner, Irving B., Alice F. Healy and Robert W. Proctor (eds.), *Handbook of Psychology*, Volume 4. *Experimental Psychology*, 2nd edn. New York, NY: Wiley. 607-630.

Greenberg, Joseph H. 1987. Relative pronouns and P.I.E. word order type in the context of the Eurasiatic hypothesis. in Winfred P. Lehmann (eds.), *Language typology 1987: Systematic balance in language: Papers from the Linguistic Typology Symposium, Berkeley 1–3 December 1987*. Amsterdam: Benjamins. 123-138.

Grice, Paul H.1975. Logic and Conversation. In Cole, P. and Morgan, J. L. (eds.), *Syntax and Semantics. Volume 3: Speech Acts*. New York: Academic Press. 41-58.

Grice, Paul H.1989. *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Halliday, Michael A. K. and Hasan, Ruqaia. 1976. *Cohesion in English*. London: Longman.

Halliday, Michael A. K. 1989. *Spoken and Written Language*. Oxford: Oxford University Press.

Hampton, James A. 2001. The role of similarity in natural categorization. In M. Ramscar, U. Hahn, E. Cambouropolos and H. Pain (eds.), *Similarity and categorization*. Oxford: Oxford University Press. 13–28.

Haspelmath, Martin and Oda Buchholz. 1998. Equative and similative constructions in the languages of Europe. In van der Auwera Johan and Dónall Ó Baoill (eds.), *Adverbial constructions in the languages of Europe.* Berlin: Mouton de Gruyter. 277–334.

Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In M. Tomasello (eds.), *The new psychology of language, vol. 2*. New York: Erlbaum. 211-243.

Haspelmath, Martin. 2004. Coordinating constructions: an overview. In Martin Haspelmath (eds.), *Coordinating Constructions*. Amsterdam: John Benjamins. 3-39.

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

Hearst, Marti A. 1992. Automatic Acquisition of Hyponyms from Large Text Corpora. In Proceedings of the Fourteenth International Conference on Computational Linguistics (COLING-92).

Hendriks Vettehen, Paul and Leo van Snippenburg. 2004. Do well-balanced exemplars in news stories provide food for thought? In Renckstorf, Karsten, Denis McQuail, Judith E. Rosenbaum, Gabi J. Schaap (eds.), *Action theory and communication research: Recent developments in Europe*. Berlin: De Gruyter. 279-290.

Heritage, John. 1984. *Garfinkel and ethnomethodology*. Cambridge: Polity.

Hobbs, Jerry R. 1985. *On the Coherence and Structure of Discourse* (Report 85–37). Stanford: Stanford University, Center for the Study of Language and Information.

Horie, Kaoru. 2000. Complementation in Japanese and Korean: A contrastive and cognitive linguistic approach. In Horie K. (eds.), *Complementation: Cognitive and Functional Perspectives*. Amsterdam: John Benjamins. 11-31.

Horie, Kaoru. 2003. What cognitive linguistics can reveal about complementation in non-IE languages: Case studies from Japanese and Korean. In Eugene H. Casad and Gary B. Palmer (eds.), *Cognitive Linguistics and Non-Indo European Languages.* Berlin, New York: Mouton de Gruyter. 363-388.

Horn, Laurence R. 1972. *On the semantic properties of the logical operators in English*. Los Angeles: University of California dissertation.

Hübler, Axel. 1983. *Understatements and Hedges in English*. Amsterdam: Benjamins.

Hyland, Ken. 2007. Applying a gloss: Exemplifying and reformulating in academia discourse. *Applied Linguistics* 28 (2). 266-285.

Ichikawa, Yasuko. 1991. Heiritsujoshi "to" to "ya" ni kansuru ichikōsatsu. *Bungei gengo Ken Kiwamu gengo-hen*, 20. 61-79

Iguchi, Yuko. (1998). Functional variety in the Japanese conjunctive particle kara 'because'. In Ohori, Toshio (eds.), *Studies in Japanese Grammaticalization - Cognitive and Discourse Perspectives*. Tokyo: Kuroshio. 99-134.

Iwabuchi, Elsutaro. 1961. *Akubun.* Tokyo: Nippon Hyooron sha.

Iwasaki, Shoichi. 1993. *Subjectivity in Grammar and Discourse: Theoretical Considerations and a Case Study of Japanese Spoken Discourse*. Amsterdam: John Benjamins.

Iwasaki, Shoichi. 2013. Japanese: Revised Edition. Amsterdam: John Benjamins Publishing Company.

Jennings, R E. 1994. *The Genealogy of disjunction*. Oxford: Oxford University Press.

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

Jucker, Andreas, Smith, Sara W. and Lüdge, Tanja. 2003. Interactive aspects of vagueness in conversation. *Journal of Pragmatics*, 35. 1737-1769.

Kaiser, Stefan, Yasuko Ichikawa, Noriko Kobayashi and Hilofumi Yamamoto. 2001. *Japanese: A Comprehensive Grammar*. New York: Routledge Grammars.

Kaltenböck, Gunther, Mihatsch, Wiltrud and Schneider, Stefan (eds.). 2010. *New Approaches to Hedging*. Bingley: Emerald.

Kamp, Hans. 1973. Free choice permission. *Proceedings of the Aristotelian Society,* 74. 57−74.

Kanazawa, Shosaburo. 1910. *The common origin of Japanese and Korean*. Tokyo

Kaneda, Hiroshi. 1962. Heiritsu no joshi tan nit suite. *Kokugakuin zasshi*, 63 (5). 66-76.

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

Kersten Alan W. and Dorrit Billman. 1997. Event category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23(3). 658--681.

König, Ekkehard and Carla Umbach. 2016. Demonstratives of Manner, of Quality and of Degree: A Neglected Subclass. In E. Boef, M. Coniglio, E. Schlachter and T. Veenstra (eds.), *Demonstratives: syntax, semantics and typology*. Berlin: de Gruyter Mouton.

Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*, 2nd edn. Chicago, IL: University of Chicago Press.

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

Lakoff, George. 1973. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic,* 2. 458 - 508.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.

Lakoff, Robin. 1971. If's, and's, and but's about Conjunction. In Fillmore, Charles and Terence D. Langendoen (eds.), *Studies in Linguistic Semantics*. New York: Holt, Reinhart and Winston, 115–150.

Lang, Ewald. 1984. *The semantics of coordination*. Amsterdam; Philadelphia: John Benjamins

Langacker, R. W. 1987a. Foundations of Cognitive Grammar. Vol. 1: Theoretical Prerequisites. Stanford: Stanford University Press.

Langacker, Ronald W. 1987b. Nouns and Verbs. *Language*, 63. 53-94.

Langacker, Ronald W. 1991a. *Concept, Image, and Symbol*. The Cognitive Basis of Grammar. Berlin, New York: De Gruyter Mouton.

Langacker, R. W. 1991b. Foundations of Cognitive Grammar, Vol. 2: Descriptive Application. Stanford: Stanford University Press.

Lee, Jong Hua. 2004. Fukujoshi "nado" ni kansuru goyoronteki bunscki. *Journal of the graduate school*. Taisho University, Taisho Daigaku Daigakuin kenkyu ronshu, 28. 322-311.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lewin, Bruno 1959. *Abriss der japanischen Grammatik auf der Grundlage der klassischen Schriftsprache*. Wiesbaden: Otto Harrassowitz.

Lischinsky, Alon. 2008. Examples as persuasive argument in popular management literature. *Discourse &. Communication,* 2(3). 243-269.

Longacre, Robert E. 1983. *The grammar of discourse*. New York, NY: Plenum Press.

Lucariello, Joan and Katherine Nelson. 1985. Slot-filler categories as memory organizers for young children. In *Developmental Psychology* 21. 272–82.

Lyons, John. 1977. *Semantics: Volume 1*. Cambridge: Cambridge University Press.

Lyons, John. 1968. *Introduction to Theoretical Linguistics*. London: Cambridge University Press.

Lyons, John D. 1989. *Exemplum: The Rhetoric of Example in Early Modern France and Italy*. Princeton: Princeton University Press.

Manzotti, Emilio. 1998. L'esempio. Natura, definizioni, problemi. *Cuadernos de Filología Italiana* 5. 99-123.

Martin, Samuel. 1975. *A Reference Grammar of Japanese*. New Haven: Yale University Press.

Masini, Francesca and Pietrandrea. Paola 2010. Magari. *Cognitive Linguistics,* 21. 75-121.

Mauri, Caterina. 2008. *Coordination Relations in the Languages of Europe and Beyond*. Berlin; New York: Mouton de Gruyter.

