

Language Encoding. Categories, Prototypes, Fuzziness & Alia. A Cross-linguistic Survey

言語エンコード

～通言語的調査によるカテゴリー、プロトタイプ、ファジー性、他をめぐって～

Alda NANNINI

アルダ・ナンニーニ

本論文では、より広い研究テーマ「世界の範疇としての言語」におけるエンコード（言語化）を論じている。「範疇」の役割と交差言語的視点から意味論的・語彙的な範囲を追求しながら、異なる言語の比較において「ファジーな認知エリア（範囲）」がどのようなプロセスで言語的形態を得るかを問題にしている。すでに知られているメカニズムを紹介するとともに、カテゴリーゼーションとエンコードに関する議論の中で限られた役割を果たしていた *deictic pivot* に基づくエンコードの役割を紹介する (Nannini 2018)。誤解なきよう、この研究の解釈について強くお断りしておきたいことがある。この分野の研究は、あらゆる形の差別や、言語に関して誤って導かれた「性質判断」を拒否するということである。この研究の最終目的は、言語形態のベースになる「カテゴリーゼーションと言語化プロセスの多様性と豊かさ」を強調することである。そのために形式的アプローチを超えて、社会的・文化的・語用論的アプローチも分析手段として導入する。

キーワード：linguistic encoding, lexicalization, translatability, 分類と範疇、語彙化

0. Preliminary remarks

Linguistics is not a monolithic discipline. Linguistically relevant sounds are studied by *phonetics*; the shape of lexical items by *morphology*; *syntax* investigates their connections in sentences; *semantics* focuses on meanings and their relationships, and the *text* is the basic unit in *textual linguistics*. *Historical linguistics* deals with variation in time and *sociolinguistics* with variation in society, in its *diatopic* (i.e., geographic) and *diastratic* (i.e., social distribution) extent; variation in context is the object of *pragmatics*. To make our investigation approachable both to specialists and non-specialists, criteria from these disciplines will be used here when relevant. This study is part of a wider research pursued over the years and still in progress, whose aim is to highlight patterns of cross-linguistic encoding through the interpretation of divergences in cross-linguistic categorization.

Unrequited judgement about the quality of a natural language is consequently strongly refuted, and any form of discrimination is firmly rejected.

Regrettably, scarce familiarity with the subject happens to raise such issues, and misinterpretations

often lead to mistaken assumptions, such as considering language *x* “better” than language *y* for features visible in the former but not in the latter. Not only are “quality judgements” biased, but they also find no space in linguistics (see, for ex., the exemplary Wierzbicka 1999). They only show poor knowledge of how languages work. The focus of this work is the vastness of solutions cognitive resources provide when some order in the world is required. Taking advantage of it in order to judge languages in terms of “quality” amounts to an unacceptable discriminatory operation: as a matter of fact, research shows that, for linguistic expression, humans rely on instruments which are as abundant as they are diverse. The first step shall be to address such an a-scientific (or better, anti-scientific) attitude by introducing a few examples.

To hear that language *x* is “more or less precise”, “more or less lexically rich” (usually expressed as: “it has more words”, “it is more refined”, or “more or less... (add any qualifying adjective here) than language *y*” is not uncommon. The positive pole of the *continuum* usually coincides with the speaker’s language, typically perceived as “more *x*” than the

other. For ex., some Japanese, relying on a patchy knowledge of a foreign language (generally English), believe that foreigners are barred from acquiring a “real” knowledge of the Japanese language, which, in turn, bars them from accessing cultural features, especially those perceived as deeply rooted in Japanese identity. Such an assumption, though, entails that speakers of *any* language, no exception made for the Japanese, could never break free from their own linguistic barriers and understand a foreign language and its culture. No academic proof is necessary to see the fallacy of this argument: foreigners and Japanese can and do achieve mutual understanding and are perfectly able to acquire the “equipment” necessary for a smooth comprehension. This awareness allows not only effective cross-cultural communication, but also successful foreign language (FL) education. Sadly, this bias is not rare: the author has talked to Germans who thought that Beethoven’s music cannot be performed properly by a non-German (although concessions were made for Europeans); some Italians, whose culture is best known to the author, have similar objections, but others are biased by an opposite and a-critical cultural xenophilia (often an ill-informed: “A is better” *because* it is *not* Italian). It can be speculated that this originates from the perceived marginality of contemporary Italian culture when compared to the glory of the past, and it is unfortunate that this perceived irrelevance may be exploited to fuel biases such as the “us (*good*) vs. them (*bad*)”. Examples could be multiplied *ad libitum* and certainly readers have had their own share of similar experiences.

Defining relevant features in cross-linguistic and cross-cultural research is not easy, because speakers are often unaware that “speaking about language” relates to factors deeply embedded in their cultural identity. It may raise extremely sensitive issues, and cursory and superficial observation must be avoided. As noted above, our aim is to highlight *the abundance of resources* humans recur to as *categorizing subjects*. This richness argues strongly against biases of any kind, and especially *against the assumption that encoding divergences can be evaluated or equated to qualitative criteria*.

A frequent assumption is that “if language *x* categorizes a *phenomenon/object/aspect of the world* and language *y* does not, the speaker of *y* is not equipped to understand what *x* expresses”. This amounts to say that speakers of different languages cannot communicate because each holds a representation of the world totally off chart for the other. Linguistics and semiotics have shown that these assumptions are not instantiated in the communicative process, which requires an approach more adherent to reality. To make this feasible, as a possible procedure, a few strategies will be discussed in the following, keeping in mind that others may also be available.

