

A dual ontology across the grammatical / conceptual divide

Isabelle Roy
Nantes Université / CNRS,
UMR 6310 LLING
isabelle.roy@univ-nantes.fr

Bridget Copley
CNRS / Université Paris 8,
UMR 7023 SFL
bridget.copley@cnr.fr

[...] the proposed dividing line is a real one, and [...] it distinguishes between what is grammatically real—structures and formal properties of functional items, and what may be very real, but not grammatically so—properties of substantive vocabulary.

Borer 2005a: 10-11

A map is not the territory it represents[.]
Korzybski 1933: 58

1 Grammar and concepts

A theory of ontology is an answer to one of the oldest questions of humanity, namely *What is there?* This answer can, however, take different forms depending on whether one considers what actually exists in the world—and which is not directly accessible—, what exists in our perception of the world—which is to a large extent shared—, or rather what exists in the ways in which we speak of the world—and which can vary greatly across the different linguistic systems. We do not have much to say about the first dimension, i.e. that of *foundational* ontology. There is a metaphysical fact of the matter but what we think of it is inconsequential for a theory of language. The other two dimensions, however, are important because they are very closely related. Our mental concepts of what is in the world, that is, what exists almost universally in our perception, are very closely related to language. There are two obvious reasons for that: we use language to talk about these concepts, and perhaps even more importantly, these concepts are in themselves the backbone of our grammatical systems. Every language, through its grammatical system, forces the speaker to express particular distinctions that are in fact an answer to the ontological question.

In neo-constructionist approaches to the syntax-semantics interface, of which Hagit Borer has been and continues to be a leading advocate, the ontology question is bound to take a very important place because morphosyntactic decomposition presupposes a prior identification of grammatical primitives (e.g., number, person, tense, and so on) that are based on related conceptual primitives. This is also true of cartographic approaches and other theories of decomposition concerned with interactions between structure and meaning.

In a simple world, grammatical primitives and conceptual primitives would go hand in hand. It is self-evident, however, that distinct linguistic systems may grammaticalize different conceptual primitives. As Roman Jakobson said "... languages differ essentially in what they *must* convey and not in what they *may* convey" (1959:236) [*italics ours*]. So, we have to recognize that different languages may grammaticalize different subsets of the set of conceptual primitives. To take just one example, at the conceptual level, anyone can form an intuition that there is a difference between reporting something for which we have direct evidence or not. Evidentiality is conceptually accessible to anyone: uttering *It's raining in Nantes* because I see the rain through my window or because I heard it from a friend in Nantes does not report exactly the same thing. In the grammar of English this conceptual distinction does not have any grammatical effect as in both cases the same sentence will be used successfully. If I want to express the absence of direct evidence I can do it (*It's raining in Nantes apparently!*) but I don't have to. In other languages, and they represent about 25% of the world languages according to Aikhenveld (2004), the source of information must be grammatically expressed. In many cases the morphosyntax of the verb will be directly affected. For instance, in Turkish, the *-di* and the *-miş* suffixes vary in the expression of direct vs. indirect evidentiality, forcing the speaker to make an ontological choice in the formation of a standard sentence:

- (1) a. Adam elmayı yedi
man apple.ACC eat.DIRECT EVIDENTIAL
'The man ate the apple' → witnessed
-

- b. Adam elmayı yemiş
 man apple.ACC eat.INDIRECT EVIDENTIAL
 ‘The man ate the apple’ → reported or inferred

(examples from Arslan et al. 2015)

To further illustrate variation across linguistic systems, consider the expression of Tense. While many languages express Tense on the verb, languages like Mandarin, for instance, do not: the verbal form remains unchanged and information relative to time (future and past, for instance) are given by satellite expressions. Mandarin speakers of course have the conceptual notions related to past, present or future even though such information is not marked by the morphosyntax but only by satellite lexical items. Conversely, Mandarin has a complex nominal classifier system that comprises a number of relatively transparent classifiers whose semantics is largely based on physical properties of the object (shape, size, entity, and so on). Again, it would be inconceivable to say that speakers of languages without classifiers do not perceive physical properties of objects. The type of cross-linguistic variations presented here are very common and many more examples could be given that involve all sorts of conceptual distinctions: definiteness, proximity, boundedness, causality, stativity, eventivity, property, degree, manner, and many others, that are grammaticalized in some languages and not in others.