Mauri, Caterina. 2014. What do connectives and plurals have in common? The linguistic expression of ad hoc categories. In J. Blochowiak, S. Durrlemann-Tame, C. Grisot and C. Laenzlinger (eds.), *Linguistic papers dedicated to Jacques Moeschler*. Genève: University of Geneva Publication.

Mauri, Caterina. 2016. Building and interpreting ad hoc categories: a linguistic analysis. In *Formal models in the study of language*, Berlin: Springer. In print.

Mauri, Caterina and Andrea Sansò. forthcoming. Strategie linguistiche per la costruzione on-line di categorie: un quadro tipologico. In *Atti del XLIX Convegno della Società di Linguistica Italiana*.

Mauri, Caterina and Anna Giacalone. 2011. *Restricted indefiniteness: the case of Italian piuttosto che*. Paper presented at the 44th Annual Meeting of the SLE, 8-11 September 2011, Logrono, Spain.

Maynard, Senko. 1989. *Japanese conversation: Self contextualization through Structure and Interactional Management*. Norwood: Ablex.

Maynard, Senko. 1990. *An Introduction to Japanese Grammar and Communication Strategies*. Tokyo: Japan Times.

McEnery, Tony and Andrew Wilson. 1996. Corpus Linguistics. Edinburgh: Edinburgh University Press

Medin, Douglas L., and Marguerite M. Shaffer. 1978. Context theory of classification learning. *Psychological Review* 85. 207-238

Medin, Douglas L., Norbert O. Ross, Scott Atran, Douglas Cox , John Coley , Julia B. Proffittf, and Sergey Blok. 2006. Folkbiology of freshwater fish. *Cognition,* 99. 237– 73.

Meyer, Charles. 2002. *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.

Mihatsch, Wiltrud. 2007. The construction of vagueness: "sort of" expressions in Romance languages. In Radden, G., Köpke, K.-M., Berg, T. and Siemund, P. (eds.), *Aspects of meaning constructiong meaning: from concepts to utterances*. Amsterdam: Benjamins. 225-245.

Mihatsch, Wiltrud. 2009. The Approximators French *comme*, Italian *come*, Portuguese *como* and Spanish *como* from a Grammaticalization Perspective. In Rossari, C., Cojocariu, C., Ricci, C. and Spiridon, A. (eds.), *Grammaticalization and Pragmatics. Facts, Approaches, Theoretical Issues*. Bingley: Emerald. 65-91.

Mihatsch, Wiltrud. 2010a. *"Wird man von hustensaft wie so ne art bekifft?": Approximationsmarker in romanischen Sprachen*. Frankfurt am Main: Klostermann.

Mihatsch, Wiltrud. 2010b. The Diachrony of Rounders and Adaptors: Approximation and Unidirectional Change. In Kaltenböck, G., Mihatsch, W. and Schneider, S. (eds.), *New Approaches to Hedging*. Bingley: Emerald. 193-222.

Morita, Yoshiyuki. 1980. *Kiso Nihongo 2*. Tokyo: Kadokawa.

Murphy, Gregory L. 2002. *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, Gregory L. and Douglas L. Medin. 1985. The role of theories in conceptual coherence. *Psychological Review*, 92. 289-316.

Nagano, Takeshi (eds.). 1958. *Gakkoo bumpoo gaisetsu*. Tokyo.

Narrog, Heiko. 2009. A diachronic dimension in maps of case functions. *Linguistic Discovery*, 8. 233–254.

Narrog, Heiko. 2012. *Modality, Subjectivity, and Semantic Change. A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Narrog, Heiko and van der Auwera, Johan. 2011. Grammaticalization and semantic maps. In Heiko Narrog and Bernd Heine (eds.), *Handbook of Grammaticalization*. Oxford: OUP. 318–327.

Nonaka Ikujiro and Noboru Konno. 1998. The Concept of "Ba": Building a Foundation for Knowledge Creation. In *California Management Review*, 40 (3). 40-54.

Ohori, Toshio. 2004. Coordination in Mentalese. in Martin Haspelmath (eds.) *Coordinating Constructions*. Amsterdam: John Benjamins. 41-66.

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

Pearson, Jennifer. 1996. The Expression of Definitions in Specialised Texts: A Corpusbased Analysis. Euralex '96 Proceedings, Part II. Papers submitted to the Seventh Euralex International Congress on Lexicography in Göteborg, Sweden. Göteborg: Göteborg University, Department of Swedish. 759-769.

Perelman, Chaim. and Lucie Olbrechts-Tyteca. 1969. *The new rhetoric: A treatise on argumentation*. Translated by J. Wilkinson and P. Weaver. Notre Dame: University of Notre Dame Press.

Peterson, David A. and Kenneth Vanbik. 2004. Coordination in Hakha Lai (Tibeto-Burman). In M. Haspelmath (eds.), *Coordinating constructions*. Amsterdam; Philadelphia, Penn, John Benjamins. 333-356.

Pietrandrea, Paola. 2005. *Epistemic modality. Functional properties and the Italian system*. Amsterdam: Benjamins.

Polanyi, Michael. 1967. *The Tacit Dimension*. New York: Anchor Books.

Pólya, George. 1945. *How to Solve It*. Princeton: Princeton University Press.

Prince, Ellen F., Frader, Joel, Bosk, Charles and Dipietro, R.J. 1982. On Hedging in Physician-Physician Discourse. In Dipietro, R. J. (eds.), *Linguistics and the Professions*. Norwood/New Jersey: Ablex. 83-97.

Ramat, Paolo. 1992. Thoughts on degrammaticalization. *Linguistics*, 30. 549-560.

Recanati, François. 2004. Literal Meaning. Cambridge: Cambridge University Press.

Reisberg, Daniel. 1997. *Cognition: Exploring the science of the mind*. New. York: Norton.

Rips, Lance J. 1989. Similarity, typicality and categorization. In S.Vosniadou & A.Ortony (eds.), *Similarity and Analogical Reasoning*. Cambridge: Cambridge University Press.

Robert, Stéphane. 2008. Words and their meanings: Principles of variation and stabilization. In Simone Vanhove (eds.), *From Polysemy to Semantic Change: A Typology of Lexical Semantic Associations*. Amsterdam: John Benjamins. 55-92.

Rodríguez Abruñeiras, Paula. 2015. *Exemplifying markers in English: Synchronic and Diachronic considerations*. PhD Dissertation.

Rosch, Eleanor. 1973. Natural Categories. In *Cognitive Psychology* 4. 328-50.

Rosch, Eleanor. 1975. Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology: General* 104(3). 192–233.

Rosch, Eleanor. 1978. Principles of Categorization. In Rosch Eleanor and Barbara B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, N. J.: Lawrence Erlbaum Associates. 27-48.

Rosch, Eleanor, Carolyn B. Mervis. 1975. Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology*, 8. 382-439.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson and Penny Boyes-Bream. 1976. Basic objects in natural categories. *Cognitive Psychology*, 1. 491-502.

Ross, John Robert (1973). Nouniness. In O. Fujimura (eds.), *Three dimensions of linguistic theory*. Tokyo: TEC. 137–257.

Rossari, Corinne and Jacques Jaynez. 1999. Par exemple: une procédure d'exemplification par la preuve", in B. Combettes, C. Schnedecker and A. Theissen (eds.), *Ordre et distinction dans la langue et le discours*. Paris: Champion. 289-298.

Sawada, Osamu. 2016. The Japanese negative *totemo* 'very': toward a new typology of negative sensitive items. Paper presented at the 52nd Annual Meeting of the Chicago Linguistic Society. April 21, 2016.

Schneider, Stefan. 2007. *Reduced parenthetical clauses as mitigators. A corpus study of spoken French, Italian, and Spanish*. Amsterdam: Benjamins.

Searle, John and Daniel Vanderveken. 1985. Foundations of Illocutionary Logic. Cambridge: Cambridge University Press.

Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.

Sinclair, John. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.

Sperber, Dan and Deirdre Wilson. 1986/1995. *Relevance*. Oxford: Blackwell.

Sperber, Dan and Deirdre Wilson. 2002. Pragmatics, modularity and mind-reading. In *Mind and Language* 17. 3-23.

Smith, Edward E., and Steven A. Sloman. 1994. Similarity- versus rule-based categorization. *Memory & Cognition*, 22, 377-386.

Smith, Linda B. 2005. Emerging idea about categories. In L. Gershkoff-Stowe and D. H. Rakison (eds.), *Building Object Categories in Developmental Time*. Malwah, NJ: Lawrence Erlbaum Associates. 159-173.

Smith, Linda B. and Larissa K. Samuelson. 1997. Perceiving and remembering: category stability, variability and development. In Koen Lamberts and David Shanks (eds.), *Knowledge, concepts and categories*. Hove: Psychology Press. 161–95.