Language and *category*, variously combined in several research fields, may sometimes sound as no more than a fancy choice of words; in different frames, though, their meaning may diverge substantially and drawing a “terminological map” may be useful. Although *categorization and language; language and categorization* (order is not irrelevant), *language* or *linguistic categories*, etc. are employed in various theoretical frames, the first term we choose to introduce is *fuzziness*. *Fuzziness* (It. usually *vaghezza*, which is rather *vagueness* ⁽¹⁾), may evoke *ambiguity* and imply a confusing and unhelpful nuisance, but language research regards it as the main instrument for expressing linguistic potential (Machetti 2011). It enables effectiveness by conferring flexibility and avoids excessive and non-sensical proliferation.

Not only lexical units, but also semantic/syntactic interactions recur to *fuzziness*. Larger speech units can (and usually *must*) be interpreted in their reciprocal interaction and context of utterance. See for ex. the Engl. *neck*-compounds: *bottleneck* (both in literal and figurative sense), *turtleneck*, etc.: neither of them is a *neck*, but they both share “something” with it. The first hints to a *metaphorical* process (“*x* is what the *neck* is to the body”) and the second to a *metonymical* one (“*x* is contiguous to the *neck*”). Context-molded interpretations can be exemplified by *You are a genius!*: usually interpreted as a compliment, certain circumstances (*contexts*) may charge it as a sarcastic insult. The same goes for

relevant social features (e.g., the *modality* property of *may* and *would* can be exploited to convey *courtesy*); environmental features; tone of the interaction and many other factors which play a role in interpretation. *Fuzziness* is shaped by *context* and the participants' socio-pragmatic *competence* is relied on to find the *adequate* interpretation: *fuzziness* plays a relevant role in *understanding*.

Univocally defined elements are certainly useful for computing (but this is not clear-cut either: *fuzzy logic* is not an *oxymoron*, it is a field of *logic*), but only *fuzziness* allows *semantic* (such as *metaphor* or *metonymy*) and *pragmatic* triggers (*irony*, *courtesy*, etc.) to access linguistic items. This reckoning explains the increasing relevance given to these fields of research in the last decades ⁽²⁾. *Fuzziness* plays a role also in *prototype theory* and *cognitive linguistics*, as a characteristic of humans as *categorizing subjects*. *Prototype theory*'s lasting developments since Rosch (1973) have shown that *categories* do not rely on strictly definitory traits, but rather allow for a great deal of *fuzziness*. Commonly quoted examples are *gradation*, and Labov (1973), which has shown that very differently shaped objects can be called *cup* as opposite to *bowl*, *mug*, *dish* or *vase*, etc. Similarly, the category of BIRD allows the coexistence of *central* instances (the *robin* being "the most bird" of all) and *peripheral* ones (the *penguin*, "the least bird" of all). Recent theories (Taylor 2003[1989], Diodato 2015), such as *situated conceptualization* and *conceptual blending* consider several factors (e.g., the already mentioned *metaphor*; *frames*, *scripts*, etc.) and point out the *dynamicity of the process* of *categorization* and *classification*, and ultimately of semantic and linguistic encoding.

1. What do language and category mean?

Language and *category* may refer to different items in different fields. Morphology and syntax relate *category* to distribution (e.g., connections / restrictions of use such as *word order*; *gender and number agreement*, *active/medial/passive voice*) and express semantic, pragmatic and textual contents. Semantic *categories* deal with relations among meanings and their legitimate interpretations (e.g.,

relationships and referential context vary for *verbs* and *nouns*: The *airplane flies* and *The flight of the airplane*). Semantic roles, cross-linguistically, may select different morphological-syntactic interfaces: in Engl. *I go/I like*, both the *agent* and the *experiencer* are in the nominative, It. *vado/mi piace* (along with Latin, the Romance languages and German) encodes only the former as *nominative* and the latter as *dative*. In Japanese, two encodings are possible: the *topic* (>*experiencer*, Jap. は) - *comment* (が) + 好き *suki* structure (common in *topic prominent* languages), retrievable also in 嫌い *kirai*, ("do not like"), 怖い *kowai* ("scare, be scared"), etc., and the 気に入る *ki ni iru* construction. *Ki* lacks further determination (*whose ki it is* is not explicit) and *iru* 入る is an alternative reading for 入る *hairu* ("enter"). It entails that "something enters the animate participant's *ki*". *Ki*, a polysemic word existing in several Eastern languages, represents in Japanese an ancient Chinese loanword the meaning of which is *vital force*, *spirit*, *psyche*. These translations, though, rely on different cultural models: Western Languages require a scientific context for *psyche* (an ancient Greek loanword meaning *soul*); *spirit* may evoke religion or *vital force*, etc.: here again *vagueness* (rather than *fuzziness*) is playing its role. The Japanese encoding entails "x entering one's *intimate sphere*" (whatever *ki* may indicate)". The *experiencer* (the animate being endowed with *ki*) *perceives* x as *entering their ki*. Native speakers indicate that the *experiencer* can be encoded either as an *indirect object* (*dative*) or as a *topic*. Which cognitive "route" enables this structure to relate to *like*? Our hypothesis is that the animate being *perceives* "liking" as "x's entrance" in their *ki* without their intervention: if compared to the opposite *ki ni iranai*, "x does not enter (one's) *ki*" : i.e., if something is not liked, "it does not penetrate one's *ki*" ⁽³⁾. These divergences seem to provide an insight of the activated cognitive areas (see Nannini 2010, 2018 for an analysis).

