

A Definitional Metalanguage for Explanatory Combinatorial Lexicography

Lucie Barque

Lattice
Université Paris 7
UFRL, case 7003,
75251 Paris cedex 5, France
lbarque@linguist.jussieu.fr

Alain Polguère

OLST — Dép. de linguistique et traduction
Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal (Qc) H3C 3J7, Canada
alain.polguere@umontreal.ca

How are lexical senses described? One can identify two main families of sense descriptions on the basis of their relation to word meaning: word paraphrases *vs.* word interpretations. The first family can itself be subdivided into three main subfamilies, according to the formal nature of sense descriptions.

1. Word paraphrases

- 1.1. “**Standard**” **lexicographic definitions** are, roughly speaking, linguistic paraphrases of individual word senses. Thanks to their relatively controlled structure (use of a “defining vocabulary”), the LDOCE’s definitions (Longman, 2003) may claim to be one of the most representative example of this approach to sense description.
- 1.2. **Analytical definitions** are more formal descriptions of senses, to be used in the scientific study of language semantics. Such definitions (like the previous ones) find their origin in Aristotle’s theory of concept definition and are used in “theoretical” dictionaries such as (Wierzbicka, 1987) or (Mel’čuk *et al.*, 1984, 1988, 1992, 1999).
- 1.3. **Logical definitions** are, so to speak, translation of word senses into logical formulas, usually based on some form of predicate calculus (Dowty, 1979), or graphical encoding of logical configurations (conceptual graph *à la* Sowa, etc.).

2. Word interpretations

Feature-value matrices are formal descriptions tailored for unification, rather paraphrase processes. They can be based on simple sets of binary features (Pottier, 1974) or be based on richer, logic-based feature systems (Pustejovsky, 1998).

We will present an approach to sense description that is somehow intermediary between analytical definitions and feature-value matrices (1.2 and 2 above). It has been developed in the process of translating the *Explanatory Combinatorial Dictionary* (ECD) definitions into a semantic database of French lexical meanings. The defining language we propose can function as some sort of “pseudo-code” for both reasoning on word senses and formally encoding their description in a machine-readable format. This new approach to sense description is applied to the construction of the *BDéf* database (Altman and Polguère, 2003) and is presently used for two types of research: 1) exploration of how lexical function relations (semantic derivations and collocational links) can be hooked up to semantic components of definitions, and 2) identification and modeling of regular patterns of polysemy.

In order to illustrate our approach in this short abstract, we will look at two senses of the French vocable MAISON (Eng. HOUSE or HOME). There is no entry for MAISON in the published ECDs. However, MAISON is accounted for in the DiCo database of French collocations and semantic derivations (Polguère, 2000; Steinlin *et al.*, 2004). Its entry in this database presently contains five different senses, of which we will consider here only the first two, MAISON_{1.1} and MAISON_{1.2}, illustrated below.

- I. 1 *Elle cherche la **maison** où habite son cousin.*
{She’s looking for the **house** where her cousin lives’}
- 2 *Elle a quitté la **maison** à l’âge de 18 ans.*
{She left **home** at the age of 18’}

As one can see, MAISON_{1.1} possesses a direct counterpart in one of the senses of HOUSE. However, MAISON_{1.2} corresponds, though probably not exactly, to one of the senses of HOME. While MAISON_{1.1} denotes a construction one lives in, MAISON_{1.2} denotes a “space” occupied by a group of people, which is inside the place these people live in. It is worth noting the MAISON_{1.2} is mainly used in the definite singular form, has illustrated in the example above. A sentence like

*Elle a quitté **sa maison** à l’âge de 18 ans.*

would mean that this young woman left her **house**, rather than her **home** at the age of 18. It could hardly be used, for instance, for someone living in an apartment at the time she left —whereas our initial sentence

(with *la maison*) could. As one can see, the difference between these two senses of MAISON is both very clear (a construction *vs.* the inside of a construction) and subtle: no “pure” metonymic relation between the two. In the following table, we present next to each other the BDéf definitions for MAISON_{I.1} and _{I.2}.

MAISON I.1	MAISON I.2, mainly definite singular
<u>Composante centrale</u> (= main component) 1: habitation de X <u>Corps de la définition</u> (= definition body) /*Dimensions*/ 2: *1 grand.relativement /*Structure*/ 3: *1 constitué de 1{niveau}/plusieurs{niveau} /*Matériau*/ 4: *1 fabriqué avec {matériau}.résistant <u>Typage des actants</u> (= actant types) X: individu	<u>Composante centrale</u> 1: lieu_occupé par X <u>Corps de la définition</u> /*Limites*/ 2: *1 être intérieur de habitation de X <u>Typage des actants</u> X: {individu}
<u>English gloss in linear form:</u> Habitation^a of person X , that is relatively big, made up of one or more levels and built with strong materials.	<u>English gloss in linear form:</u> Place used by the group of persons X, that corresponds to the inside of the habitation of X.

- a. The French word HABITATION is much more common than its English counterpart, and there is no problem defining in French *maison de X* on the basis of *habitation de X*.

These definitions call for many comments and explanations. For lack of space, we can only highlight in bold in the above-table the semantic components that implement the quasi metonymic relationship between the two senses of MAISON analyzed here. In our presentation, we will explain how BDéf-like definitions are structured (elementary semantic propositions grouped into clearly identified defining blocks) and demonstrate, using specific examples of the MAISON_{I.1} *vs.* _{I.2} type, how they can be used to model semantic relations between various senses of polysemic words.

References

- Altman J., Polguère A. (2003) La BDéf: base de définitions dérivée du Dictionnaire explicatif et combinatoire, *Proceedings of the First International Conference on Meaning-Text Theory (MTT-03)*, Paris, 43-54.
- Dowty D. (1979) *Word Meaning and Montague Grammar*, Dordrecht: Reidel.
- Longman (2003) *Longman Dictionary of Contemporary English*, 4th edition, Harlow: Longman.
- Mel’čuk I. *et al.* (1984, 1988, 1992, 1999) *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques*, vol. I-IV, Montreal: Les Presses de l’Université de Montréal.
- Pottier B. (1974) *Linguistique générale. Théorie et description*, Paris: Klincksieck.
- Polguère A. (2000) Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French, *Proceedings of EURALEX’2000*, Stuttgart, 517-527.
- Pustejovsky J. (1998) *The Generative Lexicon*, Cambridge, Mass.: The MIT Press.
- Steinlin J., Kahane S., Polguère A., El Ghali A. (2004) De l’article lexicographique à la modélisation objet du dictionnaire et des liens lexicaux, *Proceedings of EURALEX’2004*, Lorient, 177-186
- Wierzbicka A. (1987) *English Speech Act Verbs: A semantic dictionary*, Sydney: Academic.