Language Sciences 48 (2015) 70–89

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

Argument structure constructions in a Natural Language Processing environment

Alba Luzondo-Oyón [a,*], Francisco J. Ruiz de Mendoza-Ibáñez [b]

[a] Dpto. de Filologías Extranjeras y sus Lingüísticas, Facultad de Filología, Universidad Nacional de Educación a Distancia, c/ Senda del Rey 7, 28040, Madrid, Spain
[b] University of La Rioja, Philology Building, c/San José de Calasanz 33, 26004, Logroño, La Rioja, Spain

ARTICLE INFO

Article history:
Received 12 November 2014
Received in revised form 4 January 2015
Accepted 5 January 2015
Available online 29 January 2015

Keywords:
Construction Grammar
Argument-structure constructions
Natural Language Processing
FunGramKB

ABSTRACT

This paper contributes to the field of computational Construction Grammar (cf. Steels, 2012; Van Trijp, 2011) by providing a linguistically oriented formalized treatment of argument structure constructions within the architecture of a multipurpose lexico-conceptual knowledge base for Natural Language Processing systems known as FunGramKB (Periñán, 2013). More concretely, we analyze three members of the family of the English resultative, namely, the resultative (e.g. *He hammered the metal flat/into a flat sheet*), the caused-motion construction (e.g. *He shoved the canoe into the water*) and the *way* construction (e.g. *He fought his way free/to freedom*), paying special attention to lexical–constructional integration variability. Thus, the present article offers a description of the format of constructional schemata in FunGramKB as machine-tractable representations of constructions and proposes a 'splitting-like' solution (*à la* Boas, 2003) to handle the mismatches resulting from lexical-constructional fusion. We finally argue that, in the case of FunGramKB, a feasible computational implementation must be based on constructional sub-types rather than on broad-scale constructions of the Goldbergian kind.

1. Introduction

Work on Construction Grammar within Cognitive Linguistics is almost three decades old now. It has given rise to a number of approaches, which are largely compatible with one another, and whose differences are more often than not a matter of goals and areas of special focus (for an overview of the different perspectives see Dirven and Ruiz de Mendoza, 2010; Hoffmann and Trousdale, 2013; Butler and Gonzálvez-García, 2014). In the context of this productive research program, the literature has examined a large number of linguistic *constructions* (or entrenched form-meaning/function pairings) of varied nature and complexity. Among them we find the following: 'Let alone' (e.g. *Who would like this crap, let alone buy it?*; Fillmore et al., 1988); 'What's X Doing Y?' (e.g. *What's a physicist doing studying marine life?*; Kay and Fillmore, 1999); 'contrastive reduplication' (e.g. *My car isn't MINE-mine, it's my parents*; Ghomeshi et al., 2004); 'X is *so* N(P)' (e.g. *That's so Monica*; Gonzálvez-García, 2014). Constructions like these arise from attaching special meaning implications, of a personal and interpersonal kind, to their basic predication. Thus, *What's a physicist doing studying marine life?* is not an information question, despite its Wh-interrogative form, but an expression of complaint about a given state of affairs that the speaker

* Corresponding author. Tel.: +34 913988699.
E-mail addresses: aluzondo@flog.uned.es (A. Luzondo-Oyón), francisco.ruizdemendoza@unirioja.es (F.J. Ruiz de Mendoza-Ibáñez).

http://dx.doi.org/10.1016/j.langsci.2015.01.001
0388-0001/

finds to be wrong or odd. A different class of constructional configurations, known as argument-structure constructions (cf. Goldberg, 1995), is based on the linguistic expression of predicate-argument relationships. Some of the members of this latter class are the ditransitive construction (e.g. *John gave Mary a book*; cf. Barðdal et al., 2011), the resultative construction (e.g. *The boy ate himself sick*; cf. Boas, 2003), the subjective-transitive construction (e.g. *I find her so sweet*; cf. Gonzálvez-García, 2009), and the caused-motion construction (e.g. *He pushed the canoe into the* water; cf. Peña, 2009). Argument-structure constructions are the concern of the present paper.

Clearly, argument-structure constructions only represent a small portion of the full inventory of form-meaning pairings of a given language or its *constructicon* (cf. Boas, 2014, p. 94). However, the amount of work devoted to their analysis—in which complementary aspects of their complex nature have been emphasized—is impressive. Many issues have been covered in the literature, such as the universal or language-specific nature of constructions, how to determine analytical categories and their grounding in experience, or the challenges that constructions pose for specific frameworks like Role and Reference Grammar (RRG; cf. Nolan and Diedrichsen, 2013; see also Nolan, 2014). A particular area of controversy is the understanding of the principles that regulate the integration of lexical structure into constructions. In this respect, three broad analytical perspectives can be distinguished. One, which is typical of Goldberg's (1995) work, focuses its attention on capturing broad-scale generalizations on lexical–constructional behavior. A second one, which arises from work on the resultative by Boas (2003, 2008, 2011), tips the scale in favor of including more detailed information in lexical description in order to compensate for the lack of constraining power of Goldbergian generalizations. A third, more recent approach, called the Lexical Constructional Model (LCM; cf. Ruiz de Mendoza, 2013; Ruiz de Mendoza and Galera, 2014), has shed light on the constraining role of some cognitive processes, including metaphor and metonymy, for the integration of lexical structure into constructional characterizations.