Distribution, it is well known, may also entail semantic restrictions. For ex., *eat* or *sleep* require animate subjects, although inanimate subjects are admitted if pragmatic criteria of *relevance* activate a *metaphoric* process, as in (1) and (2):

- (1) And the thing is these *things have been sleeping* in the attic for 16 years (…)
 (<https://forums.moneysavingexpert.com/discussion/3048810/getting-rid-of-stuff>)
- (2) That is where *the rust has eaten* the sheet metal away.
 (<https://www.motorbiscuit.com/rusty-monday-vw-microbus-blues/>)

2. Are categorization and classification synonyms?

The very terms *categorization* and *classification*, often used as synonyms, are *fuzzy*. The Japanese rendition offers an interesting picture: frequently translated as 分類 *bunrui*, it can indicate greatly different meanings: it may refer to *grouping* and to its opposite *separation*, *distinction*. This apparent contradiction possibly frames two different perspectives, one “internal” and the other “external”. *Grouping* “gathers objects with common features” (which share “internal” properties) resulting in *distinction and separation* of groups of objects whose properties differ from one another (“external”): if so, this *fuzziness* may be ascribed to a *metonymy*. Lexemes which can *contextually* express opposite meanings are not unknown: they are considered instances of *polysemy* and are commonly defined *addad* (or *enantiosemic*, or *auto-antonym*). See, for ex., Engl. *fast* in *hold fast* “firmly” and *move fast* “quickly”; It. *tirare* in *tirare la corda* “pull” or *tirare la palla* “throw”, Jap. いい *ii* or 大丈夫 *daijōbu* which can *contextually* express both “acceptance” and “refusal”: これでいい *Kore de ii* “This will do” (“acceptance”) vs. 砂糖はいい *Satō wa ii*, “Sugar is ok = No sugar” (“refusal”). The particles *de*, *wa* or *ga* must provide contextual references *because* the lexemes themselves (*ii*, *daijōbu*) are *fuzzy*. Being highly standardized, they are not perceived as ambiguous by natives, while foreigners, at first, may find them confusing. *Bunrui* can also refer to both the *criteria* and the *results* of the process (Engl. *ordination/categorization/classification* vs. *order/category/class*).

A different route is shown by Jap. 範疇 *hanchū*, usually employed to indicate a philosophical *category*: understandably, Japanese dictionaries mention

Aristotle’s and Kant’s definitions in this entry. It does not seem to include the process and result of *individuation*: in other words, *hanchū* entails *category*, but not *categorization*, making it a better rendition for a philosophical term, univocal and strictly defined. *Library and Information Science* offers a useful distinction of *classification* and *categorization*, which explains why it can be misleading to see them as synonyms in all contexts. Jacob (2004: 522) states as follows:

Categorization is the process of dividing the world into groups of entities whose members are in some way similar to each other. *Classification* In LIS, the term “classification” is used to refer to three distinct but related concepts: a system of classes, ordered according to a predetermined set of principles and used to organize a set of entities; a group or class in a classification system; and the process of assigning entities to classes in a classification system.

Terminological complexity requires semantically univocal (*non polysemic*) terms. *Categorize* and *classify* may share certain synonymic contexts, but they convey distinct meanings because they refer to different operations. As seen above (for ex. Jacob 2004 and Diodato 2015), *categorizing* is a biological faculty, but it is not exclusive of humans: all living beings take advantage of it in order to discern *relevant features* (It. *tratti pertinenti*, Jap. 関連特徴 *kanren tokuchō*). As Diodato explains (2015: 4), it represents a significant *prerequisite* for language without being, strictly speaking, a *linguistic ability*. As an *innate cognitive* faculty embedded in the biological structure of a species, it enables perceptive and cognitive processes which account for very frequent (tendentially universal) categories, such as human physical dimensions (e.g., *up/down* or *in front/behind*). For the *embodied mind theory* (Lakoff-Johnson 1980, Johnson 1987a, 1987b), it plays a role in *metaphoric* processes (e.g., *up* is *good*, *down* is *bad*; *in front* is *the future* and *behind* is *the past*, etc.). Diodato specifies further that the *categorizing ability* does not simply “register” differences or similarities within the experiential *continuum*, but it also appears to *impact perception* itself, and languages show

that this process is not rare at all. ⁽⁴⁾ Among the meanings of *classification*, Jacob includes “a process of assigning entities to classes in a *classification system*”, which makes keeping *class* and *category* apart even more compelling, due to their different implications and link to different abilities.

We can summarize as follows: cognitive faculties make humans able to *categorize* (“individuate similarities”) the world and in so doing to organize it; as a result, *categorization* provides the ability of distinguishing *classes* (by virtue of categories) to which elements of the world are assigned.

3. Categories: Aristotle and Kant

Recalling the basics of the concept of *category*, we shall begin from the Greek philosopher Aristotle (IV sec. B.C.E.), who stated that an object belongs to a *category* if it *shares* the same features (*properties*) with all the members of that category. In other words, either an object belongs to a category, or it does not. If the properties of *dog* are *animal*, *mammal*, *quadruped*, *canid*, all dogs must share these properties, if an entity has these properties, it is a dog and if it does not, it is not a dog. As straightforward as it sounds, though, this line of thought may encounter some difficulties.



Fig. 1 <https://www.esupply.co.jp/ItemPage/EEX-CH34>

For example, in English (and in Italian) a *chair* is an object used to sit, with legs, a horizontal board attached to another board at about 90°. If so, does Fig. 1 represent a *chair*? Aristotelic *categories* are commonly assumed to represent a rigid structure of *properties* as *necessary and sufficient conditions*. This is certainly fundamental in the so-called “classic theory of categorization”, in spite of its counterintuitive results, as in the example of the

chair. To be fair, though, this statement oversimplifies Aristotle’s conceptual system, which was not as static as it seems. It also contemplated the concepts of *potence* (δύναμις *dýnamis*) and *act* (ἐνέργεια *enérgeia*), the former indicating *potentiality* of what “*can be*” but not yet “*is*”; the latter *the stage of completion and actualization* through *becoming*. This dynamic, still, does not explain why a *chair* can still be a *chair* even without legs, but Aristotle also introduces the systematic contraposition of *substance* (οὐσία *ousia*) and *accident* (συμβεβηκός *sumbebēkōs*), and accounts for certain divergences as *fundamental* vs. *occasional* properties. It can be reasonably assumed that although it was not his intention, these concepts considered together hint to *categorization* as a *dynamic process*. Aristotle’s representation of *categories* is fundamental for the entire history of philosophy as a system in which *categories* are related to *the properties of the object which is being categorized* ⁽⁵⁾.