Stenmark, Dick. 2001. *The Relationship between Information and Knowledge*. In Proceedings of IRIS 24. Ulvik, Norway, August 11-14.

Suzuki, Satoko. 1998. Pejorative connotation: A case of Japanese. In A. Jucker and Y. Ziv (eds.), *Discourse Markers: Descriptions and theory*. Amsterdam: John Benjamins. 261-76.

Tanimori, Masahiro. 1994. *Handbook of Japanese Grammar*. Vermont, Tokyo: Tuttle Publishing.

Taylor, John R. (1995). *Linguistic Categorization. Prototypes in Linguistic Theory*, Oxford: Clarendon Press.

Taylor, Shelley E. and Suzanne C. Thompson. 1982. Stalking the elusive "vividness" effect. *Psychological Review* 89. 155–181.

Taylor, Yuki. 2010. *Functions of Japanese exemplifying particles in spoken and written discourse*. Ph.D. dissertation. University of California

Taylor, Yuki. 2015. The evolution of Japanese *toka* in utterance-final position. In Sylvie Hancil, Alexander Haselow & Margje Post (eds.), *Final particles*. Berlin & New York: Mouton de Gruyter. 141-156.

Terakura, Hiroko. 1983. Noun modification and the use of *to yuu*. *Journal of the Association of Teachers of Japanese*, 18 (1). 23–55.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.

Trager, George L. 1938. "Cottonwood-tree", a south-western linguistic trait. International *Journal of American Linguistics*, 9 (2/4). 117–118.

Tversky, Amos. 1977. Features of Similarity. *Psychological Review*, 84(4). 327-352.

Tversky, Amos and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185. 1124-1131.

van der Auwera, Johan. 2013. Semantic maps: For synchrony and diachrony. In Anna Giacalone Ramat, Caterina Mauri and Piera Molinelli (eds.), *Synchrony and. Diachrony: A dynamic interface*. Amsterdam: Benjamins. 153-176.

Vanhove, Martine. 2013. Similative, equative, and comparative constructions in Beja (NorthCushitic). version avant revisions. <halshs-00920198>

Voghera, Miriam. 2013. Tipi di tipo nel parlato e nello scritto. In *Di Linguistica e di Sociolinguistica, Studi offerti a Norbert Dittmar*. Roma: Bulzoni. 185-195.

Vovin, Alexander. 2003. *A Reference Grammar of Classical Japanese Prose*. Routledge Curzon

Yamaguchi, Gyoji. 1988. Jodooshi "nado" to sono shuuhen. *Rombun*, 50. 22-31.

Yamamoto, Shohei. 2004. Nihongo no bokashi hyoogen toka to kono eigo sootoo goku. *Language communication culture*. Vol. J, 1. 77-90.

Yu, Kang. 1995. *"Nado" "nadoka" no imi yoohoo to kono hensen. Kokubungaku ko*. Hiroshima: Hiroshima Daigaku Kokugo Kokubungakkai.

Yuasa, Etsuyo and Jerry Sadock. 2002. Pseudo-subordination: A mismatch between syntax and semantics. *Journal of Linguistics*, 38. 87-111.

Williamson, Timothy. 1994. *Vagueness*. Routledge.

Wilson, Deirdre. 2004. *Relevance and lexical pragmatics*. UCL Working Papers in Linguistics 16. 343-360.

Wilson, Deirdre and Robyn Carston. 2007. A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In Noel Burton-Roberts (eds.), *Pragmatics.* London: Palgrave. 230-259.

Wittgenstein, Ludwig. 1953/1978. *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford: Basil Blackwell.

Zhang, Niina Ning. 2008. Encoding exhaustivity. *USTWPL*, 4. 133-143.

Zillmann, Dolf. 1999. Exemplification theory: Judging the whole by some of its parts. *Media Psychology* 1. 69–94.

Zillmann, Dolf. 2002. Exemplification theory of media influence. In Jennings Bryant & Dolf Zillmann (eds.), *Media effects: Advances in theory and research*. London: Routledge. 19-41.

Zillmann, Dolf and Hans B. Brosius. 2000. *Exemplification in Communication: The Influence of Case Reports on the Perception of Issues*. Mahwah, NJ: Lawrence Erlbaum Associates.

APPENDIX A – THE CORPUS

The following is an excerpt of the table used to classify corpus-data (20 occurrences for each Japanese exemplifying construction). Its aim is to provide an instance of how the classification of occurrences has been brought about, on the basis of the parameters pointed out in chapter 2.
For the full document, go: http://www.leadhoc.org/index.php/data-access/ (available from February 2017)

OCCURRENCES OF *YA*

YA	Presence of the label	Syntactic types of labels	Position of the category label	Linguistic links	Semantic properties label	Syntactic properties of the example	Semantic properties of the example	Animacy	Number of examples	Exhaustivity	nado
LLC1	x	N	inside post	no	S	np	object	inanimate entity	2	no	x
LLC2	x	adj N	inside post	no	S	np	object	inanimate entity	3	no	x
LLC3	0					np	action	inanimate entity	3	no	
LLC4						np	object	animate entity	2	X	
LLC5	0					np	object	inanimate entity	2	no	x
LLC6	x	N	outside post	coma	S	np	object	inanimate entity	2	no	x
LLC7	0					np	action	inanimate entity	2	no	x
LLC8	x	adj N	inside post	no	S	np	object	inanimate entity	2	no	x
LLC9						np	item	inanimate entity	2	X	
LLC10	x	rel N	inside post	nado	S	np	object	inanimate entity	2	no	x
LLC11	0					np	animate	animate entity	3	no	x
LLC12	0					np	item	inanimate entity	3	no	x
LLC13	0					np	item	inanimate entity	2	no	
OR1	0					np	item	inanimate entity	3	no	
LLC14	0					np	object	inanimate entity	2	no	x
LLC15	x	N	inside post	no	S	np	item	inanimate entity	2	no	x
LLC16	0					np	object	inanimate entity	2	no	x
LLC17	0					np	action	inanimate entity	2	no	x
LLC18	0					np	object	inanimate entity	2	X	
LLC19	0					np	object	inanimate entity	2	X	x

YA	EC Syntactic structure	Aspect	Modality	Topic continuity	Types of SA	Types of categorization	Function
LLC1	n YA n NADO no CL	perfective	realis	category	declarative	similarity	categorization
LLC2	n YA n, n NADO no CL	perfective	realis	category	declarative	similarity	categorization
LLC3	n YA, n, n	perfective	epistemic		declarative	frame	categorization
LLC4	n YA n	iterative	realis		declarative		separative conjunction
LLC5	n YA n NADO	perfective	epistemic		declarative	similarity	categorization
LLC6	n YA n NADO, CL	perfective	epistemic	category	declarative	frame	categorization
LLC7	n YA n NADO	perfective	realis	category	declarative	similarity	categorization
LLC8	n YA n NADO no CL	perfective	realis		declarative	similarity	categorization
LLC9	n YA n	perfective	realis		declarative		separative conjunction
LLC10	n YA n NADO no CL	perfective	realis	category	declarative	frame	categorization
LLC11	n YA n, n NADO	perfective	epistemic	category	declarative	similarity	categorization
LLC12	n YA n, n NADO	perfective	epistemic		declarative	similarity	categorization
LLC13	n YA n	perfective	epistemic		declarative	similarity	categorization
OR1	n YA n, n	perfective	epistemic		directive	similarity	categorization
LLC14	n YA n NADO	perfective	realis		declarative	similarity	categorization
LLC15	n YA n NADO no CL	perfective	realis	category	declarative	frame	categorization
LLC16	n YA n NADO	perfective	realis		declarative	similarity	categorization
LLC17	n YA, n NADO	perfective	epistemic		directive	frame	categorization
LLC18	n YA n	perfective	realis		declarative	frame	categorization
LLC19	n YA n NADO	perfective	realis	category	declarative	similarity	categorization

