Grammatical functions: a problematic fundamental concept of LFG?

Agnieszka Patejuk, **Adam Przepiórkowski**

{aep,adamp}@ipipan.waw.pl

Introduction

While LFG emphasises that grammatical functions (GFs) are first-class linguistic entities, not defined via tree-configurational or any other primitives, there is surprisingly little agreement on the definition of particular grammatical functions; the only function investigated in some depth is SUBJ, with a proposal of Falk 2006 to decompose it into two separate (but co-extensive in many of the familiar languages) functions: the most prominent argument of a verb and the argument that is accessible cross-clausally.[1] As the discussion in Dalrymple 2001, pp. 19–24, makes clear, there is no single cross-linguistically valid definition of object in LFG, not even one relating to passivisation; rather, as put in Dalrymple and Nikolaeva 2011, p. 24, "[d]iagnostics targeting nonsubject grammatical functions, specifically objects, also [i.e., as in case of subjects] vary from language to language"; this questions the usefulness of OBJ, etc., as language-independent notions. The aim of this paper is to re-examine the fundamental LFG notion of grammatical functions, summarise known and present less known problems that it faces, and propose a solution in the spirit of HPSG.

GFs are largely redundant

Dalrymple 2001, p. 9, assumes the following repertoire of universally available grammatical functions in LFG (see also Bresnan *et al.* 2015, pp. 97–100): "SUBJect, OBJect, OBJ_θ, COMP, XCOMP, OBLique$_\theta$, ADJunct and XADJunct", where grammatical functions with the θ subscript are not atomic grammatical functions such as SUBJ or OBJ, but they "represent families of relations indexed by semantic roles, with the θ subscript representing the semantic role associated with the argument". In this paper, we concentrate on the governable (i.e. argument) grammatical functions SUBJ, OBJ, OBJ_θ, OBL_θ, COMP and XCOMP.

The first problem, the lack of operational cross-linguistic definition, is mentioned above. Nevertheless, since most previous work, also within LFG, concerns subjects (Falk 2006) and objects (Dalrymple and Nikolaeva 2011), it is relatively clear how they could be defined in many languages: in case of subjects, they could be defined as the only arguments that agree with the verb, and additional criteria could include visibility outside of the minimal domain of the head (e.g., for the purposes of control and raising) and binding. In case of (primary) objects, passivisation could be taken as the most common test, with object agreement and perhaps other phenomena as supporting factors. Such criteria are undoubtedly syntactic and do not boil down to information already present elsewhere in the syntactic or semantic representation, i.e., subjects and objects defined this way are indeed primitive notions. In particular, subjects and objects may have different categorial realisations, including as clauses (see e.g. Dalrymple and Lødrup 2000 and references therein), and they are not defined or restricted thematically (as is evidenced, e.g., by psych verbs).

However, the same cannot be said about other argument grammatical functions which, in the usual LFG practice, are conglomerates of independent syntactic (categorial) and semantic (thematic) properties. Once we exclude subjects and (primary) objects, nominal arguments deterministically map into OBJ_θ, prepositional arguments – into OBL_θ, finite clauses (CPs) – into COMPs and infinitival clauses – into XCOMPs:

(1)

XP:	NP	PP	CP	InfP
GF:	OBJ_θ	OBL_θ	COMP	XCOMP

Interestingly, the range of θ is qualitatively different in case of OBJ_θ than in case of OBL_θ: in the former, the subscript must indicate the thematic role of the (secondary) object, while in the latter – it may also represent a specific preposition[2] (in case of non-semantic prepositions; see e.g. Dalrymple 2001, pp. 26–27, 82, 249, 281). In either case, the information indicated by such a subscript is already present elsewhere in the representation of the sentence: in the s-structure in case of the thematic role (Asudeh and Giorgolo 2012, Asudeh *et al.* 2014) or in the f-structure in case of the idiosyncratic form of the preposition.

While the above considerations apply to English, analyses of other languages seem to have adopted similarly redundant "definitions" of grammatical functions, with the exception of mapping some NPs to OBL_θ on the basis of particular morphological cases (e.g., King 1995, p. 180). An even clearer illustration of this kind of redundancy is provided by Australian languages, where grammatical functions may be defined on the basis of morphological cases: since, as discussed in Nordlinger 1998, § 3.3.3, case features are required in such languages independently of grammatical functions, the question should be whether in such languages grammatical function features are required independently of case.