Aristotle’s perspective was reversed by Immanuel Kant (1724-1804), who claimed that human beings can obtain knowledge only through *experience*: the “thing in itself” (Germ. *Das Ding an sich*) is not accessible (“thinkable”) without experience, the “thing” being any object in the world. No object of knowledge can exist independently from our perception: in other words, perception is the prerequisite for knowledge. In Kant’s hypothesis, perception activates elements (*categories*) existing in our *intellect* and this process enables us to organize the experiential data. It starts with the *intuitions* (phenomena, *the perceived*) which proceed from the *senses* and provide the experienced data. The *perceived* activates *categories*, thought by Kant as *pure concepts* (the “manner our intellect functions”) which remain *empty* if not activated by the phenomena provided by experience. If activated, our intellect proceeds *to unify* the multiplicity of empirical data through such *categories* and this process enables us *to judge* (i.e., “to think and speak about”) an object by making it “thinkable” to us ⁽⁶⁾. Denying that an object’s properties are independent from our knowledge, as it was in Aristotle’s system, Kant brings about a fundamental change of perspective: categorization does not rely on the

object but shifts to the *intellect*, where the process of knowing is activated.

4. Reevaluating fuzziness: context, prototypes, embodied mind, frames and scripts. Linguistic encoding.

Understanding diversity, plurality and variability in the way knowledge can be organized enriches our knowledge of the human being and legitimizes our previous statement that the perspective adopted here rejects all kinds of judgment and discrimination. Humans have the biological necessity to bring order in the continuum of reality, because without this operation they would not even survive. They deal with the world through perception, and even if, as in the Kantian approach, “objects in themselves” are unreachable, they have the ability of setting into motion the *categories* their *intellect* is endowed with and reach knowledge. *Categorizing* allows living beings to make choices, from the basic distinction between what is edible and what is not, all the way up to higher cognitive functions.

Just to be sure, let us clarify that this research does not aim at the “Universal Grammar”, as it is conceived in generativist research, in which syntax (“grammar” in that approach) is considered the mechanism, innate and governed by constant laws, which accounts for human language and may be a possible representation of certain aspects of the human linguistic ability, although neurological research seems to be inconclusive (a thorough examination in Cowie 2017). Our study is not concerned with explaining how correct sentences are structured, or which mechanisms provide the ability of producing them. We are contemplating a larger cognitive approach and focusing on how cognitive contents can and are *differently* elaborated and organized. This *variation* may offer access to the repertory of resources human beings rely on for organizing the world, both in its material and psychological extent. Every phenomenon requires an explanation, and different languages offer multiple and equally valid solutions to similar issues.

5. Categorization and encoding. Similarities and divergences. Encyclopedia, Dictionary and related competence

Imagine some people in a room, which is in total disarray and must be tidied up. All of them will perform *the same operation*, but different people will organize the space differently, choosing different *places* and *proximity* among the objects to obtain different *effects*. It is reasonable to assume that they will respect *similar restrictions*: nobody will put a knife, or the bed, in the bathroom: they will be assigned (*classified* as “belonging there”) respectively to the kitchen and the bedroom. Some may place the PC on the left side of a desk, others on the right, but no one will be precluded from understanding that it is a PC and not a flower vase. Some may like the desk closer to a window, as the position best suited to work or study; others may place there a coffee table and a sofa instead: a peaceful spot for relax. Each area will be just perfect for each person’s day to day life. Clearly this example does not cover the complexities of language, but it may show that a different spatial organization does not imply incomprehensible underlying criteria. Similarly, many linguistic features may not overlap completely and may not show a 1:1 ratio among languages and examples of this will be given in the following.

An example for historical and cultural differences is to imagine Christian terminology poured as is into Buddhist terminology or vice versa: clearly the whole conceptual systems would be misrepresented (a discussion of Lat. *peccatum* in Nannini 2015, Fujitani 2000). Still, acknowledging divergences *does not exclude mutual understanding* (Japanese Christians, as well as Buddhist Westerners, exist).

Languages have many strategies to overcome these shortcomings. In this respect, good examples are provided by *material culture*: historical and cultural reasons explain why *knife* and *fork*, in Japanese, are *loanwords*: ナイフ *naifu*, フォーク *fōku*. To exemplify semantic restriction as a common feature of loanwords, consider Engl. *opera* or *tempo* which are music technical terms, while in Italian they

cover a vast semantic area. On the other hand, Jap. 漫画 *manga* and アニメ *anime* are universally known in Western languages.