OCCURRENCES OF *NADO*

NADO	Presence of the label	Syntactic types of labels	Position of the category label	Linguistic links	Comma	Semantic properties label	Syntactic properties of the example	Semantic properties of the example	Animacy	Number of examples	Exhaustivity	tari	toka	ya
LLC1	x	adj N	outside post		comma	G	vp	action		2	no	x		
LLC2	x	N	inside post	no		S	np	object	inanimate entity	2	no			x
LLC3	x	adj N	inside post	no		S	np	object	inanimate entity	3	no			x
LLC4	x	N	outside pre			S	np	item	inanimate entity	1	no			
LLC5	x	mix	outside post		comma	S	vp	action		2	no	x		
LLC6	0						vp	action	inanimate entity	2	no	x		
LLC7	0						np	object	inanimate entity	2	no			x
LLC8	0						vp	action		1	no			
LLC9	x	N	outside post		comma	S	np	action	inanimate entity	2	no			
LLC10	0						np	action	inanimate entity	2	no			x
LLC11	x	adj N	inside post	no		S	np	item	inanimate entity	2	no			x
LLC12	x	NN	outside pre			S	np	object	inanimate entity	3	no			
LLC13	x	rel N	inside post		comma	S	np	object	inanimate entity	2	no			x
LLC14	0						np	object	inanimate entity	1	no			
LLC15	0						np	item	inanimate entity	1	no			
OR1	0						np	object	inanimate entity	2	no			
LLC16	0						np	object	inanimate entity	2	no			
LLC17	0						np	item	inanimate entity	4	no			x
LLC18	x	NN	inside post	no		S	np	item	inanimate entity	3	no			
LLC19	x	gen N	Inside post	no		G	vp	action		1	no			

NADO	EC Syntactic structure	Aspect	Modality	Topic continuity	Types of SA	Types of categorization	Function
LLC1	v TARI v NADO, CL	perfective	evidentiality	category	declarative	frame	categorization
LLC2	n YA n NADO no CL	perfective	realis	category	declarative	similarity	categorization
LLC3	n YA n, n NADO no CL	perfective	realis	category	declarative	similarity	categorization
LLC4	n NADO	perfective	realis	category	declarative	similarity	categorization
LLC5	v TARI v NADO, CL	perfective	realis	category	declarative	frame	categorization
LLC6	v TARI v TARI toki NADO	habitual	realis		declarative	frame	categorization
LLC7	n YA n NADO	perfective	epistemic	category	directive	similarity	categorization
LLC8	v NADO	perfective	realis		declarative	similarity	categorization
LLC9	n YA n NADO, CL	perfective	epistemic	category	declarative	frame	categorization
LLC10	n YA n NADO	imperfective	realis	category	declarative	similarity	categorization
LLC11	n YA n NADO no CL	perfective	realis	category	declarative	similarity	categorization
LLC12	n NADO	perfective	realis	category	declarative	similarity	categorization
LLC13	n YA n NADO no CL	perfective	realis	category	declarative	frame	categorization
LLC14	n NADO	perfective	realis	category	declarative	similarity	categorization
LLC15	n NADO	perfective	realis	category	declarative	similarity	categorization
OR1	n, n NADO	perfective	epistemic	category	declarative	similarity	categorization
LLC16	n, n NADO	perfective	epistemic	category	declarative	similarity	categorization
LLC17	n YA n, n NADO	perfective	epistemic		declarative	similarity	categorization
LLC18	n, n, n NADO no CL	perfective	realis	category	declarative	similarity	categorization
LLC19	V NADO no CL	perfective	realis		declarative	frame	categorization

OCCURRENCES OF *TARI*

TARI	Presence of the label	Syntactic types of labels	Position of the category label	Linguistic links	Semantic properties label	Syntactic properties of the example	Type of Strategy	Exhaustivity	Number of examples	toka	nado	EC Syntactic structure
LLC1	0					vp	connective	no	3			vp TARI vp TARI vp
LLC2	0					vp	connective	no	3			vp TARI vp TARI vp suru
LLC3	0					vp	connective	no	3			vp TARI vp TARI suru
LLC4	0					vp	connective	no	4			vp TARI vp TARI suru
LLC5	x	adj N	outside post	comma	G	vp	connective	no	2		x	vp TARI vp NADO, CL
LLC6	0					vp	connective	no	2			vp TARI vp TARI suru
LLC7						vp	connective	X	2			vp TARI vp TARI suru
LLC8	0					vp	connective	no	2			vp TARI vp TARI suru
LLC9	0					vp	connective	no	2			vp TARI vp
LLC10	0					vp	connective	no	2			vp TARI vp TARI suru
LLC11	x	mix	outside post	comma	S	vp	connective	no	2		x	vp TARI vp suru NADO, CL
LLC12	0					vp	connective	no	2			vp TARI vp TARI suru
LLC13	0					vp	connective	no	2			vp TARI vp suru
LLC14	0					vp	connective	no	2			vp TARI vp TARI toiu CL
LLC15						vp	general extender	X	1			vp TARI suru
LLC16	0					vp	connective	no	2			vp TARI vp TARI suru
LLC17	0					vp	connective	no	2			vp TARI vp suru
LLC18	0					vp	connective	no	2			vp TARI vp TARI suru
LLC19						vp	connective	X	2			vp TARI vp TARI suru
LLC20	0					vp	connective	no	2			Vp TARI vp TARI

TARI	position suru	Aspect	Modality	Topic continuity	Types of SA	Types of categorization	Function
LLC1	no	habitual	realis	category	declarative	similarity	categorization
LLC2	suru	perfective	epistemic		declarative	similarity	categorization
LLC3	linked suru	imperfective	realis	category	declarative	frame	categorization
LLC4	linked suru	perfective	realis		declarative	similarity	categorization
LLC5	no	perfective	evidentiality	category	declarative	frame	categorization
LLC6	linked suru	perfective	realis		declarative	similarity	categorization
LLC7	linked suru	perfective	epistemic		declarative		free alternative
LLC8	linked suru	perfective	epistemic		declarative	similarity	categorization
LLC9	no	iterative	epistemic		declarative	frame	categorization
LLC10	linked suru	iterative	realis		declarative	frame	categorization
LLC11	suru	perfective	realis	category	declarative	frame	categorization
LLC12	linked suru	perfective	realis		declarative	similarity	categorization
LLC13	suru	iterative	epistemic	category	declarative	frame	categorization
LLC14	no	iterative	epistemic	category	declarative	frame	categorization
LLC15	linked suru	perfective	realis		declarative		hedging
LLC16	linked suru	iterative	realis		declarative	frame	categorization
LLC17	suru	perfective	epistemic		declarative	similarity	categorization
LLC18	linked suru	perfective	epistemic		declarative	similarity	categorization
LLC19	linked suru	iterative	epistemic		declarative		free alternative
LLC20	no	perfective	epistemic		directive	similarity	categorization

OCCURRENCES OF *TOKA*

TOKA	Presence of the label	Syntactic types of labels	Position of the category label	Linguistic links	Semantic properties label	Syntactic properties of the example	Semantic properties of the example	Animacy	Type of Strategy	Exhaustivity	Number of examples	tari
LLC1	0					np	object	inanimate entity	general extender	no	1	
LLC2						vp	item	inanimate entity	connective	X	2	
LLC3	x	N	inside post	nado	G	vp/np	action	inanimate entity	connective	no	2	
LLC4	0					np	object	inanimate entity	general extender	no	1	
OR1	0					np	item	inanimate entity	connective	no	2	
OR2	x	N	outside pre		G	np	action	inanimate entity	general extender	no	3	x
OR3	0					np	item	inanimate entity	general extender	no	1	
OR4	0					np	item	inanimate entity	connective	no	3	
OR5	x	Rel N	outside pre		S	np	object	inanimate entity	connective	no	2	
OR6	x	adj N	outside post	comma	S	np	object	inanimate entity	connective	no	2	
OR7	x	adj N	outside post	comma	S	np	object	inanimate entity	connective	no	2	
LLC5	x	adj N	outside pre		S	np	item	inanimate entity	general extender	no	3	
LLC6	0					np	item	inanimate entity	general extender	no	1	
LLC7	x	gen N	outside post	comma	S	np	item	inanimate entity	general extender	no	1	
LLC8	0					vp	action	inanimate entity	connective	no	2	
LLC9	0					np	item	inanimate entity	connective	no	2	
LLC10						np	object	inanimate entity	connective	X	2	
LLC11						vp	action		general extender	X	1	
LLC12	0					vp	action	inanimate entity	connective	no	2	
LLC13	0					np	object	Inanimate entity	general extender	no	1	

TOKA	nado	EC Syntactic structure	Aspect	Modality	Types of SA	Topic continuity	Types of categorization	Function
LLC1		n TOKA	habitual	realis	declarative	category	similarity	categorization
LLC2		n TOKA n TOKA	perfective	epistemic	directive			free alternative
LLC3	x	v TOKA, n NADO, CL	perfective	realis	declarative		similarity	categorization
LLC4		n TOKA	habitual	realis	declarative		similarity	categorization
OR1		n TOKA n toitta KIND	perfective	realis	declarative	category	similarity	categorization
OR2		v TARI, v TARI n TOKA	perfective	realis	declarative	category	frame	categorization
OR3		n TOKA	perfective	realis	declarative	category	similarity	categorization
OR4		n TOKA, n, n	perfective	realis	declarative		similarity	categorization
OR5		n TOKA n TOKA	perfective	realis	declarative	category	similarity	categorization
OR6		n TOKA n TOKA, CL	perfective	realis	declarative	category	similarity	categorization
OR7		n TOKA n TOKA, CL	perfective	epistemic	directive	category	similarity	categorization
LLC5		n, n, n TOKA	perfective	realis	declarative	category	similarity	categorization
LLC6		n TOKA	perfective	realis	declarative	category	frame	categorization
LLC7		n TOKA, CL	perfective	realis	directive	category	similarity	categorization
LLC8	x	n TOKA n TOKA	iterative	epistemic	declarative	category	similarity	categorization
LLC9		n TOKA n TOKA	perfective	realis	declarative	category	frame	categorization
LLC10		n TOKA n TOKA	habitual	realis	declarative			separative conjunction
LLC11		v TOKA	perfective	realis	declarative			approximator
LLC12	x	n TOKA n TOKA	iterative	realis	declarative		similarity	categorization
LLC13		n TOKA	perfective	realis	declarative		similarity	categorization

APPENDIX B – THE QUESTIONNAIRE

The following is the questionnaire used to verify the specific readings identified, both in the Japanese and in the English version. Tables include the percentages of responses for each paraphrase.
1 = not appropriate; 2 = neutral; 3 = very appropriate.

