

Distributional restrictions based on word content and their place in dictionaries

Michele Prandi
University of Genoa
michele.prandi@unige.it
<http://www.micheleprandi.net/>

In A. Orlandi, L. Giacomini (Eds.), *Defining Collocation for Lexicographic Purposes*, Peter Lang, Bern: 99 – 121.

In this paper, I shall examine a general question that cuts across the different subjects discussed in the present volume, that is, the distributional constraints based on word content. The label ‘content-based restrictions’ covers a wide array of heterogeneous, although interconnected structures. The main aim of this paper is to turn it into an orderly map.

When one thinks of content-based restrictions, the first notion that comes to mind is collocation. In this paper, however, I do not speak of collocations for two reasons. First of all, although often used to refer to a variety of different data, this cover notion easily leads one to think of a unitary set of structures, flattening relevant differences and discouraging an in-depth differential analysis. Secondly, some current definitions of collocations (see for instance Grossman / Tutin 2002, 2003; Mel’čuk 2003, Mel’čuk / Polguère 2006; Orlandi, in this volume, § 1.1, for an overview) include both phraseology and word formation, which however are located outside the domain of distribution for both phraseological units and compound words have a distribution but not a distributional structure. They occur in constructions but their inner structure is not the outcome of a combinatory process. Like formal syntactic constraints, by contrast, content-based constraints presuppose construction and distributional structure, a datum that leaves no room for phraseological combinations of lexemes on the one hand and word formation on the other.

After examining some different ideas on lexicon to be found within contemporary linguistics and their manifold relations with conceptual contents and syntactic structures (§ 1), I shall suggest some criteria that enable us to identify three layers of content-based distributional restrictions, namely lexical solidarities, consistency criteria and cognitive models (§ 2). In the final section, I shall hypothesise that the three layers form a hierarchy, analyse the relation of each with lexical structures, lexical contents and formal syntax and examine the place of each in dictionaries (§ 3).

1. Conceptual contents, lexicon and syntax

Lexicon is far from being a univocal concept. First, there is not a shared notion of lexicon in linguistics. Furthermore, to speak of distributional restrictions implies

analysing lexemes in use in syntagmatic structures, a methodological stance that rises some specific questions unknown to more traditional analyses based on systematic lexical values out of use. On the one hand, lexemes in use carry with them a nebula of information that goes far beyond the borderline of anything it is reasonable to call lexicon. On the other, when dealing with content-based distributional constraints it is impossible not to call into question the relationship with formal distributional constraints belonging to syntax.

If we consider the scientific literature of the last century, the documented notions of lexicon range between two opposite ends. According to the structural tradition (Trier 1931(1973), 1932(1973); Lyons 1963; 1977, Coseriu 1967; 1968), lexicon is a language-specific structure like phonology or grammar. As Lyons (1963: 37) puts it, “Each language must be thought of as having its own semantic structure, just as it has its own phonological and grammatical structure”. The consequence of such a statement is that the formal organisation of lexical structures is highlighted at the expenses of substantive conceptual contents. At the opposite end, some cognitive linguists share the idea that lexicon contains any kind of conceptual and factual information connected in some way or other to the use of words. This methodological stance questions the borderline between lexical and encyclopaedic information on the one hand and the borderline between the symbolic and the indexical dimension on the other. The former point is underlined by Haiman (1980: 331) when he claims that “dictionaries *are* encyclopaedias”. The latter idea is shared by some cognitive linguists, among others Croft & Cruse (2004) and Fauconnier (1997), who anchor the meaning of sentences in contingent uses, leaving aside the distinction between long-lasting structures and contingent events and therefore the distinction between the symbolic and indexical dimension: “A language expression E does not have a meaning in itself; rather, it has a *meaning potential* (Fauconnier 1992), and it is only within a complete discourse and in context that meaning will actually be produced” (Fauconnier 1997: 37).

The relationship between lexicon and syntax is in turn intricate. In recent years, the study of lexicon has been moved from the paradigmatic dimension to the syntagmatic one, that is, from the correlations between lexemes within lexical fields to the relations between lexemes within the combinatory structure of model sentences. According to Gross (2012: Ch. 2), for instance, the relevant unit for lexical analysis coincides with a structure formed by a predicative term saturated by its arguments, that is, with a model sentence. Against this methodological stance, however, the object of lexicography ends up overlapping with the object of syntax to a certain extent. Therefore, the question of their interaction comes to the foreground. Insofar as lexical meanings have a combinatory dimension, and therefore a distributional structure, what is the exact relation between the syntax of meanings, so to speak, and the syntax of forms? Once again, the developments of the last decades in linguistics suggest two opposite answers.

Within a typical formal paradigm, as documented by the classical version of generative grammar (Chomsky 1957, 1965), syntactic structures are assumed to be independent of any kind of conceptual constraint: “grammar is autonomous and