The strategy employed in Japanese for verbal neologisms is to attach する *suru* (“do”) to a *katakana* word (e.g., クリックする *curikku suru*, スキャンする *sukyan suru*). Loanword-originated Italian verbal neologisms appear as new members of the *-are* conjugation (for ex. *clickare*, *scannerizzare*). Germ. suffix *-ieren* marks verbs borrowed from Romance languages (*studieren*, *rasieren*), while older ones were assumed or produced with the typical *-en* (*kaufen* “buy” > Lat. *caupo*, *-onis* “seller”). It can be seen that loanwords are a powerful strategy to acquire a necessary linguistic item and several types have been individuated. Jap. 鉄道 replicates Germ. *Eisenbahn* and so does It. *ferrovia* (*ferro* “iron”, *via* “road”), which goes against Italian word order rules. Italian linguistic research defines this pattern *structural loanwords* (*prestiti strutturali*): it recurs to existing elements to form a new lexical unit, shapes it on the borrowed one even if, in so doing, it “ignores” the rules. Interestingly, word order compound rules are not broken by It. *grattacielo* (*gratta-* “scrape” *cielo* “sky”), shaped on *skyscraper*. The *semantic loanword* (*prestito semantico*) strategy borrows the *meaning* from a different language and assigns it to an existing element of the lexicon: the etymological meaning of It. *realizzare* is “to make *x* real”, but in modern Italian it also accommodates Engl. *realize* (for which *accorgersi di* also exists).

Abstract concepts are also variously classified and encoded. Non-Italian speakers may not find immediate criteria to adequately tell *amare* from *volere bene* or *amore* from *affetto*. These rely on different categorizing criteria and no Italian speaker considers them synonyms. Synonymic contexts certainly exist, but a solid *encyclopedic competence* of Italian culture and language is required for a correct production and interpretation. In addition, restrictions may appear when “the object of love” is a person, an animate or an inanimate noun. Comparing translations of *amare/voler bene* and Engl. *love/be fond of* shows rules of acceptability (or lack thereof) in *unmarked* contexts (i.e., preferred interpretations). When

acceptability is not ascertained, further contextual features will be required for interpretation.

(3) I love/am fond of my friends. I love/am fond of my parents. I love/am fond of dogs.

(3a) [?]Amo/Voglio bene ai miei amici. Amo/Voglio bene ai miei genitori. Amo i cani/[?]Voglio bene ai cani.

(4) He loves/is fond of money

is acceptable in English, but *voler bene* would be unacceptable in Italian without further explanation:

(4a) Ama il denaro/ *Vuole bene al denaro.

Jap. 愛する *aisuru* or 好き *suki* show distribution and restrictions which diverge from both Italian and English. So do Germ. *lieben*, *mögen*, *gefallen*, *schmecken* and *gern*. Some contexts may also recur to It. *piacere* (commonly translated with Engl. *like*, Jap. *suki*), which has been excluded here to avoid further complexity which goes beyond our present scope. This limited comparison illustrates deep connections between language and cultural values: not the same objects are appropriate for both *amare* and *voler bene* (a non-preferred, or *marked*, interpretation may even convey a negative evaluation) ⁽⁷⁾. Values emerge in society and in its expressions such as art, music, dance, cinema, etc., but only language can express and explain them. The interrelations existing among these elements are defined in Italian research as *encyclopedic competence* (It. *competenza enciclopedica*) and *dictionarial competence*. With the latter, we translate It. *competenza dizionariale*, which lacks an English rendition. Taylor (1989 [2003]: 84ff) defines it *linguistic knowledge* and refers to it as *dictionary* in the discussion, but *linguistic knowledge* can be very generic, so the Italian terminology will be followed here. Originated in semiotic theory, the former defines *a certain language speaker’s knowledge of the world*, and the latter refers to *the correct linguistic knowledge which structures meaning*. Their relevance in cross-cultural communication and FL education in Japan (with an introduction of Eco’s investigation (1975 [1991]: 143ff, 1984 and many others) is discussed in Nannini (2009, 2010).

The lack of a term for a material or an abstract item which is not part of a culture is not at all rare.

The *ki* seen above is an example and so are Jap. *wabisabi* or *sabishii*, which are perceived not only as lexical items, but also as features of the very Japanese identity, cultural traits that, from a native's point of view, foreigners may fail to grasp. Usually voiced by non-linguists, this attitude relies on "visible" discrepancies, although research shows otherwise. Tullio De Mauro dedicated part of his scientific production (e.g., De Mauro 1982, 2002) to the *onnipotenza semantica* (lit. *semantic omnipotence*) or *indeterminatezza semantica* (*semantic indetermination*) related to *vaghezza* (*vagueness*, similar but not necessarily the same of *fuzziness*), which had been already relevant in Wittgenstein's thought (Machetti 2006, 2011). De Mauro's *Dictionary* defines *onnipotenza semantica* as follows: "the property of a semiologic code to include in its contents any other code's contents, acknowledged by A. Tarski and L. Hjelmlev as a specific characteristic of languages"⁽⁸⁾. More recent studies prefer the term *plurifunzionalità* (Engl. *plurifunctionality*) with a somewhat more limited extent.

Above, some "paths" have been outlined which allow mutual understanding (Machetti 2011: 202ff)⁽⁹⁾, and the idea that people can understand each other only if they share the same culture has been challenged as simplistic and not scientifically grounded. As a matter of fact, in Japan foreigners happen to be told: *You seem Japanese!*, *You are more Japanese than a Japanese!* with surprised expressions of appreciation. The most puzzling experience is to be told that *if people didn't see your face, they would believe you are Japanese*: this comment is certainly appreciative, but instead of emphasizing someone's understanding, it rather conveys that someone "is not supposed to be, look, etc." that way. All humans have a language *which can speak about any other language*, and the wealth of its strategies can be observed in encoding, which seems particularly adequate to identify the "paths" referred to above. This perspective makes *untranslatability* and *inexpressibility* less arguable, because any language (*semiotic code*) is an instrument (language as *metalanguage*: i.e., "the use of language to speak about itself") *endowed with the capability of*

explaining any other language.

6. Role of the *Deictic Pivot in Linguistic Encoding.*

A further tool for the interpretation of encoding processes (*paths*) can be added to the dynamics and strategies to overcome the difficulties of *translatability/untranslatability*. A language *x* can speak about any other language and we have discussed that it does not require its elements to totally overlap the ones of language *y* and that no 1:1 relationship is necessary: languages always tap into their potential. Jap. *wabisabi*, Germ. *gemütlich*, It. *amore/affetto*; *amare/voler bene* will be best conveyed through contextual choices, paraphrases and the other strategies seen above, all of which exploit *vagueness*.