JAPANESE VERSION

1. そのたびに占い師や霊感の強い人に見てもらったり、お払いをしてもらったりするのですが、相変わらず不運は続きます。 **source: http://news.mynavi.jp/column/smb/008/**

	3 (very appropriate)	2	1 (not appropriate)	Response Total
ある時は占い師のところへ行く。またある時は、霊感の強い人のところへ行く。	48.3% (14)	51.7% (15)	0.0% (0)	29
ある時は占い師のところへ行く。またある時は、霊感の強い人のところへ行く。またある時は、その他の人のところへ行く。	44.8% (13)	34.5% (10)	20.7% (6)	29
占い師、霊感の強い人、その他の人のところへ同時に行くことが、時々ある。	6.9% (2)	31.0% (9)	62.1% (18)	29

2. 日本全体で温室効果ガスの排出を減らす２０２０年までの中期目標について、麻生首相は８日、緩やかな削減幅を主張する経済界や労働組合の代表相次いで会談し、意見を聴いた。 **source: http://www.asahi.com/eco/TKY200906080324.html**

	3	2	1	Response Total
首相は経済界の代表と会談し、それから労働組合の代表と会談した。	86.2% (25)	10.3% (3)	3.4% (1)	29
首相は経済界の代表と会談し、それから労働組合の代表と会談した。それから、その他の人とも会談した。	6.9% (2)	48.3% (14)	44.8% (13)	29

3. 異常が検知された場合は、メールによる通知やセッション切断するといった対応も可能という。 **source:**
http://cloud.watch.impress.co.jp/epw/cda/software/2008/06/09/13128.html

	3	2	1	Response Total
ある場合はセッションが切断される。また、別の場合はメールで通知される。	62.1% (18)	27.6% (8)	10.3% (3)	29
ある場合はセッションが切断され、別の場合はメールで通知されるが、その他の対応がとられることもある。	31.0% (9)	44.8% (13)	24.1% (7)	29
メール通知に加えてセッション切断が同時に行われる。	10.3% (3)	41.4% (12)	48.3% (14)	29
メール通知、セッション切断、その他の対応が、同時に行われる。	0.0% (0)	31.0% (9)	69.0% (20)	29

	3	2	1	Response Total
このポータルサイトは、PC 初心者、中級者の双方にわかりやすい。	72.4% (21)	27.6% (8)	0.0% (0)	29
このポータルサイトは、PC 初心者、中級者、それ以外の人にもわかりやすい。	37.9% (11)	44.8% (13)	17.2% (5)	29
このポータルサイトは、PC 初心者、中級者、上級者にわかりやすい。	17.2% (5)	69.0% (20)	13.8% (4)	29

	3	2	1	Response Total
ゴム状硫黄は、純度 99%の硫黄華から作られることがある。純度 98%の硫黄粉末から作られることもある。	69.0% (20)	31.0% (9)	0.0% (0)	29
ゴム状硫黄は、純度 99%の硫黄華から作られることがある。純度 98%の硫黄粉末から作られることもある。 さらに、他のタイプの硫黄から作られることもある。	34.5% (10)	10.3% (3)	55.2% (16)	29
ゴム状硫黄は、黒または褐色である。	69.0% (20)	17.2% (5)	13.8% (4)	29
ゴム状硫黄は、黒、褐色、その他の色の場合がある。	20.7% (6)	24.1% (7)	55.2% (16)	29

	3	2	1	Response Total
自治体の担当者と意見交換をした。その後、政府の担当者と意見交換をした。	44.8% (13)	51.7% (15)	3.4% (1)	29
自治体の担当者と意見交換をした。その後、政府の担当者と意見交換をした。そしてその後、その他の人々と意見交換をした。	10.3% (3)	48.3% (14)	41.4% (12)	29
自治体の担当者に加えて政府の担当者と同時に意見交換をした。	20.7% (6)	58.6% (17)	20.7% (6)	29
自治体の担当者、政府の担当者、およびその他の人々と同時に意見交換をした。	10.3% (3)	24.1% (7)	65.5% (19)	29
自治体と政府は、復興、防災行政に携わっている。	58.6% (17)	27.6% (8)	13.8% (4)	29
自治体と政府は、復興、防災行政、その他のことに携わっている。	24.1% (7)	37.9% (11)	37.9% (11)	29

7. 「 戦地 」 への 派遣 なるため 国会 で 議論 と なった が 、 政府 は 「 自衛隊 が 活動 する 地域 は 非 戦闘 地域 」 など と 主張 。 source: http://blog.goo.ne.jp/neo_japan21/e/a2f93761eba4ab6846bcf757d5726407

	3	2	1	Response Total
政府は「自衛隊が活動する地域は非戦闘地域」に類似することを主張した。	17.2% (5)	34.5% (10)	48.3% (14)	29
政府は「自衛隊が活動する地域は非戦闘地域」という主張に加え、その他の主張をした。	48.3% (14)	34.5% (10)	17.2% (5)	29
政府は「自衛隊が活動する地域は非戦闘地域」に類似することを主張し、加えてその他の主張をした。	6.9% (2)	37.9% (11)	55.2% (16)	29
政府は「自衛隊が活動する地域は非戦闘地域」と主張した。	51.7% (15)	44.8% (13)	3.4% (1)	29

8. 国民 自由党 ０７年 の 欧州 連合 （ ＥＵ ） 加盟 や ９％ の 経済 成長 率 など の 実績 強調 した 、 貧富 の 差 拡大 など に対する 国民 の 不満 響いた 。

	3	2	1	Response Total
民衆は様々な問題、特に貧富の差の拡大に、不満を持っていた。	69.0% (20)	27.6% (8)	3.4% (1)	29
民衆は様々な問題、つまり貧富の差の拡大やその他の問題に、不満を持っていた。	48.3% (14)	48.3% (14)	3.4% (1)	29

9. 自民党側からは「機構は雇用保険の無駄遣いのシンボル。廃止出来なければ、自民党の真価が問われる」（渡辺喜美元行革担当相）などと批判が噴出した。 source: http://blog.goo.ne.jp/mamorukai_kunren/c/b810ad59916608cdf725a1b424bc6f09/3

	3	2	1	Response Total
「機構は雇用保険の無駄遣いのシンボル。廃止できなければ、自民党の真価が問われる」という批判があった。	41.4% (12)	55.2% (16)	3.4% (1)	29
「機構は雇用保険の無駄遣いのシンボル。廃止できなければ、自民党の真価が問われる」という批判に加え、その他の批判があった。	62.1% (18)	17.2% (5)	20.7% (6)	29

10. こうした投資のアプローチについて社会責任投資（SRI）などと呼ばれています。 source: http://cloud.watch.impress.co.jp/epw/cda/software/2008/06/09/13128.html

	3	2	1	Response Total
SRI に似た、ひとつの名前でのみ呼ばれている。	6.9% (2)	17.2% (5)	75.9% (22)	29
SRI と、その他の名前で呼ばれている。	41.4% (12)	48.3% (14)	10.3% (3)	29
SRI に似た名前と、その他の名前で呼ばれている。	0.0% (0)	17.2% (5)	82.8% (24)	29

10. こうした投資のアプローチについて社会責任投資（**SRI**）などと呼ばれています。 **source: http://cloud.watch.impress.co.jp/epw/cda/software/2008/06/09/13128.html**

	3	2	1	Response Total
様々な名で呼ばれるが、特に SRI と呼ばれている。	44.8% (13)	44.8% (13)	10.3% (3)	29
SRI と呼ばれている。	37.9% (11)	58.6% (17)	3.4% (1)	29