independent of meaning” (Chomsky 1957: 17) and “*uniquely* determines [...] semantic interpretation” (Chomsky 1966: 5). Syntax is conceived of as formal not only in that it has a static form but also, and above all, because it imposes a form on the organised concepts. At the opposite end, within functional and cognitive approaches, syntax is assumed to be instrumental and iconic or, more specifically, diagrammatic. In this light, syntactic structures have neither an autonomous form nor a shaping power, but simply mirror independent networks of conceptual relations. The active role of syntax is confined to profiling conceptual relations (Langacker 1987), that is, to imposing a perspective on them (Fillmore 1977). According to Dik (1989(1997: 8)), for instance, “Semantics is regarded as instrumental with respect to pragmatics, and syntax as instrumental with respect to semantics. In this view there is no room for something like an ‘autonomous’ syntax”. The iconic view of syntax, in turn, is split into a weak and a strong version. With regard to a weak version, the structure of sentences mirrors the structure of long-lasting conceptual models: “the linguistic form is a diagram of conceptual structure” (Haiman 1985: 2). With regard to a strong version, the structure of sentences is assumed to depict contingent experiential situations. According to Langacker (1991(1992: 35)), for instance, the meaning of a sentence is the “image” of “a particular event known in full detail”: “When we use a particular construction or grammatical morpheme, we thereby select a particular image to structure the conceived situation for communicative purposes” (Langacker 1991(1992: 12)). Once again, the borderline between the symbolic and the indexical dimensions is blurred. In any case, if we assume that syntactic structures mirror the structure of complex concepts, syntax is not an autonomous structural level but is reduced to the relational dimension of lexicon making room for complex conceptual structures and therefore taking the model sentence as its relevant unit. According to Langacker (2000: 18), “lexicon and grammar form a continuum, structures at any point along it being fully and properly described as symbolic in nature”: grammar is a repository of complex structures each of which, like a lexeme, has a meaning, or a family of interconnected meanings. In a similar light, construction grammar suggests that “Constructions themselves carry meaning” (Goldberg 1995: 1). The meaning of a ditransitive construction, for instance, “can be argued to be the sense involving successful transfer of an object from an agent to a recipient, with the referent of the subject agentively causing this transfer”.

2. The content-based restrictions: selection restrictions, lexical solidarities and cognitive models

This whole range of issues concerning the relation between conceptual structures, lexical structures, lexical contents and syntax is developed throughout the history of Twentieth Century linguistics, from Porzig to cognitive linguistics.

The first linguist who isolated syntagmatic lexical structures, and therefore a first type of distributional restrictions grounded on contents, was Porzig (1934), who suggested the concept of lexical solidarity soon after Trier (1931) published his seminal work on lexical fields, which are paradigmatic lexical structures. In German,

for instance, the difference between the verbs *essen* and *fressen* depends on the nature of the subject, which is a human being in the former case and an animal in the latter. A lexical solidarity is an asymmetric kind of structure, which connects a determinant term with a determined one. The determined term is a relational, unsaturated concept – the verb *essen*, for instance – whereas the determinant term is provided by a class of arguments appropriate for its saturation: for instance, the class of human subjects. A lexical solidarity contributes to determine the value (Saussure 1916(1974: 116-117)) of a relational lexeme.

Unlike lexical solidarities, which belong to formal lexical analysis, the discovery of selection restrictions is one outcome of syntactic research, and in particular of the study of the constraints imposed on the distribution of formal classes of words and phrases within the structure of the sentence. When looking for the constraints that govern the distribution of such formal categories as noun, verb, noun phrase, verb phrase, distributional analysis inevitably meets with an independent layer of constraints that depend on conceptual contents. A verb like *pour*, for instance, requires a second argument whose form and conceptual content are equally constrained: in formal terms, it has the form of a noun phrase – the direct object; in conceptual terms, it refers to a concrete liquid substance. According to Harris (1946: 178), formal constraints belong to syntax, whereas conceptual constraints belong to lexicon, which is seen as a “semantic” complement of formal syntax:

there are further limitations of selection among the morphemes so that not all sequences provided by the *formulae* [that is, the formal schemata of syntagmatic structures] actually occur. Individual limitations of selection cannot be described in these *formulae*; at best, the most important among them can be stated in special lists or in the dictionary.

Unlike Harris, Chomsky (1965) considers selection restrictions as a kind of syntactic constraints. This methodological stance transfers into linguistics Carnap’s idea of a logical grammar (Carnap 1932) that includes both formal and conceptual distributional restrictions. Whereas the function of formal constraints is to account for the agrammaticality of such ill-formed combinations as *This and is perfectly*, the function of conceptual constraints is to filter out such inconsistent conceptual combinations as for instance *Colorless green ideas sleep furiously*, *And Winter pours its grief in snow* (Emily Brontë), *They sleep, the mountain peaks* (Alcman¹).

The idea that selection restrictions belong to syntax was immediately rejected by generative semantics (McCawley 1970(1971), Lakoff 1971), which claimed the “semantic”, and therefore lexical, nature of selection restrictions. According to

¹. The first line of Alcman’s famous *Nocturne*, *eúdousin d’oréon koryphai* is quoted in the English translation by M. L. West, *Greek Lyric Poetry*, Oxford World’s Classics, Oxford: Oxford University Press, 1993: 35. As Jakobson (1959(1971)) immediately underlined, Chomsky’s move is in conflict with the idea of a formal grammar that, like Husserl’s ‘pure grammar’, should be independent of any substantive conceptual constraint. If syntactic connections are independent of the connected concepts, and in the first place of conceptual consistency, why should formal syntax filter out inconsistent combinations? As we shall see below, conflictual complex meanings provide both the most powerful argument in favour of formal syntax and a functional justification for it.

McCawley (1970(1971: 218), selection restrictions “are not restrictions on how lexical items may be combined but rather restrictions on how semantic material may be combined”. Within the framework of generative semantics, the adjective *semantic* is used in a broad sense; it does not refer to language-specific lexical structures but roughly means *conceptual*, or *notional* in Jespersen’s (1924: Ch. III) sense. The same attitude was inherited some decades later by cognitive and functional scholars: according to Lakoff (1987: 539), for instance, “the meanings are concepts in a given conceptual system”. The idea that selection restrictions belong to lexicon is now a commonplace sanctioned by handbooks (see for instance Leech 1974; Palmer 1976; Lyons 1977).