2 The grammatical / conceptual divide

The picture sketched above is one of a shared set of basic conceptual primitives, with languages each grammaticalizing a subset of this set. We need at least such a picture, but as Hagit Borer has argued, we need to further elaborate the picture because of the widespread existence of cases where a similar primitive is used at the conceptual and the grammatical levels but they exhibit a kind of mismatch.

Cases of mismatch are also wide-spread. That a category (in the broad sense) exists at the conceptual level and at the grammatical level does not mean it covers the same reality nor that it distinguishes the same classes of elements. Consider, for instance, gender, masculine and feminine. These two categories refer to a conceptual distinction, whereas linguistically, in languages that have a dual masculine/feminine nominal class system, as for instance French, any conceptual distinction (whether chromosomal, hormonal, social, etc.) is largely ignored: *une sentinelle* (fem.) ‘a sentinel’ designates a standing guard, until recently most typically a man guard; *le bébé* (masc.) remains masculine independently of the actual gender of the baby (**une bébé* (fem.) does not exist, and we will say that *Lila (fem.) est un joli bébé (masc.)*); *la tortue* (fem.) ‘turtle’, *la sauterelle* (fem.) ‘grasshopper’, *le moustique* (masc.) ‘mosquito’, *le pou* (masc.) ‘louse’ have only one grammatical gender, while the animal belongs to two.

Another example (Borer 2005a: 115) concerns plurality. Plural, which is taken to mean “more than one” in English, is linguistically marked on nominals by the *-s* suffix (e.g., *birthday* / *birthdays*). As Borer has shown in a very interesting manner, the mapping between the semantics (‘more than one’) and the morphosyntax (suffix *-s*) is imperfect: quantities smaller than one also take the so-called ‘plural’ morpheme, despite the fact that they do not express a plurality, and not even a full unity: e.g., *zero books*/**zero book*, *0,5 books*/**0,5 book*. Borer concludes from these facts that the English *-s* does not realize number but rather a classifier related to countability. In a rather similar fashion, *pluralia tantum* (i.e., nouns that do not have a singular variant, e.g., *rapids*, *blinders*, *scissors*, *glasses*) are not really plurals but denote singular objects; and certain singulars (e.g., *crowd*, *team*, *herd*) designate a plurality / collection of entities.

Across languages, the same concept (or conceptual primitive) may also correspond to different grammaticalization phenomena. For instance, Talmy (2000 e.g.) has shown that the expression of manner of motion events can differ according to whether the language expresses the manner on the verb and the location by the satellite (e.g., *danced into the room*) or reversely, the motion is expressed by the verb and the manner by the satellite (e.g., *entrer dans la salle en dansant* (lit. *enter the room dancing*), in French).

Similarly, but in another domain, in different languages nominals may be classified differently as either mass or count (Borer 2005a), but the classification may not be associated with any conceptual difference. For instance, *furniture*, *spaghetti* and *scaffolding* are mass terms in English, while their equivalents in French (*meuble*, *spaghetti* and *échafaudage* respectively) are count terms. Again, here, the difference between the two languages is located at the grammatical level and not the conceptual one. The terms in English and French are strictly equivalent, except that *furniture*, *spaghetti* and *scaffolding* do not pluralize, while *meubles*, *spaghettis* and *échafaudages* do. Thus, while we say *We're having (gluten-free) spaghetti with Hagit tonight* in English, in French we will use a plural form *On va manger des spaghettis (sans gluten) avec Hagit ce soir*. Such cases are legion.

Finally, to round out our overview of mismatch situations, we should also mention cases of grammatical distinctions that have been neutralized or bleached conceptually. In fact, gender, e.g. in Romance languages, as presented above, is such a case. The conceptual distinction masculine / feminine is mostly bleached, except maybe for certain individuals and animals where grammatical gender is more closely associated with the corresponding conceptual gender (e.g., *un acteur* 'actor' / *une actrice* 'actress'; *un chat* 'male cat' / *une chatte* 'female cat').