11. まず 最初 に 、今 の ところ 非常 に まだまだ 治す の が むずかしい 病気 とか ケガ で若い 人 や 子ども さん かかる 可能 性 の ある もの 三つ だけ ご 紹介 します 。 **source: http://ameblo.jp/regenerative-kyoto/theme29-10006816417.html**

	3	2	1	Response Total
若い人や子どものかかる可能性があるものは、ケガによるもの、または病気によるものである。	31.0% (9)	34.5% (10)	34.5% (10)	29
若い人や子どものかかる可能性があるものは、ケガによるもの、病気によるもの、まはたその他のものによるものである。	24.1% (7)	44.8% (13)	31.0% (9)	29
若い人や子どもは、ケガ、病気に同時にかかる可能性がある。	10.3% (3)	37.9% (11)	51.7% (15)	29
若い人や子どもはケガ、病気、そして他のものに同時にかかる可能性がある。	3.4% (1)	24.1% (7)	72.4% (21)	29

12. 僕、高校の頃から、理想とか、自分に合うものを探してたんです **source: http://www.e-labo.net/meister/2008/02/seo.php**

	3	2	1	Response Total
僕は何かを探していて、それは例えば理想だった。	72.4% (21)	20.7% (6)	6.9% (2)	29
僕は理想に似た何かを探していた。	6.9% (2)	31.0% (9)	62.1% (18)	29
僕は何かを探していて、それは理想やその他のものだった。	44.8% (13)	41.4% (12)	13.8% (4)	29
僕は理想を探していた。（ほかにはない）	3.4% (1)	31.0% (9)	65.5% (19)	29

13. なお 、I 型 （インスリン 依存 型） とか II 型 （インスリン 非 依存 型） という 表現は 現存 は 使い ませ ん ので 注意 して ください 。 **source: http://allabout.co.jp/gm/gc/300397/**

	3	2	1	Response Total
I 型、または II 型という表現は使わない。	96.6% (28)	0.0% (0)	3.4% (1)	29

13. なお 、 Ⅰ型 （ インスリン 依存 型 ） とか Ⅱ型 （ インスリン 非 依存 型 ） という 表現 は 現存 は 使い ませ ん ので 注意 して ください 。 source: http://allabout.co.jp/gm/gc/300397/

	3	2	1	Response Total
Ⅰ型、またはⅡ型、またはその他の表現は使わない。	0.0% (0)	27.6% (8)	72.4% (21)	29

14. ネット上でも、「最近顔が丸くなった」「薬とかの副作用じゃないかと心配」「アルコールの飲みすぎでむくんだ？」と様々な憶測を呼んでいる (gossip about a famous woman who got fat recently) source: http://www.j-cast.com/2008/04/28019559.html

	3	2	1	Response Total
ネット上の人々は、薬やその他の副作用を心配している。（いくつか薬や薬物を飲んでいたのでは）	65.5% (19)	31.0% (9)	3.4% (1)	29
ネット上の人々は、薬の副作用を心配しているが、「薬」とはっきり言いたくない。	31.0% (9)	31.0% (9)	37.9% (11)	29

15. 地元 消防 団員 と 町 職員 計 ７０人 が 徒歩 と 車 で 入り 、 浸水 した 家屋 や 道路 に 散乱 した 流木 など を 片づけ たり 、 食料 の 配給 を したりした 。 source: http://blog.goo.ne.jp/jun-propela/e/c1a900c4aac7f13df8f57648e4b7304f

	3	2	1	Response Total
地元消防団員と町職員は流木を片付け、食糧を配給し、その他の活動を行った。	79.3% (23)	20.7% (6)	0.0% (0)	29
地元消防団員数人と町職員数人は流木を片付けた。ほかの地元消防団員と町職員は食料を配給した。	13.8% (4)	37.9% (11)	48.3% (14)	29
地元消防団員数人と町職員数人は流木を片付けた。ほかの地元消防団員と町職員は食料を配給した。そして、残りの地元消防団員と残りの町職員はその他の活動を行った。	6.9% (2)	37.9% (11)	55.2% (16)	29

16. また、ゆったりとくつろげる音楽を聴いたり、自分が心地良いと感じる香りを嗅ぐのもおすすめ。さらに肌寒い季節には、ホットミルクなどの温かい飲み物を飲むと、体が温まり、寝つきが良くなります。 source: http://allabout.co.jp/1/225879/1/product/225879_08.htm

	3	2	1	Response Total
香りを嗅ぐ、音楽を聴く、この両方をしても、どちらかだけでもよい。	79.3% (23)	17.2% (5)	3.4% (1)	29
香りを嗅ぐ、音楽を聴く、その他のこと、これら全部をしても、どれかだけでもよい。	24.1% (7)	44.8% (13)	31.0% (9)	29
香りを嗅ぐ、音楽を聴く、この両方をしてもよい。	34.5% (10)	48.3% (14)	17.2% (5)	29
香りを嗅ぐ、音楽を聴く、その他のことをしてもよい。	13.8% (4)	48.3% (14)	37.9% (11)	29

16. また、ゆったりとくつろげる音楽を聴いたり、自分が心地良いと感じる香りを嗅ぐのもおすすめ。さらに肌寒い季節には、ホットミルクなどの温かい飲み物を飲むと、体が温まり、寝つきが良くなります。 source: http://allabout.co.jp/1/225879/1/product/225879_08.htm

	3	2	1	Response Total
香りを嗅ぐか、音楽を聴くという、二つの選択肢がある。	20.7% (6)	17.2% (5)	62.1% (18)	29
香りを嗅いで音楽を聴くという、一つしか選択肢がない。	3.4% (1)	0.0% (0)	96.6% (28)	29

17. これにより、新規ユーザを追加したり同僚の連絡先を調べたりする際に、当て推量をしなくてもよくなった。 source: http://sourceforge.jp/magazine/07/08/21/0134229

	3	2	1	Response Total
新規ユーザーを追加する時、または、同僚の連絡先を調べる時のいずれか。	34.5% (10)	41.4% (12)	24.1% (7)	29
新規ユーザーを追加する時、同僚の連絡先を調べる時、またはその他のことをする時のいずれか。	48.3% (14)	24.1% (7)	27.6% (8)	29
新規ユーザーを追加する時、または、同僚の連絡先を調べる時の、２つの場合。	34.5% (10)	51.7% (15)	13.8% (4)	29
新規ユーザーを追加し、かつ同僚の連絡先を調べる時。	0.0% (0)	13.8% (4)	86.2% (25)	29
新規ユーザーを追加し、同僚の連絡先を調べ、かつその他のことをする時。	0.0% (0)	6.9% (2)	93.1% (27)	29

18. しかし この よう な リーダーシップ は 、欲求 不満 を 満たす 手段 を 失っ たり 、あるいは エスカレート し すぎ た とき に 、崩壊 へ の 道 を たどり ます 。 source: http://allabout.co.jp/gm/gc/294079/

	3	2	1	Response Total
以下の二つの場合がある。欲求不満を満たす手段を失った時、または、エスカレートしすぎた時。	93.1% (27)	6.9% (2)	0.0% (0)	29
以下のいずれか一つの場合である。欲求不満を満たす手段を失った時、または、エスカレートしすぎた時。	6.9% (2)	44.8% (13)	48.3% (14)	29
以下のいずれか一つの場合である。欲求不満を満たす手段を失った時、または、エスカレートしすぎた時、またはその他のことがおきた時。	3.4% (1)	37.9% (11)	58.6% (17)	29
欲求不満を満たす手段を失い、かつ、エスカレートしすぎた時。	0.0% (0)	10.3% (3)	89.7% (26)	29
欲求不満を満たす手段を失い、エスカレートしすぎ、かつその他のことがおきた時。	0.0% (0)	6.9% (2)	93.1% (27)	29

19. あなた の コンピュータ に は 空き 領域 (緑 の 領域) は 十分 に あり ますか？。 そう で ない なら 、 スワップ サイズ を 大きく し たり 、 物理 メモリ を 増やす こと で 対処 でき ます 。

	3	2	1	Response Total
スワップサイズを大きくする、または、物理メモリを増やす。この二つの対処法があり、二つをやってもいいし、どちらか一つでもいい。	96.6% (28)	0.0% (0)	3.4% (1)	29
スワップサイズを大きくする、または、物理メモリを増やす、またはその他のことをする。これら全てをやってもいいし、一つだけをやってもいい。	10.3% (3)	41.4% (12)	48.3% (14)	29
スワップサイズを大きくし、物理メモリを増やす。この一つの対処法しかない。	0.0% (0)	10.3% (3)	89.7% (26)	29
スワップサイズを大きくし、物理メモリを増やし、その他のことをする。この一つの対処法しかない。	0.0% (0)	3.4% (1)	96.6% (28)	29