GFs and dependent sharing

What test may be used to decide whether two different predicates require arguments of the same syntactic function? Dalrymple 2001, p. 366, cites the following contrast, originally from Barbara Partee's dissertation (Hall 1965, p. 66), apparently showing that dependents shared between two coordinated verbs must "must bear the same grammatical function in both conjuncts":

(2)   John washes and polishes his car in the garage.

(3)   *John washes and keeps his car in the garage.

---

[1] See Sag 2007 and references therein for a related discussion within HPSG.

[2] Strictly speaking, this also means that – contrary to textbook expositions – there is no "repertoire of universally available grammatical functions", as they may bear language-specific indices.

While in (2) the locative phrase is an adjunct to both WASH and POLISH, in (3) it is still an adjunct to WASH, but an (oblique) argument to KEEP – hence the ungrammaticality.

However, closer inspection shows that this test regularly contradicts dominant LFG analyses. For example, a locative phrase is required in case of verbs such as RESIDE (McConnell-Ginet 1982, p. 166), so it must be treated as its argument, if one adopts the prevailing view that required dependents are arguments. On the other hand, in case of DIE, such a locative phrase is a prototypical optional adjunct. Hence, the following attested sentences should be ungrammatical, and for the same reason as (3):

(4) If a person resided and died in a foreign country and had assets in US, can the estate be probated in US?[3]

(5) Prime Minister Sir Winston Churchill resided and died in Number 28 on the street called Hyde Park Gate…[4]

(6) We assessed data on Medical Examiner-certified suicide victims aged 65 years or older from 2001 through 2004 who had resided and died in New York City…[5]

Another problematic case is illustrated with the following examples:

(7) I will devour this cake.

(8) I will give Mary this cake.

(9) I will either devour or give Mary the carrot cake my mother baked yesterday.

In (7), *this cake* is the passivisable OBJ, while in (8) – it is an OBJ_θ, as the OBJ position is taken by the passivisable *Mary* (Dalrymple 2001, p. 22). But these two different grammatical functions may be shared, as (9) illustrates.

The problem also occurs in languages different than English. For example, Patejuk 2015, p. 51, discusses the following examples from Polish:

(10) Marek   manipuluje i wysługuje się Marysią.
      Marek.NOM manipulates and lackey   REFL Marysia.INST
      'Marek manipulates and lackeys Marysia.'

(11) Marysia   lubi ale też boi   się Marka.
      Marysia.NOM likes but also be afraid REFL Marek.ACC/GEN
      'Marysia likes but at the same time is afraid of Marek.'

The natural definition of object in Polish is as the passivisable argument; if so, in each example the non-subject argument (*Marysią* in (10) and *Marka* in (11)) bears the OBJ function only in relation to one of the conjoined verbs (to *manipuluje* 'manipulates' and to *lubi* 'likes', respectively). This again violates the claim that shared dependents must bear the same grammatical function in relation to conjoined verbs. One way to attempt to defend this claim would be to revert to the more traditional understanding of the direct object, as the argument in the accusative case. If so, neither of the verbs in (10) takes an OBJ (the shared argument is in the instrumental). However, in (11), one verb, *lubi* 'likes', takes such an accusative object and the other verb, *boi się* 'fears', takes a genitive argument, so the shared argument *Marka* still simultaneously fills two different grammatical function slots.[6] As there is no other reasonable way of defining OBJ in Polish, we must conclude that either it makes no sense (or at least there is no need) to posit OBJ in Polish, or the coordination test based on the contrast from Hall 1965, p. 66, does not work.

In fact, the latter seems to be the case. Without attempting to provide an exhaustive analysis, let us note that in all the grammatical examples where a dependent bearing different grammatical roles is shared, it actually has the same (or sufficiently similar) semantic role in relation to the conjoined verbs. In particular, in the *resided and died* examples, the locative phrase, while obligatory in case of RESIDE and optional in case of DIE, has the semantic role of event location, the same as the locative *in the garage* in Partee's grammatical (2). On the other hand, while the phrase *in the garage* still expresses location in case of (3), it arguably bears two rather different semantic roles with respect to WASH and KEEP, namely, event location in case of the former, but participant location in case of the latter.[7]

Let us finally note that the fact that two predicates may assign different grammatical functions to their shared dependent is not a problem for LFG; as verified in an XLE implementation, all that is required is the assignment of grammatical functions in c-structure rules via functional uncertainty, as in (12), rather than via separate equations, as in (13):

(12) $(\uparrow \{\text{GF1}|\text{GF2}\})=\downarrow$

(13) $(\uparrow \text{GF1})=\downarrow \lor (\uparrow \text{GF2})=\downarrow$

So the only conclusion of this section is that shared dependents do not provide a test for the sameness of grammatical functions, contra Dalrymple 2001, p. 366.