On the other hand, within the framework of cognitive linguistics, which neutralises the distinction between lexical contents and both cognitive structures, encyclopaedic information and contingent data, selection restrictions are seen either as restrictions on shared cognitive models (Fillmore 1977) or as beliefs about the world (Haiman 1980: 345):

semantic constraints and beliefs about the world are not to be distinguished. Thus, selection restrictions: the sentence ‘the rock is pregnant’ violates a selection restriction. But the categorization of rocks as inanimate and hence, *a fortiori*, barren, is a belief about the world, and one which is not necessarily shared by everyone.

In addition to that, it should be stressed that functional and cognitive linguists who are acquainted with the structural tradition (see for instance Dik 1989(1997: 91) and Wierzbicka 1980: 87) equate selection restrictions to lexical solidarities. According to Geeraerts (1991: 38), for instance, “syntagmatic semantic relations, known in transformational grammar as selection restrictions [...] had actually been discussed earlier by Porzig (1934)”.

To sum up, linguists use three labels to refer to content-based constraints on the distribution of lexemes: lexical solidarities, selection restrictions and restrictions on cognitive models. According to some linguists, these labels simply constitute three different ways to refer to the same phenomena. The aim of this paper is to argue for the opposite hypothesis: lexical solidarities, selection restrictions and cognitive constraints, in spite of some surface analogies, represent three profoundly distinct categories which are linked by a web of complex and revealing connections. The distinction is relevant for the structure and aims of dictionaries. As we shall see below, in particular, selection restrictions, unlike lexical solidarities, are never stated in current dictionaries, whose aim is to give definitions of isolated words. By contrast, if a dictionary is meant to account for the use of words in model sentences – if it is “generative” in Gross’s sense, selection restrictions become an essential component.

2.1. Selection restrictions and lexical solidarities

Whereas the discovery of selection restrictions is connected to a distributional approach to syntactic structures, Porzig's work lies outside syntax. His idea was not to describe distributional limitations based on word content but to identify a syntagmatic kind of formal and language-specific lexical structure. As pointed out by Coseriu (1967), the function of lexical solidarities is to provide the language-specific formal organisation of lexical fields with a peculiar kind of differential features, which are not inherent but relational. The couple *essen* vs *fressen*, for instance, organizes the conceptual area of eating in German thanks to a differential dimension that is not based on some inherent features of the verbal content but depends on its relation with a specific kind of subject². This premise already suggests that lexical solidarities and selection restrictions are radically different. In order to examine this point in detail, we shall now observe some significant instances.

The lexical paradigm organising the conceptual area of killing in English contains such values as *murder*, *assassinate*, *slaughter*, *exterminate*, *execute*, *slay*, *butcher*, and *massacre*. Each of these different lexemes is subjected to language-specific lexical solidarities, which impose rather arbitrary restrictions on the kinds of being that can hold as patient. *Murder* is restricted to persons;

assassinate adds the restriction that the object must be a person in a position of political importance and that the agent has a political motive for killing. *Slaughter* and *butcher* seem to be terms used primarily for the killing of animals for food [...] *Slay* is applied to humans or higher animals, overlapping somewhat with *slaughter*, but it has an archaic, especially biblical, connotation. *Exterminate* is usually used for intentionally killing in order to get rid of fairly low forms of animal life, e. g. insects, or animals that are considered pests, e. g. rats [...] *Massacre* adds the feature that the object consists of a group of people [...] *Execute* is like *kill*, and adds the qualification that the act is a punishment for a crime and is carried out according to the laws or mores of a social group" (Lehrer 1974: 123-124).

As the examples show, a language is sovereign when imposing specific restrictions on the use of words, but it is so on one condition: all these restrictions are internal to the boundaries of consistency, and presuppose it. English can freely legislate about what kinds of being can be murdered, slaughtered or massacred, but on the preliminary condition that all these beings are mortal, and therefore living beings. It is not the task of English to state what kinds of being can die, and therefore be killed. Whereas the relation between *assassinate* and its direct object, which has to refer to a person of political importance, is a clear instance of lexical solidarity, the constraint

². Lounsbury 1964: 1073-1074) identifies a root meaning correlated with some oppositive dimensions: "We shall regard as a paradigm any set of linguistic forms wherein: (a) the meaning of every form has a feature in common with the meanings of all other forms of the set, and (b) the meaning of every form differs from that of every other form of the set by one or more additional features. The common feature will be said to be the ROOT MEANING of the paradigm. It defines the semantic field which the forms of the paradigm partition. The variable features define the OPPOSITIVE DIMENSIONS of the paradigm". The classical distinction, however, is not appropriate here for the generic concept of eating is not a meaning in German because there does not exist a hypernym such as English *eat*.

that restricts death to living beings is a clear case of selection restriction. Whereas the lexical solidarity is restricted to the English language, selection restrictions belong to a more general system of concepts that is shared by a very large community.