Nannini (2018) puts forward a criterion the extent of which can potentially offer a wider insight: the *deictic pivot*. Already known as one factor of linguistic encoding, it plays a role in interpreting personal pronouns, time/place references, verbal tenses, and many other features which have been and still are investigated. But, as Nannini (2018) hypothesizes, its influence may be ever more consequential than previously thought. The *deictic pivot* conveys *subjective vs. objective encoding* and appears to function as a criterion in the selection of *have vs. existential* in *possessive encoding*. This hypothesis goes beyond the patterns introduced by Heine (1997) and draws an outline of the cognitive space of *possession* in various languages: Latin, Ancient Greek, English, Italian and Japanese (a typical *topic prominent language*). Italian and Latin; or Italian and English do not share the same encoding. Heine has shown a number of patterns, such as the so-called *habere* construction (Engl. *have*, Germ. *haben*, It. *avere*, etc.) and the *existential* construction (*possessee* in the *nominative case*, *possessor* in the *dative case* and *existential verb*, or *copula* in its existential meaning): i.e., that *the possessee is / exists to / for the possessor*. This structure exists in Anc. Greek and in Lat. (*mihi sunt duo pueri*, "I have two sons") as well as in Jap. (人) に〜がある・いる (*hito*) *ni* ~ *ga aru/iru*). But

Jap. also employs the $\sim wa \sim ga X$ construction ($X = i$ -adjective / na -adjective + copula, cp. *Zō wa hana ga nagai*, *Watashi wa kami no ke ga kuro da*, etc.), encoding a particular predication of properties, related to the construction defined in Japanese linguistics *unagi-bun*. Copula and copular elements encoding require further research, but *possession* divergences appear to depend on how certain cognitive areas are categorized. In other words, divergent encodings will be expected depending on which features are given priority (i.e., *pragmatic relevance*). Nannini (2018) concludes that the *habere* and the *existential* encoding differ in the relevance given to the *participant which coincides with the deictic pivot of the utterance* (not necessarily the subject)⁽⁹⁾. When it is given prominence the *habere* construction will be selected, but when it is entrusted to the context, *existential/copular* encodings will appear. See, for ex., the following utterances which have formal correspondence but divergent (*unmarked*) interpretation:

(4a-b) There are three brothers; There is a class
/ Ci sono tre fratelli; C'è una lezione.

(4c) Kyōdai ga san nin iru (兄弟が3人いる);
Jugyō ga aru (授業がある)

(4a-b) do not refer to the *deictic pivot* (presumably the speaker, if unmarked), while (4c) is the Japanese “default encoding” for *possession* (*I have three brothers, I have a car*). Extreme prominence is given to the *pivot* by Italian and other Romance languages, in which even *age* is encoded with *habere*, while English, German and Japanese encode it as a property with a *copular* predicate. This partial overlap can be verified in (5a-e).

(5a-b) Sono italiano, Ho ventidue anni. / Je suis
italien, J'ai vingt-deux ans.

(5c-d) I am Italian, I am twenty-two (years old)
/ Ich bin Italiener, Ich bin zweiundzwanzig
Jahre alt.

(5e) Watashi wa Itariajin da / Watashi wa nijūni
sai da (私はイタリア人だ / 私は22歳だ)

Further evidence is provided by the *zero marking* when the *deictic pivot* conveys the default interpretation, as in the encoding of “home”: It. *a casa* (“at home/home”), Engl. *go home*, Germ. *heimgen*

(etymologically as Engl.); *gehen zu Hause*, etc. No mark specifying whose *home* it is is required, and in Jap. “go home/return home” are entrusted to the context so that the simple *kaeru* (“leave”) can imply 家に *ie / uchi ni*.

(6) Vai a casa? (*your house*, the point of view is the 2nd person) (lit. “Do you go home?”)

(7) Vieni a casa? (*my home* is also *your home*: *venire* “come”, deictically implies a movement towards a 1st or 2nd person) (lit. “Do you come home?”) / Vieni a casa *mia*? (lit. “Do you come to *my* home?”)

If *my home* is *not your home*, the possessive *mia* is required. In Engl. *Do you want to come over?* can convey “to my house”, *Do you want to come home?* usually implies that speaker and the listener live in the same house (cp. also *Do you want to go home?*). The two instances of (7) can be synonymic only under specific pragmatic circumstances, such as the speaker’s intention to express closeness with the listener (“as if it were your house too”). For Japanese, cp. also (8) :

(8) Jap. 早く帰ってお風呂に入ろう! *Hayaku kaette o furo ni hairou!* (“Let’s go home quickly and have a bath!”)

The unmarked interpretation of *kaeru* requires nothing else, and seems to encode two metonymical areas: one is *leave* and the other is *go back to where the animate agent is staying*: the context determines which one is relevant. This explains also a different *actionality*: it can be a *punctual* (“leave”) or an *accomplishment* (“go back”) verb. The role of *vagueness* relying on the contexts is confirmed by the lack of ambiguity perceived by native speakers. *Kaeru* can be “one’s living place” but can be extended to a hotel, or even more culturally relevant, “the family house in one’s hometown” (実家 *jikka*), as in:

(9) 正月に帰る *Shogatsu ni kaeru* (“I am going back (*scil.* to my family home) for New Year’s”), etc.

Foreigners are then often asked 夏休みに帰るんですか *Natsuyasumi ni kaeru n desu ka* (“Do you go back (*scil.* to your country) for summer vacation?”). The semantically close 戻る *modoru* (“go back”)

though, shows different restrictions in place and the *deictic pivot* cannot be retrieved from the context.