Beyond purely theoretical pursuits, some Construction Grammar accounts increasingly show greater interest in the formalization and computational implementation of many of the notions put forward in Construction Grammar and Cognitive Linguistics. Fluid Construction Grammar (Steels, 2012), Embodied Construction Grammar (Bergen and Chang, 2005), and Sign-Based Construction Grammar (Boas and Sag, 2012) share this goal (see also Chang et al., 2012; Van Trijp, 2013 for comparative descriptions). This paper adds to this research trend by presenting readers with a linguistically oriented computational treatment of argument-structure constructions. More specifically, we study several members of the resultative family of constructions (cf. Goldberg and Jackendoff, 2004; Luzondo, 2014; Ruiz de Mendoza and Luzondo, 2014).[1] This family codes changes of state or location resulting from verbal action (e.g. *He kicked Bob black and blue*, *He kicked the ball into the net*, etc.). Van Trijp (2011), for example, has illustrated how argument-structure characterizations can be handled in the context of Fluid Construction Grammar. In his proposal, van Trijp offers a "design pattern" (see Steels, 2012, p. 16) that is capable of dealing with many of the complexities of argument structure and he operationalizes such a pattern within Fluid Construction Grammar. The design pattern is aimed at capturing the multilayered and indirect nature of form-meaning mappings (Van Trijp, 2011, pp. 2–3). To illustrate, note that the same syntactic role may appear in a different surface form depending on the context. Thus, in *He saw him crossing the street*, the masculine pronouns that function as Subject of the main clause and as Subject of the subclause are respectively expressed as *he* and *him*. To account for this and other cases of variable behavior, van Trijp proposes a two-stepped solution. In the first step, lexical and phrasal constructions specify their semantic and combinatorial potential. In the second, argument-structure constructions select an actual value from this potential and implement the linking between semantics and syntax, thus yielding specific realizations (Van Trijp, 2011, pp. 8–9). For example, for the verb *send*, the transitive construction requires the selection of the following values: (i) at the semantic pole, 'sender-Agent' and 'sent-Patient' (while the semantic roles 'Recipient/Goal' are discarded); (ii) at the syntactic pole, Subject and Direct Object (DO) (the Indirect Object (IO) and the Oblique are also left out). This yields utterances like *Jill sent a letter*. The ditransitive construction, on the other hand, selects Agent, Patient and Recipient, which are then mapped onto Subject, IO and DO in *John sent Mary a letter*. What this approach does not seem to cover, however, is the fact that when a given verb fuses with a particular argument-structure construction, the syntactic pattern may impose different values for each realization. A case in point is that of *sweep* and the transitive-resultative construction. In *He swept the floor clean*, the construction selects the semantic-syntactic pairs "Agent-Subject" and "Patient-DO", which arise from lexical projection (cf. *He swept the floor*), and adds the result ingredient (i.e. *clean*). In *He swept the broom to pieces*, by contrast, both "the broom" and "to pieces" are arguments supplied by the construction, as evidenced by the ungrammaticality of **He swept the broom*. Here, the Patient role, which is mapped onto the object of the causal action, is the instrument employed to perform the goal-oriented activity of sweeping. This phenomenon cannot be treated computationally by merely specifying combinatorial potentials in the verbal construction, since there may not always be a straightforward match between lexical and constructional properties, even in cases in which the same verb and the same construction interact with each other (cf. *She ate the cob clean*,[2] *My youngest brother ate himself sick with the grapes*[3]). From a linguistic perspective, the mismatch requires a solution based on the specification of the factors that license the incorporation of "the broom" as an object into the resultative construction. A

[1] The interested reader is further referred to Jiménez and Luzondo (2013) for an RRG account of the English resultative. Departing from Van Valin's (2005, p. 239) description of the resultative pattern, the article offers a fine-grained analysis of such a configuration by drawing on the work on constructional schemas carried out in Diedrichsen (2010, 2011), and Nolan (2011a, 2011b).

[2] Google Books. *Grow Fruits & Vegetables the Way They Used to Taste*. John Festus Adams. 1988. Accessed on Oct. 16, 2014.

[3] Google Books. *White Turtle*. Merlinda Bobis. 1999. Accessed on Oct. 18, 2014.