The relation between selection restrictions and lexical solidarities should now be clear: lexical solidarities draw further language-specific restrictions within conceptual areas that are by definition consistent, and therefore within the conceptual borderline drawn by selection restrictions. If this is true, selection restrictions cannot possibly be seen as language-specific lexical structures. Selection restrictions are consistency requirements that belong to a natural ontology, namely to a sort of conceptual constitution shared by a very large number of people that cuts across many different linguistic communities and governs not only linguistic expression and thought but also, and in the first place, everyday behaviour. The reasons that lead one to think that a sentence such as *The moon smiles* (Blake) has an inconsistent meaning are the same that prevent one from asking questions and giving orders to the moon or to a tree (Prandi 2004: Ch. 8).

The differences between lexical solidarities and consistency criteria are confirmed by the observation of conceptual conflicts.

The violation of a lexical solidarity is a lexical mistake. Even when it leads to a conflict, it does not end up in inconsistency. The utterance *John murdered a spider*, for instance, is conflictual but not inconsistent. In spite of the inappropriate lexical choice, the action itself is consistent in that a spider can be killed. The barrier between the spider and the act of murdering is not conceptual, but formal lexical. This is the reason why lexical conflicts can be considered shallow conflicts. The violation of a consistency criterion, on the other hand, gives rise to a process that is essentially and irreversibly inconsistent and does not admit a consistent framing. The utterance describing a smiling moon, for instance, ascribes to the satellite an action that is inconsistent with its essential properties as an inanimate being. What is wrong in the utterance is not the choice of the word but the action itself. The barrier between the moon and the smile is not just lexical: it is conceptual. This is a case of true inconsistency.

As any consistent conceptual relation can by definition be framed in consistent words, lexical mistakes can be removed by substitution. In the presence of a lexical mistake, it is always possible to find a more or less direct path through lexical structures, leading to at least one alternative non-conflicting lexeme on the basis of a common root meaning. Given an example such as *John murdered a spider*, for instance, the verb can be replaced by *kill*, the generic hypernym of the field containing *murder*. On the other hand, if we take into account the utterance *John murdered a calf*, the verb can be replaced either by the generic verb *kill* or by the more specific *slaughter*, which applies a correlative lexical solidarity to the common root meaning 'kill': *John slaughtered a calf*.

As it is a question of conceptual lawfulness, by contrast, inconsistency cannot be removed by lexical substitution. If a poet describes the moon as a smiling creature, no consistent alternative formulation can be attained by lexical substitution. There is no English verb that is at the same time correlative or structurally connected to *smile* and

fit for inanimate beings. What is barred to the moon is not a given lexeme, but the whole conceptual area of human expression, no matter what lexeme is used to describe it.

A last argument in favour of this line of reasoning is offered by translation. Lexical structures are by definition language-specific, and so are lexical mistakes, which implies that translation does not necessarily preserve them. The anomaly of the English utterance *The horse is mewling* is preserved in French – *Le cheval miaule* – and Italian – *Il cavallo miagola* – because the same lexical solidarity is shared by all three languages. The anomaly of the German sentence *Hans frißt*, by contrast, disappears in English, French and Italian. The translation of an inconsistent meaning, for its part, has no effect on the conceptual conflict. In one of his sonnets, the French poet Charles Baudelaire attributes to the moon the inconsistent experience of dreaming: *Ce soir la lune rêve avec plus de paresse*, which is translated into English word by word as *The moon tonight dreams vacantly*³. On the sole condition that the same system of consistency criteria is shared, whatever language the utterance is translated into the inconsistency does not disappear. For the utterance to lose its inconsistency, it is not enough to imagine another lexical structure; one has to imagine another conceptual landscape – a conceptual picture of world which, unlike ours, would allow celestial bodies to dream. For a person grown up within such a strange ontology, on the other hand, this utterance would in any case be taken as consistent, no matter what language is used to express it.

2.2. Selection restrictions and cognitive models

Cognitive models (Holland & Quinn 1987) impose substantive conceptual restrictions on the structure of typical kinds of things and states of affairs. Like lexical solidarities and unlike consistency criteria, cognitive models belong to the territory of consistent concepts, the same that is circumscribed by consistency criteria. Like consistency criteria and unlike lexical solidarities, cognitive models are conceptual structures shared by a very large group including many different linguistic communities.

Cognitive models share an essential property with consistency criteria: neither conceptual system depicts a picture of the world as it actually appears, but rather a picture of some possible worlds. The two pictures, however, are depicted according to different relevance criteria, and therefore each bears a specific relationship to the world of real experience.

The rule of cognitive modelling is typicality. Cognitive models draw a simplified picture of what our shared world would look like if it were inhabited only by typical beings behaving in a typical way. The typical world is a poorer version of ours, containing nothing more and much less. The typical world, for instance, does not contain a single bird unable to fly.

³. Baudelaire, Ch., *The flowers of evil*, transl. by James Mc Gowan, Oxford: Oxford University Press, 2008.

The rule of consistency is substantive possibility, and therefore conceptual lawfulness. Consistency criteria uncover an indefinite set of worlds, including any possible worlds that combine in any imaginable way all kinds of being and properties compatible with conceptual lawfulness. The inverted world of Baroque poetry, for instance, is as admissible as its real counterpart in terms of consistency. Strolling through consistent worlds, one could meet swimming birds, flying feathered fish – *De l’océan de l’air les poissons emplumés* (Chevreau) – and floating clouds – *I wandered lonely as a cloud / That floats on high o’er vales and hills* (Wordsworth) – under a black sun⁴:

Le feu brûle dedans la glace
le soleil est devenu noir (Théophile)

Fire burns within ice
the sun has become black

Though held to be generally true, cognitive models are sensitive to actual experience. They not only admit the possibility of being falsified by experience but also actually entail it, for the question about typicality arises only insofar as experience makes room for non-typical instances. Most instances that falsify the expectations raised by cognitive models are taken as natural: this is the case, for instance, of birds that cannot fly. Some others would be seen as somehow odd: to come across a feathered fish, for instance, would certainly be an amazing experience. All these kinds of being, however, are equally conceivable, because they do not cross the boundaries of consistency: cognitive modelling is internal to consistency. By contrast, there is no room for inconsistent beings or processes in either actual or possible experience. Inconsistent beings and processes such as the dreaming moon or sleeping mountains are conceivable only as complex meanings of significant expressions, that is, as semantic structures of the symbolic order.