The relevance of *deictic pivot* has been further framed by defining *subjectivity* as its *starting point*, the *perspective* from which the utterance is produced, in contrast with *objectivity*, which is not related to the *pivot's* perspective. If the former is pertinent, it influences *deixis* and defines the pattern of lexical selection. One further example is the encoding of movement in It. *andare/venire* which is not coextensive with Jap. 行く *iku*/来る *kuru*. The latter reminds partially Germ. *gehen/kommen*, but they do not overlap with Engl. *go/come*. All these verbs convey a movement A > B but a differently prioritized *deictic pivot* produces different categorizing features. The Italian verbs have two deictic centers, i.e., the 1st and the 2nd person and both work as *pivot*: movements toward the place where the *emitter* or the *receiver* are (were, or will be) will be encoded by *venire*, otherwise *andare* will be selected. Another interesting example which may be related to the encoding of the *deictic pivot* is Jap. *zero marking* vs. *-garu*, as in *i*-adj.: 悲しい vs 悲しがる *kanashii/kanashigaru*; *na* adj. 不思議 vs 不思議がる *fushigi/fushigigaru*, 食べたい vs 食べたがる *tabetai/tabetagaru*. The *zero marking* shows that the *deictic pivot* is to be retrieved from the context. One interesting example is the distinctive use of Italian *bello* and *buono*:

(10) un *bel* film / un *buon* film

The *bello* “beautiful” version is not an aesthetic evaluation, it conveys instead the psychological involvement of the speaker, while the *buon* “good” alternative acknowledges the technical accomplishment and quality. This alternative is often surprising for non-Italians:

(11a-b) una *bella* bistecca (“a beautiful steak”) vs. una *buona* bistecca,

(12a-b) un *bel* caffè (“a beautiful coffee”) vs. un *buon* caffè

(10) highlights psychological involvement, but the *bella/bel* alternative of (11) and (12) conveys the speaker’s *anticipation*. Similar instances of *bello* co-occur with the *medium voice* and the so-called *ethic dative* (e.g., *mangiare* > *mangiarsi*), which, as shown

above, consistently encodes the *experiencer* role. For example:

(13) *Ci siamo mangiati* un *bel* panino al prosciutto (“We treated ourselves to a great ham sandwich”).

(14) *Mi sono bevuto* un *bel* caffè forte e mi sono rimesso a lavorare. (“I enjoyed a cup of strong coffee and got back to work”).

Such occurrences, as those discussed in 2., can be cross-linguistically verified, and justify the assumption that *subjectivity/objectivity* (as tentatively defined above) conveyed by the *deictic pivot* encode the relevance of the “involvement with the contents of the utterance” and seems to represent a powerful encoding tool and to provide a remarkable outcome in lexical semantics.

7. Conclusions

Several instances of the relationship between categorization and linguistic encoding have been introduced, and issues related to *untranslatability* (of items taken in isolation) have been addressed in this paper. It has been shown that *contextualization* (which goes beyond the linguistic environment or *co-text*) provides powerful coordinates in outlining conceptual areas and represents a notable tool for *translatability* and interpretation. Equally powerful is *semantic omnipotence* as a means to relate codes with one another. Adequate knowledge of sociocultural features (*encyclopedic competence*) along with several factors intervening in linguistic encoding have also been examined. *Subjectivity* vs *objectivity* related to the pragmatic concept of *deictic pivot* has been pointed out as a trait which linguistic encoding may or may not entrust to *con-text* (and *co-text*) for interpretation and disambiguation and examples have been given on the basis of the analysis introduced in Nannini (2018). Long known linguistic strategies, virtually present in every language, have also been reconsidered. Among those, particularly noteworthy are *loanwords*, both for their motivation (e. g., cultural and/or social necessity) and for their structure (*formal* or *semantic loanwords*). The theoretical frame is inspired by contextual relevance and the *scripts* and *frames theory*. A deeper

knowledge of the human faculty of categorization and of the mechanisms ruling the interaction between categorization and linguistic encoding can offer a significant insight into actual lexical repertoires: this approach is not only meaningful for linguistic and cross-cultural comprehension, but it also offers a meaningful contribution to a more efficient and motivating language teaching and learning.