To take another example of a conceptual distinction that is no longer associated with the corresponding grammatical distinction, consider the deictic expressions *-ci* and *-là* in French. Historically, the former was related to a proximal location (cf., *here* and *this* in English), and the later to a distal location (cf., *there* and *that*). Nowadays, this distinction has largely disappeared: one can use *je suis là* (lit. *I am there*) to refer to the precise location where they are standing (thus by definition proximal), and *celui-ci là-bas* (lit. 'this.one-here over-there') to designate an object far away.

Considering such mismatch cases that are found across languages, and noting how prevalent they are in grammar, we believe that Hagit Borer is correct to propose a sharp divide between what is grammatical and what is conceptual. The divide between the conceptual and the grammatical domains surfaces in Hagit Borer's exoskeletal approach (2003, 2005a, 2005b, 2013) in two places. First, it surfaces in the division of labor between the lexicon and the computational system: the lexicon is composed of roots, which are associated with an encyclopedic knowledge (in essence, a concept) but are devoid of any formal grammatical information (e.g., number, aspect, but also argument structure, and even category). Second, it also surfaces in the formal division between 'encyclopedic' roots that simply relate to an encyclopedic concept and a functional lexicon, that comprises essentially, grammatical features (e.g., [+pl], [+pst]), as well as grammatical items that, themselves, carry such features (e.g., the determiner <*the*[+Def]>) (Borer 2003:34).

3 A dual ontology

Here we propose to take the idea of the grammatical/conceptual divide to what we think is one of its logical conclusions.

Consider a predicate.¹ A predicate—which is a linguistic object—is not a property—which is something in the world. But while a predicate is not a property, a predicate does stand in some relation or mapping to a property. As linguists, we generally encounter this mapping as the evaluation function (e.g., Heim & Kratzer 1998) represented by double brackets "[[]]" around a predicate, as in (2):

$$(2) \quad \llbracket \text{PREDICATE} \rrbracket(x) = 1 \text{ iff it's true that } \textit{property} \text{ holds of } x$$

If we take seriously Hagit Borer's point (Borer 2005a: 11) that "grammar only cares about its own", however, there is something strange about (2). The strange thing is that x is the argument of a predicate, but also the holder of a property in the world. But in the same way that a predicate is not a property, we should not be forced to say that the argument *is* the individual. That is, a predicate does not (cannot) take an object in the world as its argument.

So, just as a predicate is not the very same object as a property, but merely maps to it; in the same way, the term x is not itself the very same object as the holder of the property, but merely maps to it. Following this train of thought, and contrary to usual practice, the evaluation function should properly also include terms such as x in its domain, not merely predicates. So, instead of writing (2), we should rather write (3), where $\llbracket x \rrbracket = \xi$:

$$(3) \quad \llbracket \text{PREDICATE} \rrbracket(\llbracket x \rrbracket) = 1 \text{ iff it's true that } \textit{property} \text{ holds of } \xi$$

What all this suggests is that for each language, there is not one ontology, but rather two ontologies: a conceptual ontology that we assume, as a null hypothesis, to be common to humans in general; and also, a language-specific grammatical ontology. These stand in relation to each other at the grammatical/conceptual divide (Roy, 2015). Following Copley & Harley (2015) and Copley (2018), we can call this proposal a "dual ontology" proposal. In the conceptual ontology there are entities and properties that can hold of such entities; in the grammatical ontology there are terms that refer to entities as well as terms that refer to properties. The phenomenon of mismatch can then be thought of as a mismatch between the map (the grammatical ontology) and the territory (conceptual ontology), as in Korzybski's (1933) aphorism.

We may suppose that the reason that linguists usually indiscriminately write the same variable to refer to both the argument of a predicate and the entity in the world is that the grammatical/conceptual distinction has not been thought to matter for entities. We do think this point may matter in some cases, however. What mediates between grammar and the conceptual level should be some mapping between the predicate and the property on the one hand; *and* one between the argument term and the object on the other. An appreciation of this point opens the theoretical space to interesting possibilities that would otherwise remain unexplored.