20. 写真 を 精査 すれ ば 多少 の ノイズ 感 や 細部 の 潰れ が ある もの の 「 キレイ だ から オッケー !!」 と 思え たり し て 。 source: http://k-tai.impress.co.jp/cda/article/stapa/36839.html

	3	2	1	Response Total
「キレイだからオッケー！！」に類似したことを思えるかもしれない。	41.4% (12)	41.4% (12)	17.2% (5)	29
「キレイだからオッケー！！」に加え、その他のことを思えるかもしれない。	3.4% (1)	34.5% (10)	62.1% (18)	29
「キレイだからオッケー！！」に類似することに加え、その他のことを思えるかもしれない。	10.3% (3)	20.7% (6)	69.0% (20)	29
「キレイだからオッケー！！」と思えるかもしれない。	72.4% (21)	20.7% (6)	6.9% (2)	29

307

ENGLISH VERSION

	1 (not appropriate)	2	3 (very appropriate)	Response Total
On some occasions, he visited fortune-tellers. On some occasions, he visited people with a strong inspiration.	0.0% (0)	66.7% (4)	33.3% (2)	6
On some occasions, he visited fortune-tellers. On some occasions, he visited people with a strong inspiration. On some occasions, he visited other similar people.	50.0% (3)	33.3% (2)	16.7% (1)	6
Several times, he visited both fortune-tellers, people with a strong inspiration and other similar people. Every time he visited them at the same time.	100.0% (6)	0.0% (0)	0.0% (0)	6

	1	2	3	Response Total
The prime minister spoke with the representatives of the labour unions. Then, he spoke with the representatives of the business community.	0.0% (0)	33.3% (2)	66.7% (4)	6
The prime minister spoke with the representatives of the labour unions. Then he spoke with the representatives of the business community. Then, he spoke with other similar people.	33.3% (2)	33.3% (2)	33.3% (2)	6

	1	2	3	Response Total
In case an irregularity is detected, on some occasions, software support consists in the cutting of the session. On some other occasions, software support consists in notification by email.	0.0% (0)	83.3% (5)	16.7% (1)	6
In case an irregularity is detected, on some occasions, software support consists in the cutting of the session. On some other occasions, software support consists in notification by email. On some other occasions, software support consists in other similar things.	33.3% (2)	16.7% (1)	50.0% (3)	6
In case an irregularity is detected, software support consists in one solution: cutting of the session together with notification by email.	83.3% (5)	16.7% (1)	0.0% (0)	6
In case an irregularity is detected, software support consists in one solution: cutting of the session, notification by email and other similar things as well together at the same time.	100.0% (6)	0.0% (0)	0.0% (0)	6

4. PC 初心者 や 中級 者 も 理解 し やすい よう に 、 **HDD** に関する 基礎 知識 や 、 **HDD** の 増 設 ・ 交換 ・ 自作 方法 、 トラブル 時 の 対処 方法 や データ の バックアップ 方法 といった 実用 的 な 知識 を 、 図 や 写真 を 交え て 分かり やすく 紹介 する ポータルサイト 。 **source: https://osdn.jp/magazine/08/07/02/2253207**

	1	2	3	Response Total
The internet portal is easy for PC beginners and also for intermediate learners.	0.0% (0)	16.7% (1)	83.3% (5)	6
The internet portal is easy for PC beginners, for intermediate learners and other people.	83.3% (5)	16.7% (1)	0.0% (0)	6
The internet portal is easy for PC beginners, intermediate learners and advanced learners.	66.7% (4)	33.3% (2)	0.0% (0)	6

5. 純度 ９８％ の 硫黄 粉末 や ９９％ の 硫黄 華 で 作っ た ゴム 状 硫黄 は 褐色 や 黒色 で 、 試験 管 に 黒い 物質 が 残っ た 。 **source: http://raitu.tumblr.com/post/68665329/5-98-99**

	1	2	3	Response Total
Sometimes rubber-like Sulphur is made by flower of Sulphur 99% pure. On other occasions, it is made by Sulphur powder 98% pure.	16.7% (1)	33.3% (2)	50.0% (3)	6
Sometimes rubber-like Sulphur is made by flower of Sulphur 99% pure. On other occasions, it is made by Sulphur powder 98% pure. On other occasions, it is made by other types of Sulphur powder.	83.3% (5)	16.7% (1)	0.0% (0)	6
Rubber-like Sulphurs are sometimes black and sometimes brown.	0.0% (0)	50.0% (3)	50.0% (3)	6
Rubber-like Sulphurs are sometimes black, sometimes brown and sometimes they are of other colours.	83.3% (5)	0.0% (0)	16.7% (1)	6

6. 中越 地震 の 復興 や 防災 行政 に 携わる 自治体 や 政府 の 担当 者 と 意見 交換 する ほか 、 地 震 で 被害 を 受け た 柏崎 原発 や 防災 施設 など も 視察 する. **source: http://blog.livedoor.jp/tson314/archives/52215791.html**

	1	2	3	Response Total
They had a meeting with the representatives of the municipalities. Then they had a meeting with the representatives of the Government.	0.0% (0)	66.7% (4)	33.3% (2)	6
They had a meeting with the representatives of the municipalities. Then they had a meeting with the representatives of the Government. Then they had a meeting with other similar people.	66.7% (4)	16.7% (1)	16.7% (1)	6
They had a meeting with the representatives of the municipalities and the representatives of the Government at the same time.	16.7% (1)	50.0% (3)	33.3% (2)	6
They had a meeting with the representatives of the municipalities, the representatives of the Government and other similar people at the same time.	66.7% (4)	16.7% (1)	16.7% (1)	6

	1	2	3	Response Total
Municipalities and the Government are involved in the administration of disaster prevention and in the reconstruction of the earthquake center.	16.7% (1)	33.3% (2)	50.0% (3)	6
Municipalities and the Government are involved in the administration of disaster prevention, in the reconstruction of the earthquake center and in other similar things.	66.7% (4)	16.7% (1)	16.7% (1)	6

	1	2	3	Response Total
The Government claimed something similar to "自衛隊 が 活動 する 地域 は 非 戦闘 地域 ".	0.0% (0)	50.0% (3)	50.0% (3)	6
The Government claimed "自衛隊 が 活動 する 地域 は 非 戦闘 地域 " and other similar statements.	50.0% (3)	16.7% (1)	33.3% (2)	6
The Government claimed something similar "自衛隊 が 活動 する 地域 は 非 戦闘 地域 ", and other similar statements.	16.7% (1)	33.3% (2)	50.0% (3)	6
The Government claimed "自衛隊 が 活動 する 地域 は 非 戦闘 地域 ".	16.7% (1)	66.7% (4)	16.7% (1)	6

	1	2	3	Response Total
The public discontent focused on different problems, but in particular on the expansion of the gap between poor and rich.	33.3% (2)	0.0% (0)	66.7% (4)	6
The public discontent focused on different problems, the expansion of the gap between poor and rich and other similar problems.	0.0% (0)	83.3% (5)	16.7% (1)	6

	1	2	3	Response Total
Criticisms such as "機構は雇用保険の無駄遣いのシンボル。廃止出来なければ、自民党の真価が問われる"	16.7% (1)	50.0% (3)	33.3% (2)	6

9. 自民党側からは「機構は雇用保険の無駄遣いのシンボル。廃止出来なければ、自民党の真価が問われる」（渡辺喜美元行革担当相）などと批判が噴出した。 source: http://blog.goo.ne.jp/mamorukai_kunren/c/b810ad59916608cdf725a1b424bc6f09/3

	1	2	3	Response Total
Criticisms such as "機構は雇用保険の無駄遣いのシンボル。廃止出来なければ、自民党の真価が問われる" and other similar criticisms.	16.7% (1)	16.7% (1)	66.7% (4)	6

10. こうした投資のアプローチについて社会責任投資（SRI）などと呼ばれています。 source: http://cloud.watch.impress.co.jp/epw/cda/software/2008/06/09/13128.html

	1	2	3	Response Total
It is called by one name: something similar to "SRI".	100.0% (6)	0.0% (0)	0.0% (0)	6
It is called by different names: "SRI" and other similar names.	16.7% (1)	33.3% (2)	50.0% (3)	6
It is called by different names: something similar to "SRI" and other similar names.	50.0% (3)	50.0% (3)	0.0% (0)	6
It is called by different names: but in particular it is called "SRI".	0.0% (0)	0.0% (0)	100.0% (6)	6
It is called "SRI".	0.0% (0)	83.3% (5)	16.7% (1)	6