3. Content-based constraints, lexicon and syntax

Consistency criteria, lexical solidarities and cognitive models are not different labels for the same structures but radically different structures; furthermore, they are not on a level but form a hierarchy. Selection restrictions hold as general consistency criteria whose function is to delimit the territory of shared consistent concepts. Within the territory of consistent concepts, each language builds up a network of language-specific lexical relations and correlations, among which lexical solidarities find their place. Within the same territory of consistent concepts, general cognitive models, which are shared by a very large group including many linguistic communities, impose substantive constraints on the structure of typical states of affairs to be found in common experience. Against this background, we are now ready to answer the questions asked at the beginning of this paper: what is the place of each kind of constraint within the lexicon or, to put it another way, what idea of lexicon is consistent with each kind? What is the place of formal lexical structures,

⁴. The examples are taken from Genette (1966: § 1).

cognitive models and consistency criteria in a dictionary? What is the relationship between content-based restrictions and formal syntax?

3.1. Content-based constraints and lexicon

Like semantic fields, lexical solidarities belong to the formal, language-specific layer of lexical structures. Although relevant for the description of the inner structure of a given language, the formal idea of lexicon is not up to facing the challenge of syntax, that is, to providing the conceptual contents needed for the construction of complex meanings of complex expressions. In order to perform this task, formal lexical structures have to be filled with substantive conceptual contents, which however are not language-specific but shared by a larger community. Like Janus, a functionally adequate lexicon has two dimensions: it has a linguistic specific form but also a more general conceptual content. The formal lexicon is a system of relations and correlations between values that belongs to the specific structure of a linguistic system. The functional lexicon is a shared repository of consistent and substantive conceptual contents.

The relevant relations and correlations that build up the formal lexical structure of a given language can be identified against the backdrop of any kind of substantive information thanks to the commutation test. The only relevant relations and correlations that are retained as lexical structures are those that are correlated to differences on the expression plane (Hjelmslev 1943(1961); Coseriu 1967; 1968). In Latin, for instance, in the area of old age, the conceptual distinction between humans, things, and a composite class including animals and vegetables is relevant in that it provides the distinction between the meanings of *senex*, *vetus* and *vetulus* with its differential dimension (Duchacek 1965: 58). A point is worth underlining here: although relevant on linguistic grounds, such classes, and in particular the grouping of animals and vegetables, bear no cognitive import. More generally, such lexical classes are devoid of conceptual relevance⁵.

Once formal lexical structures have been isolated, we can assume that the functional side of the lexicon is an organised repository of concepts and conceptual relations. In order to draw a more exact portrait of a functional lexicon, however, we have to face two orders of questions. On the one hand, we have to take control over two centrifugal drifts that challenge its firmness, that is, the drift towards occasional data belonging to the indexical dimension and the drift towards encyclopaedic information. On the other hand, within the core of functional lexicon, we have to find

⁵. The gap between lexical classes and consistent conceptual classes is underlined by metaphorical lexical extensions of the meaning of such relational lexemes as verbs or adjectives. In English, for instance, one can *harbour*, *gratify*, *obey*, *espouse* a desire, which can *fuel* a project. Since they are shared by English speakers, all these uses are consistent: as we have remarked above, lexical structures are tautologically consistent. This however implies that the inclusion of wishes, boats and human beings in the same lexical class bears no consequence on ontological classification.

a proper place for either layer of substantive conceptual structures, that is, for cognitive models and consistency criteria.

One condition a substantive content must comply with in order to be included into functional lexicon is that it is both long lasting and shared by a wide community of speakers. This criterion ideally filters out any kind of contingent datum bound to a contingent speech event and occasionally shared by its contingent actors. Let us suppose that in order to understand a given message one has to realise that red is the colour of Ann's shirt. Many pieces of contingent information of precisely this kind have to be part of the stock of data available to the interpreter when contingent interpretative acts are performed within the borders of an occasional interpretation field (Prandi 2004: Ch. 1). But it is obvious that the piece of occasional information about Ann's shirt cannot possibly find its place within the lexical definition of the word *red*. The long-lasting datum that red is the colour of blood, by contrast, is a relevant cognitive anchor for circumscribing the shared concept. The activation of contingent information is a condition for understanding a contingent message, but not a condition for understanding the long-lasting meaning of a word such as *red*.