For instance, nominalizations may present an interesting case of what we may call "entity mismatch". 'Simple Event Nominals' (Grimshaw 1990), e.g., *film*, *meeting*, refer to an event at the conceptual level (e.g., *The meeting took place at 12 and lasted for 2 hours*), but behave syntactically like an entity, that is, as type e , not type v (Roy & Soare 2013, 2014). This grammatical entity happens to be mapped to an event in the real world, but it does not exhibit properties of grammatical events. AS-Ns (Borer 2012, 2013; also termed 'Complex Event nominals' in Grimshaw 1990; e.g., *destruction of the city*), by contrast, refer

¹ Or equivalently for our purposes, and more in line with Hagit Borer's work, consider a semantic feature.

to an event both at the conceptual level and also at the grammatical level, and thus can take event-related modifications: compare *the destruction of the city in two hours* vs./ **the meeting for/in two hours*. So, a grammatical entity, which crucially is not a grammatical event, can nevertheless be mapped to a conceptual event.

Another case of entity mismatch is proposed by Copley & Harley (2015), where forces –that is, inputs of energy–in our conceptual representation of the world are mapped to functions from situations to situations. Just as we can be confident that a predicate is not a property, we can be equally confident that a function is not an input of energy. In a dual ontology perspective, of course, there is nothing untoward about either of these ideas. Representing forces as functions from situations allows for a natural way to represent both launching and entrainment causation (Michotte 1946/1963, Shibatani 1973), which ultimately may illuminate and simplify the syntax-semantics interface in the verb phrase (Copley & Harley, 2020 ms.).

While Hagit may not endorse these particular implementations, we want to thank her for her inspirational work on the grammatical/conceptual divide that was instrumental in making a dual ontology perspective thinkable. Any formal linguistic theory that incorporates a means to distinguish conceptual information from grammatical information must be on the right track, and this is most particularly true of Hagit's exo-skeletal research program. Thank you, Hagit!

Isabelle Roy, Bridget Copley

References

- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. OUP Oxford.
- Arslan, S., Bastiaanse, R. and Felser, C. 2015. Looking at the evidence in visual world: eye-movements reveal how bilingual and monolingual Turkish speakers process grammatical evidentiality. *Frontiers in psychology*, 6, p.1387.
- Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: syntactic projections and the lexicon. In Moore J. & Polinsky M. eds, *The Nature of Explanation in Linguistic Theory*, Chicago, CSLI / The University of Chicago Press, pp. 31-67.
- Borer, Hagit. 2005a. *Structuring Sense vol. I: In name only*, Oxford: Oxford University Press.
- Borer, Hagit. 2005b. *Structuring Sense vol. II: The normal course of events*, Oxford: Oxford University Press.
- Borer, Hagit. 2012. In the event of a nominal. in Everaert, M., M. Marelj, M. and T. Sioni (eds.) *The Theta System: Argument Structure at the Interface*. Oxford: Oxford University Press.
- Borer, Hagit. 2013. *Structuring Sense vol. III: Taking form*, Oxford: Oxford University Press.
- Copley, Bridget. 2018. *The grammatical and conceptual ingredients of what happens next* (HDR, Université de Nantes).

- Copley, Bridget and Harley, Heidi. 2015. A force-theoretic framework for event structure. *Linguistics and Philosophy*, 38(2), pp.103-158.
- Copley, Bridget and Harley Heidi. 2020. What would it take to tame the verbal Hydra? CNRS/Paris 8 and University of Arizona (ms.)
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
- Heim, Irene, and Kratzer, Angelika. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
- Jakobson, Roman. 1959. On linguistic aspects of translation. In Reuben A. Brower (ed.), *On translation*, Cambridge, Mass.: Harvard University, Press. 232-239. Reprint New York: Galaxy Books 1966.
- Korzybski, Alfred 1933. *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*. International Non-Aristotelian Library Publishing Company.
- Michotte, Albert 1946/1963. *The perception of causality*. Methuen: London.
- Roy, Isabelle A. 2015. *Fonction combinatoire du langage: Structure et interprétation*. Habilitation à diriger des recherches, Université de Nantes.
- Roy, Isabelle and Soare, Elena. 2013. Event related Nominalizations. In G. Iordachioaia, I. Roy and K. Takamine eds., *Categorization and Category Change*, CSP, 121-152.
- Roy, Isabelle and Soare, Elena. 2014. On the internal event properties of *-er* nominals. *Lingua*, 141, 139–156.
- Shibatani, Masayoshi. 1973. A linguistic study of causative constructions. Ph.D. thesis, University of California Berkeley.
- Talmy, Leonard. 2000. *Towards a Cognitive Semantics*. 2 vols. Cambridge MA: MIT Press.