11. まず 最初 に 、今 の ところ 非常 に まだまだ 治す の が むずかしい 病気 とか ケガ で 若い 人 や 子ども さん かかる 可能 性 の ある もの 三つ だけ ご 紹介 します 。 source: http://ameblo.jp/regenerative-kyoto/theme29-10006816417.html

	1	2	3	Response Total
Some of them are caused by injuries. Some of them are caused by diseases.	16.7% (1)	16.7% (1)	66.7% (4)	6
Some of them are caused by injuries. Some of them are caused by diseases. Some of them are caused by other similar things.	33.3% (2)	33.3% (2)	33.3% (2)	6
They are caused by injuries and diseases together at the same time.	100.0% (6)	0.0% (0)	0.0% (0)	6
They are caused by injuries, diseases and other similar things at the same time.	100.0% (6)	0.0% (0)	0.0% (0)	6

12. 僕、高校の頃から、理想とか、自分に合うものを探してたんです source: http://www.e-labo.net/meister/2008/02/seo.php

	1	2	3	Response Total
I was looking for something, such as an ideal.	0.0% (0)	33.3% (2)	66.7% (4)	6
I was looking for something similar to an ideal.	66.7% (4)	16.7% (1)	16.7% (1)	6

311

12. 僕、高校の頃から、理想とか、自分に合うものを探してたんです source: http://www.e-labo.net/meister/2008/02/seo.php

	1	2	3	Response Total
I was looking for something, an ideal or other similar stuff.	0.0% (0)	50.0% (3)	50.0% (3)	6
I was looking for an ideal.	50.0% (3)	50.0% (3)	0.0% (0)	6

13. なお 、 I 型 （ インスリン 依存 型 ） とか II 型 （ インスリン 非 依存 型 ） という 表現 は 現存 は 使い ませ ん ので 注意 し て ください 。 source: http://allabout.co.jp/gm/gc/300397/

	1	2	3	Response Total
Expressions such as type I or type II.	0.0% (0)	16.7% (1)	83.3% (5)	6
Expressions such as type I, type II and other similar expressions.	50.0% (3)	50.0% (3)	0.0% (0)	6

14. ネット上でも、「最近顔が丸くなった」「薬とかの副作用じゃないかと心配」「アルコールの飲みすぎでむくんだ？」と様々な憶測を呼んでいる (gossip about a famous woman who got fat recently) source: http://www.j-cast.com/2008/04/28019559.html

	1	2	3	Response Total
People are worried that it is a side effect of medicines and other similar stuff. She probably took several medicines/drugs.	16.7% (1)	50.0% (3)	33.3% (2)	6
People are worried that it is a side effect of a medicine/drug. People want to be less direct about the word "薬".	16.7% (1)	33.3% (2)	50.0% (3)	6

15. 地元 消防 団員 と 町 職員 計 ７０人 が 徒歩 と 車 で 入り 、 浸水 した 家屋 や 道路 に 散乱 した 流木 など を 片づけ たり 、 食料 の 配給 をしたりした 。 source: http://blog.goo.ne.jp/jun-propela/e/c1a900c4aac7f13df8f57648e4b7304f

	1	2	3	Response Total
After entering into the town, firefighters and the town officials put in order the driftwood, distributed food and did other similar activities.	0.0% (0)	33.3% (2)	66.7% (4)	6
After entering into the town, some firefighters and some town officials put in order the driftwood, while some other firefighters and some other town officials distributed food.	33.3% (2)	16.7% (1)	50.0% (3)	6
After entering into the town, some firefighters and some town officials put in order the driftwood, while some other firefighters and some other town officials distributed food, while other firefighters and town officials did other similar things.	16.7% (1)	16.7% (1)	66.7% (4)	6

312

16. また、ゆったりとくつろげる音楽を聴いたり、自分が心地良いと感じる香りを嗅ぐのもおすすめ。さらに肌寒い季節には、ホットミルクなどの温かい飲み物を飲むと、体が温まり、寝つきが良くなります。 **source: http://allabout.co.jp/1/225879/1/product/225879_08.htm**

	1	2	3	Response Total
You can do all these actions or just one of them: smelling scents or listening to music.	0.0% (0)	33.3% (2)	66.7% (4)	6
You can do all these actions or just some of them: smelling scents, listening to music or doing similar things.	33.3% (2)	16.7% (1)	50.0% (3)	6
You can all do these actions: smelling scents and also listening to music.	16.7% (1)	66.7% (4)	16.7% (1)	6
You can all do these actions: smelling scents and also listening to music and other similar activities as well.	50.0% (3)	16.7% (1)	33.3% (2)	6
You have two possibilities: smelling scents or listening to music.	50.0% (3)	50.0% (3)	0.0% (0)	6
You have one possibility: smelling scents and also listening to music.	100.0% (6)	0.0% (0)	0.0% (0)	6

17. これにより、新規ユーザを追加したり同僚の連絡先を調べたりする際に、当て推量をしなくてもよくなった。 **source: http://sourceforge.jp/magazine/07/08/21/0134229**

	1	2	3	Response Total
In each of these cases: when you check your colleagues' addresses or when you add a new user.	0.0% (0)	50.0% (3)	50.0% (3)	6
In each of these cases: when you check your colleagues' addresses or when you add a new user or some other similar actions.	0.0% (0)	16.7% (1)	83.3% (5)	6
In these two cases: when you check your colleagues' addresses or when you add a new user.	50.0% (3)	50.0% (3)	0.0% (0)	6
In this one case: when you check your colleagues' addresses and also add a new user.	100.0% (6)	0.0% (0)	0.0% (0)	6
In this one case: when you check your colleagues' addresses and also add a new user and do other similar activities as well.	100.0% (6)	0.0% (0)	0.0% (0)	6

18. しかし この よう な リーダーシップ は 、欲求 不満 を 満たす 手段 を 失っ たり 、あるいは エスカレート し すぎ た とき に 、崩壊 へ の 道 を たどり ます 。 **source: http://allabout.co.jp/gm/gc/294079/**

	1	2	3	Response Total
In these two cases: when it loses the means to satisfy the frustration or when it escalates too much.	0.0% (0)	0.0% (0)	100.0% (6)	6
In each of these cases: when it loses the means to satisfy the frustration or when it escalates too much.	50.0% (3)	33.3% (2)	16.7% (1)	6
In each of these cases: when it loses the means to satisfy the frustration, when it escalates too much or when something similar happens.	33.3% (2)	0.0% (0)	66.7% (4)	6
In this one case: when it loses the means to satisfy the frustration and also it escalates too much.	100.0% (6)	0.0% (0)	0.0% (0)	6

18. しかし この よう な リーダーシップ は 、 欲求 不満 を 満たす 手段 を 失っ たり 、 あるいは エスカレート し すぎ た とき に 、 崩壊 へ の 道 を たどり ます 。 **source: http://allabout.co.jp/gm/gc/294079/**

	1	2	3	Response Total
In this one case: when it loses the means to satisfy the frustration, it escalates too much and other something similar happens as well.	100.0% (6)	0.0% (0)	0.0% (0)	6

19. あなた の コンピュータ に は 空き 領域 (緑 の 領域) は 十分 に あり ます か ？ 。 そう で ない なら 、 スワップ サイズ を 大きく し たり 、 物理 メモリ を 増やす こと で 対処 でき ます 。

	1	2	3	Response Total
There are two possible solutions: you can do both or just one of them. You can increase the swap size or you can increase the physical memory.	0.0% (0)	0.0% (0)	100.0% (6)	6
There are some solutions: you can do all of them or just some of them. You can increase the swap size or you can increase the physical memory, or you can do other similar things.	16.7% (1)	33.3% (2)	50.0% (3)	6
There is one solution: to increase the swap size and increase also the physical memory.	100.0% (6)	0.0% (0)	0.0% (0)	6
There is one solution: you can increase the swap size, increase the physical memory and do other similar things as well.	100.0% (6)	0.0% (0)	0.0% (0)	6

20. 写真 を 精査 すれ ば 多少 の ノイズ 感 や 細部 の 潰れ が ある もの の 「 キレイ だ から オッケー !!」 と 思え たり し て 。 **source: http://k-tai.impress.co.jp/cda/article/stapa/36839.html**

	1	2	3	Response Total
You might think something similar to "キレイ だ から オッケー !!".	50.0% (3)	16.7% (1)	33.3% (2)	6
You might think something similar to "キレイ だ から オッケー !!" and other similar thoughts.	83.3% (5)	16.7% (1)	0.0% (0)	6
You might think "キレイ だ から オッケー !!" and other similar thoughts.	83.3% (5)	16.7% (1)	0.0% (0)	6
You might think "キレイ だ から オッケー !!".	0.0% (0)	0.0% (0)	100.0% (6)	6