The borderline between lexical contents and encyclopaedic information is at one and the same time elusive, as stressed by Haiman (1980: 329), and inescapable: "One does not expect to find in a dictionary a compendium of everything that is known about horses: if one did, the entry for 'horse' alone would be considerably longer than an entire dictionary. But where exactly does one stop? And, more important, *why* does one stop?". Haiman's question presupposes that lexical and encyclopaedic information are similar kinds of data, so that their difference is a pure matter of granularity and any boundary an arbitrary choice. However, the difference between a lexical definition and an encyclopaedic description is not simply a matter of degree within a continuum – "where exactly does one stop?", as Haiman puts it – but of relevance. An encyclopaedic description ideally contains anything people are assumed not to know and would like to know about the objects referred to by words, and takes for granted anything that can be assumed as shared by everybody. An encyclopaedic description of a car, for instance, is not expected to explain at length its socially assumed function, whereas one would be surprised not to find in it a set of technical and historical data, and an accurate description of different types and famous models of cars. This considerable amount of data about cars is likely to enrich one's empirical knowledge. The definition of a concept, by contrast, tries to make explicit the assumptions that are ideally shared by everybody, namely everything everybody is supposed to take for granted and rely upon when using a word. As the nature of these tacit assumptions is not empirical, their explication does not expand our body of positive knowledge. This is the reason why a concept, even if it necessarily contains some positive information about entities, is not simply a selection from encyclopaedic data but belongs to an incommensurable order of magnitude.

Once the core of the functional lexicon is restricted to a repository of long-lasting conceptual structures shared and relied upon by its users, the following step is to

define the place of the two independent orders of long-lasting and shared conceptual structures, that is, cognitive models and consistency criteria.

If we move from the structure of lexical information to the structure of dictionaries, the first point to be stressed is that consistency criteria are never stated in dictionaries. A good dictionary of English, for instance, is expected to state that the verb *murder* denotes “the deliberate and unlawful killing of a person” (Cobuild Collins) but not that the consistent use of the verb requires an animate object. A possible reason could be that consistency criteria are not language-specific. However, this hypothesis is falsified by cognitive models, which, although generally shared, are incorporated into definitions. In fact, the real reason lies at a deeper level. If it is true that both lexical structures and cognitive models belong to the territory of consistency, and that the function of consistency criteria is to circumscribe such a territory from outside, the conclusion is that the place of consistency criteria is not within the definitions of lexical contents but among the presuppositions of the act of defining and more generally of the lexical enterprise. This point is underlined by Black (1952(1954: 32)), who compares the consistency conditions of lexical definitions to the felicity conditions of speech acts. Just as an act of promise fails if it is addressed to a celestial body, the definition of such verbs as *murder*, *assassinate*, *exterminate*, *kick the bucket* or *slaughter* only makes sense against the presupposition – the tautology, in a sense – that death can only be predicated of animate beings. In the same way that the presuppositions of consistent actions are never stated but simply relied upon when acting, the presuppositions of definitions are not stated but simply relied upon when defining. Consistency criteria remain “outside the jurisdiction of the definition” is Black’s conclusion.

Black’s remark certainly holds for the traditional game of defining, which assumes that the object of lexical definitions is provided by consistent conceptual structures. In particular, it holds for both formal lexical structures and consistent conceptual models of things and processes, which are equally assumed to be consistent. However, if the scope of lexical description is moved from the individual lexeme to the structure of the model sentence and the aim of a dictionary is not simply to describe formal and substantive lexical contents but to account for the distribution of lexemes within sentence structures – for uses (*emplois*), in Gross’s terminology – consistency can no longer be taken for granted but is in turn called into question. In the same light, consistency criteria can no longer be kept in the background but have to be focused on as objects of lexical analysis. When distributional analysis crosses the borderline of purely formal syntax to include the syntax of concepts, so to speak, Harris’s idea of lexicon as a conceptual complement of formal syntactic distribution becomes relevant. In formal distributional terms, namely in terms of formal syntactic structures, there is no difference between such a sentence as *John poured wine into Mary’s glass* and *And Winter pours its grief in snow* (Emily Brontë): both are perfectly well formed. The difference is that the former is consistent whereas the latter is not. The former builds up a complex meaning that corresponds to a shared conceptual model independently accessible to consistent thought, whereas the latter builds up a complex meaning that finds no place among consistent concepts. If its

task, among others, is to account for such a difference, a dictionary seen as a systematic description of lexical structures and contents cannot help but contain a systematic analysis of consistency criteria. Gross' model of "generative lexicon", for instance, makes room for them (Gross 2012). In Gross' model, consistency criteria are dealt with in terms of such "hyperclasses" as 'abstract', 'concrete', 'human', 'animate' and 'vegetable' and thus distinguished from distributional restrictions internal to consistency, including both lexical solidarities and cognitive models, dealt with in terms of "object classes". Gross' approach suggests that the difference between consistency criteria and distributional restrictions internal to consistency is a matter of granularity and thus somehow overlooks the essential difference between structures that account for consistency and structures that are by definition consistent. In spite of this, the room for consistency criteria in a dictionary that focuses on the syntax of concepts is out of question.

3.2. Content-based constraints and formal syntax

At this point, we are ready to deal with the last question: the relationship between content-based distributional restrictions and formal syntax.

If it is conceived of as a systematic description of the consistent distribution of lexemes within the structure of model sentences, a dictionary contains a syntax of its own – a syntax of concepts. This being the premise, the relevant question is whether the syntax of concepts, namely the set of relations displayed by consistent conceptual structures and made explicit in a dictionary, is the same as the syntax of forms, namely the structure of model sentences. As mentioned above, this question receives a positive answer within the framework of cognitive linguistics, and in particular of construction grammar: if it is true that they mirror the structure of consistent conceptual models, the formal syntactic structures of model sentences coincide with the relational dimension of the lexicon. Within such a model, there is no room for an autonomous formal syntax, that is, for a distribution of formal classes of expressions independent of the consistent distribution of concepts. Borrowing Husserl's (1901(1970: 511)) words, there is no room for "that *a priori* system of the formal structures which leave open all material specificity of meaning".