REFERENCES

- Cowie F. 2017. Innateness and Language. In Zalta ed. https://plato.stanford.edu/entries/innateness_language/ (retr. 2021/7/26)
- De Saussure F. 1996[1970]. *Corso di linguistica generale (Cours de Linguistique Générale, 1922)*, transl. De Mauro, Laterza.
- De Mauro T. 1982. *Minisemantica dei linguaggi non verbali e delle lingue*, Laterza.
- Id. 2002. *Prima lezione sul linguaggio*, Laterza.
- Diodato F. 2015. *Teorie della categorizzazione. Dal modello classico ai prototipi e oltre*, Liguori
- Heine B. 1997. *Possession. Cognitive Sources, Forces, and Grammaticalization*. Cambridge
- Eco U. 1975[1991]. *Trattato di semiotica generale*. Bompiani
- Id. 1984. *Semiotica e filosofia del linguaggio*. Bompiani
- Fujitani M. 2000. *Shinkyoku. Il canto divino. Leggere Dante in Oriente*. Univ. di Trento.
- Jacob E. K. 2004. Classification and Categorization: A Difference that Makes a Difference. In "Library Trends", 52, 3. https://www.researchgate.net/publication/32956263_Classification_and_Categorization_A_Difference_that_Makes_a_Difference (retr. 2021/2/20)
- Labov W. 1973. The boundaries of words and their meanings. In Bailey and Shuy (eds.), *New Ways of Analyzing Variation in English*. Georgetown, Univ. Press, 340-371.
- Johnson M. 1987a. *Women, fire and Dangerous Things. What Categories Reveal about the Mind*. Chicago Univ. Press
- Johnson M. 1987b. *The Body in the Mind. The Bodily Basis of Meaning, Imagination and Reason*. Chicago Univ. Press
- Lakoff G., Johnson M. 1980. *Metaphors We Live by*. Chicago Univ. Press
- Machetti S. 2006. *Uscire dal vago. Analisi linguistica della vaghezza nel linguaggio*. Laterza
- Ead. 2011. La vaghezza linguistica come problema della pragmatica. Questioni teoriche e dati a confronto. "Esercizi filosofici", 6: 195-213. (<https://www2.units.it/eserfilo/art611/machetti611.pdf>, retr. 2021/2/8)
- Nannini A. 2002. Itte Ii Koto, Itte wa Ikenai Koto. What to Say and What not to Say, in "Studies in Communication Sciences", II-1, 51-68.
- Ead. 2009. Noto e non noto linguistico e socioculturale nella didattica dell'italiano a giapponesi. In Ferreri (ed.). *Plurilingualism, Multiculturalism and Language Learning: Comparing Japan and Italy*. Viterbo
- Ead. 2010. Con-versando. Italiano di giapponesi, giapponese di italiani. In Gesuato, Crisafulli (eds.) *Una lingua per amica*. Tokyo, 7-21.
- Ead. 2015. Which Latin for Music Students? in "Kunitachi Ongaku Daigaku Kenkyu Kiyō", 50, 139-150.
- Ead. 2018. Sulla codificazione linguistica del rapporto di possesso in italiano e in giapponese. In "Studi italiani di linguistica teorica e applicata", XLVII, 2, 357-374.
- Rosch E. 1975. Cognitive Representation of Semantic Categories. In "Journal of Experimental Psychology". 104: 192- 233
- Ead. 1978. Principles of Categorization, in Rosch, Lloyd B. (eds.) *Cognition and Categorization* 27-48, Erlbaum, https://commonweb.unifr.ch/artsdean/pub/gestens/f/as/files/4610/9778_083247.pdf, (retr. 2021/1/25)
- Studtmann Paul. 2019. *Aristotle's Categories*, in Zalta (ed.), (retr. 2020/9/12) <https://plato.stanford.edu/archives/fall2018/entries/aristotle-categories/>
- Taylor J. R. 1989[2003]. *Linguistic Categorization* (third ed.), Oxford University Press.
- Thomasson A. (2019) *Categories*, in Zalta (2019) (retr. 2021/2/20) <https://plato.stanford.edu/archives/sum2019/entries/categories/>
- Wierzbicka A. 1999 *Emotion across Languages and Cultures. Diversity and Universals*. Cambridge
- Zalta E. N. 2019 (ed.), *The Stanford Encyclopedia of Philosophy*, The Metaphysics Research Lab Center for the Study of Language and Information, Stanford Univ., Summer 2019 Edition. <https://plato.stanford.edu/index.html>

NOTES

- (1) Italian linguist Tullio De Mauro (e.g., 2002) regards *vagueness* as one of the main features of language and his vision has inspired many scholars (for an introduction, Machetti 2006).
- (2) *Arbitrariness* (De Saussure 1996[1970]: 85ff, 158ff, 412), although challenged in its first formulation (e.g., Taylor 2003[1989]: 6ff), can be considered an instantiation of *fuzziness*. Taylor (1989[2003]: 41-77: 3, 4) also provides a thorough introduction to *prototype theory*. Cognitive and linguistic studies have offered interesting developments of this theory (for one of the first examples, see Labov 1973).
- (3) *Zero marking* is a frequent encoding for contextually retrievable animate participants: Nannini (2018), and below (7). Natives distinguish 入る as follows: *hairu* is a prototypical movement into a closed space by an *animate agent*; *iru* is rather a “direction”, does not entail physical space and allows *inanimate agents*. The *temporal aspect* is also different: the present-perfect 気に入った *ki ni itta* means “I liked it (and still like it)” and 好きだった *suki datta* excludes this interpretation. As for the *lexical aspect* (*Actionality*, *Aktionsart*) *iru* is a verb of *accomplishment* and, as a *telic* verb, accepts resultative (present-perfect) interpretations; *suki da*, instead, is a *state* indicated by the *copula* and as such an *atelic* verb which excludes the present-perfect interpretation: “I used to like it (but I do not anymore)”.
- (4) The Sapir-Whorf hypothesis of *linguistic relativity* will not be discussed here, because priority will be given to recent studies which also take it into consideration.
- (5) Aristotle’s categories are: (1) substance; (2) quantity; (3) quality; (4) relatives; (5) somewhere; (6) sometime; (7) being in a position; (8) having; (9) acting; and (10) being acted upon (Cat.1b25–2a4) (Studtmann 2019).
- (6) Thomasson (2019) explains: “Kant ultimately distinguishes twelve pure concepts of the understanding (A80/B106), divided into four classes of three: *Quantity*: Unity, Plurality, Totality; *Quality*: Reality, Negation, Limitation; *Relation*: Inherence and Subsistence (substance and accident), Causality and Dependence (cause and effect), Community (reciprocity); *Modality*: Possibility, Existence, Necessity”.
- (7) *Pragmatics* examines this kind of *inferred meaning* in *conversational implicatures* and its connection with literal meanings.
- (8) Available online: <https://dizionario.internazionale.it/parola/onnipotenza-semanticita> (retr. 2020/9/4)
- (9) Machetti explains that *vagueness* is not *lack of control*: if users associate anything with anything else a code would collapse and communication would be impossible.