In my opinion, the idea of a formal syntax coinciding with the relational dimensions of lexical contents is falsified by an inescapable datum belonging to the common experience of *homo loquens*: syntax is not only an instrumental device in the service of independent and consistent conceptual structures, but also a creative device capable of connecting concepts in unexpected ways.

The complex conceptual structures that circumscribe the distribution of shared concepts are tautologically consistent. The distribution of syntactic classes in formal syntactic structures, by contrast, is not constrained by a requirement of consistency. The same formal syntactic structures that mirror consistent combinations of concepts when used in a purely instrumental way – as for instance in *John poured wine into Mary's glass* – are capable of connecting concepts in such a way as to challenge the shared conceptual lawfulness: *And Winter pours its grief in snow. A fortiori*, the

formal structures of syntax are not constrained by lexical solidarities – *John murdered a spider* – and by cognitive models: *Fire burns within ice / the sun has become black*. As Husserl (1901(1970: 511-512)) points out, inconsistent meanings do not belong to meaninglessness but are the outcome of a successful formal connection of significant parts to form a meaningful whole, or “unified meaning” (“*einheitliche Bedeutung*”). Since they connect concepts in unexpected ways, they both document the autonomy of formal syntactic structures and provide a functional justification for it (Prandi 1987; 2004). Formal autonomy of syntax is the necessary condition for creative connection of concepts. If it is true that lexicon has a syntax, syntax goes far beyond lexicon.

References

- Black 1952(1954) = Black, Max: Definition, presupposition and assertion”. In *The Philosophical Review* 61. Repr. In: Black, Max (ed.): *Problems of Analysis. Philosophical Essays*, London: Routledge & Kegan Paul, 24-45.
- Carnap 1932(1959) = Carnap, Rudolph: *Überwindung der Metaphysik durch logische Analyse der Sprache*. In *Erkenntnis* II, 219-241. Engl. transl.: *The elimination of metaphysics through logical analysis of language*. In Ayer, Alfred (ed.): *Logical Positivism*, New York: Macmillan: 60-81.
- Chomsky 1957 = Chomsky, Noam Avram: *Syntactic Structures*, The Hague – Paris: Mouton.
- Chomsky 1965 = Chomsky, Noam Avram: *Aspects of the Theory of Syntax*, Cambridge / Mass: The MIT Press.
- Chomsky 1966 = Chomsky, Noam Avram: *Topics in the theory of generative grammar*. In Sebeok, Thomas (ed.): *Current Trends in Linguistics. Vol. III: Theoretical Foundations*. The Hague – Paris, Mouton, 1-60.
- Coseriu 1967 = Coseriu, Eugenio: *Lexicalische Solidaritäten*. In *Poetica* 1, 293-303.
- Coseriu 1968 = Coseriu, Eugenio: *Les structures lexématiques*. In *Zeitschrift für Französische Sprache und Literatur*, Beiheft 1, 3-16.
- Croft / Cruse 2004 = Croft, William / Cruse, D. Alan: *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Dik 1989(1997) = Dik, Simon C.: *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht - Providence. 2nd ed.: Berlin / New York: Mouton De Gruyter.
- Duchacek 1965 = Duchacek, Otto: *Sur quelques problèmes de l’antonymie*. In *Cahiers de lexicologie* VI – 1, 55-66.
- Fauconnier 1992 = Fauconnier, Gilles : *Sens potentiel: grammaire et discours*. In de Mulder, Walter / Schuerewegen, Franc / Tasmowski, Liliane (eds.): *Enonciation et Parti pris*. Amsterdam: Rodopi, 159-171.
- Fauconnier 1997 = Fauconnier, Gilles: *Mappings in Thought and Language*. Cambridge: Cambridge University Press.

- Fillmore 1977 = Fillmore, Charles J.: The case for case reopened. In Cole, Peter / Sadock, Jerrold Murray (eds.): *Syntax and Semantics. 8: Grammatical Relations*. New York / San Francisco / London: Academic Press, 59-81.
- Geeraerts 1991 = Geeraerts, Dirk: La grammaire cognitive et l'histoire de la sémantique lexicale. In *Communications* 53, 17-50.
- Goldberg 1995 = Goldberg, Adele E.: *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago / London: The University of Chicago Press.
- Gross 2012 = Gross, Gaston: *Manuel d'analyse linguistique. Approche sémantico-syntaxique du lexique*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Grossman / Tutin 2002 = Grossman, Francis / Tutin, Agnès: Collocations régulières et irrégulières: esquisse de typologie du phénomène collocatif. In: *Revue française de linguistique appliquée, Lexique: problèmes actuels* 7, 1. 2002, 7-25.
- Grossman / Tutin 2003 = Grossman, Francis / Tutin, Agnès (edd.): *Les collocations. Analyse et traitement*. Amsterdam: De Werelt, 2003 (Travaux et recherches en linguistique appliquée)
- Haiman 1980 = Haiman, John: Dictionaries and encyclopaedias. In *Lingua* 50, 329-357.
- Haiman 1985 = Haiman, John: Introduction. In Haiman, John (ed.): *Iconicity in Syntax*. Amsterdam: John Benjamins, 1-7.
- Harris 1946 = Harris, Zellig: From morpheme to utterance. In *Language* 22, 161-183.
- Hjelmslev 1943(1961) = Hjelmslev, Luis: *Omkring sprogteoriens grundlaeggelse*. Copenhagen. Engl. transl.: *Prolegomena to a Theory of Language*. Madison: The University of Wisconsin Press.
- Holland / Quinn 1987 = Holland, Doroty / Quinn, Naomi (eds.): *Cultural Models in Language and Thought*. Cambridge: Cambridge University Press.
- Husserl, Edmund 1901(1970 = Husserl, Edmund: *Logische Untersuchungen*. Band 1: Halle 1900; Band II: Halle 1901. Rev. edition: Band 1 (Prolegomena), Band II, Teil I (Res. I - V), Halle 1913; Band II, Teil II (Res. VI), Halle 1921. Critical edition: *Husserliana*, Vol. XVIII (1975) - XIX, I - II (1984). The Hague: Nijoff. Engl. transl.: *Logical Investigations*. London: Routledge & Kegan Paul.
- Jakobson 1959(1971 = Jakobson, Roman: Boas' view of grammatical meaning. In *The Anthropology of Franz Boas: Essays on the Centennial of his Birth, Memoir LXXX*. Stanford: American Anthropological Association. Repr. in Jakobson, Roman: *Selected Writings*. Vol. II: *Word and Language*. The Hague / Paris: Mouton, 489-496.
- Jespersen 1924 = Jespersen, Otto: *The Philosophy of Grammar*. London: George Allen & Unwin.
- Lakoff 1971 = Lakoff, George: Presuppositions and relative well formedness. In Steinberg, Danny D. / Jakobovits, Leo A. (eds.): *Semantics*. Cambridge: Cambridge University Press, 329-340.
- Langacker 1987 = Langacker, Ronald: *Foundations of Cognitive Grammar. I*. Stanford: Stanford University Press.

- Langacker 1991 = Langacker, Ronald: Foundations of Cognitive Grammar. II. Stanford: Stanford University Press.
- Langacker 1991(1992) = Langacker, Ronald: Concept, Image and Symbol. The Cognitive Basis of Grammar, 2nd ed.. Berlin / New York: Mouton – De Gruyter.
- Langacker 2000 = Langacker, Ronald: Grammar and Conceptualization. Berlin / New York: Mouton – de Gruyter.
- Leech, J. 1974 = Leech, Jeffrey: Semantics. Harmondsworth: Penguin Books.
- Longacre 1985(2006 = Longacre, Robert E.: Sentences as combinations of clauses. In Shopen, Tymoty (ed.): Language typology and syntactic description. Vol. 2: Complex constructions. 2nd ed.. Cambridge: Cambridge University Press, 372-420.
- Lounsbury 1964 = Lounsbury, F. L.: The structural analysis of kinship semantics. In Lunt, Horace Gray (ed.): Proceedings of the Ninth International Congress of Linguists. The Hague: Mouton, 1073-1093.
- Lyons 1963 = Lyons, John: Structural Semantics. Oxford: Blackwell.
- Lyons 1977 = Lyons, John: Semantics. Cambridge: Cambridge University Press.
- McCawley 1970(1971) = McCawley, John D.: Where do noun phrases come from?. In Jakobs, R. / Rosenbaum, P. S. (eds.): Readings in English Transformational Grammar, Waltham / Mass.: Blaisdell. Revised version in Steinberg, Danny D. / Jakobovits, Leo A. (eds.): Semantics. Cambridge: Cambridge University Press, 217-231.
- Mel'čuk 2003 = Mel'čuk, Igor: Collocations dans le dictionnaire. In: Szende, Thomas (ed.): Les écarts culturels dans les Dictionnaires bilingues. Paris: Champion, 2003, 19-64.
- Mel'čuk / Polguère 2006 = Mel'čuk, Igor / Polguère, Alain: Dérivations sémantiques et collocations dans le DiCo/LAF. In: Langue française 150. 2006, 66-83.
- Palmer 1976 = Palmer, Frank R.: Semantics. Cambridge: Cambridge University Press.
- Porzig 1934 = Porzig, Walter: Wesenhafte Bedeutungsbeziehungen. In Beiträge zur deutschen Sprache und Literatur 58, 70-97.
- Prandi 1987 = Prandi, Michele: *Sémantique du contresens*. Paris: Les Editions de Minuit.
- Prandi 2004 = Prandi, Michele: The Building Blocks of Meaning. Ideas for a Philosophical Grammar. Amsterdam / Philadelphia: John Benjamins.
- Saussure 1916(1974) = Saussure, Ferdinand de: Cours de linguistique générale. Paris: Payot. Critical edition by T. de Mauro.1972. Engl. transl.: Course in General Linguistics. London: Fontana & Collins.
- Trier 1931(1973) = Trier, Jost: Der deutsche Wortschatz im Sinnbezirk des Verstandes. Die Geschichte eines sprachlichen Feldes. Part I: Von den Anfängen bis zum Beginn des 13. Jahrhunderts, Heidelberg: Winter. Repr. in Trier 1973, 40-65.
- Trier 1932(1973) = Trier, Jost: Sprachliche Felder. In Zeitschrift für Deutsche Bildung 8. Repr. in Trier 1973, 93-109.
- Trier 1973 = Trier, Jost: Aussätze und Vorträge zur Wortfeldtheorie, ed. by van der Lee A. / Reichmann, O. The Hague / Paris: Mouton.

Wierzbicka 1980 = Wierzbicka, Anna: *Lingua Mentalis*. Sidney / New York: Academic Press.