GFs and coordination of unlikes

One of the arguments for treating at least some finite clauses as subjects or objects concerns the possibility to coordinate them with uncontroversial SUBJs and OBJs (Sag *et al.* 1985, p. 165):

(14) That Himmler appointed Heydrich and the implications thereof frightened many observers.

(15) Pat remembered the appointment and that it was important to be on time.

---

[3] http://www.avvo.com/legal-answers/if-a-person-resided-and-died-in-a-foreign-country--206311.html
[4] http://www.apeksdevelopments.co.uk/famous-hyde-park-residents-throughout-history/
[5] http://europepmc.org/abstract/MED/19210947
[6] See Dalrymple and Kaplan 2000 and, especially, Dalrymple *et al.* 2009 on how *Marka* may be analysed as accusative and genitive at the same time.
[7] See, e.g., Koenig *et al.* 2003 on this distinction.

This issue was discussed in Dalrymple and Lødrup 2000 and in Alsina *et al.* 2005 (as well as in other publications). The conclusion of the first paper was to treat COMP as an elsewhere grammatical function, assigned only if the clause is the only possible realisation of the given argument, while the other paper argued in favour of getting rid of the COMP altogether and imposing relevant categorial constraints instead. But the same problem arises for every combination of categories – each category maps to different grammatical functions and a common grammatical function must be chosen somehow.

This problem gets more complicated when one considers the distinction between closed and open grammatical functions. This distinction is problematic for the analysis of examples such as the following, discussed in Patejuk and Przepiórkowski 2014:

(16) Nie chciał pić   ani kanapki.
      NEG wanted drink.INF nor sandwich.GEN

      'He didn't want to drink nor (did he want) a sandwich.'   (Kallas 1993)

In (16) the non-subject argument is a coordinate phrase consisting of an infinitive and a nominal – according to (1), the infinitive corresponds to XCOMP, an open grammatical function, while the nominal corresponds to OBJ (it is not restricted semantically and it passivises), which is a closed grammatical function. Such examples undermine the distinction between closed and open grammatical functions – by analogy with (14)–(15), XCOMP should probably also be treated as an elsewhere function, leaving the OBJ as the common grammatical function. Furthermore, such examples lend support to the suggestion made in Alsina *et al.* 2005 that XCOMP should be removed from the "repertoire of universally available grammatical functions" in the same way as COMP.

Obvious alternative

There is an obvious solution to the problems discussed as above: just as LFG does not distinguish between various types of adjuncts (locative, temporal, manner, etc.), there is no need for it to distinguish between various types of arguments: they could be gathered in a single list-valued attribute, akin to HPSG's ARG-ST. The advantage of such a list representation is that it encodes directly the hierarchy of grammatical functions; while various LFG analyses refer instead to the hierarchy of semantic roles (i.e. to the thematic hierarchy), some – including control, pronominal and anaphor binding, and *wh*-movement (see Dalrymple 2001, pp. 345, 412, Bresnan *et al.* 2015, chs. 9–10, and references therein) are sensitive to this syntactic hierarchy. Without such an encoding of this hierarchy, the principles that make use of it must either remain on the informal level, or the formalisation of each principle must encode it separately – a clear case of a missed generalisation.

To the extent that some core grammatical functions, i.e., functions which take part in processes such as agreement and passivisation, need to be distinguished and cannot be predicted from the position in the argument list, they could be singled out as values of separate attributes (apart from being present on the argument list). This would again follow the HPSG practice; e.g., Heinz and Matiasek 1994 single out the deep subject as the value of the DA (designated argument) attribute, Sag 2007 proposes to encode the argument visible outside of the maximal projection (i.e., roughly, Falk's 2006 PIVOT) as XARG, etc. For example, in case of Polish, where – unlike in English, where each verb has a syntactic subject – some verbs do not have any subjects, not even expletive or PRO subjects, it might make sense to have a separate SUBJ attribute.[8] One such a verb is MDLIĆ 'nauseate', which may occur in two different valency frames illustrated by the two examples below:

(17) Zapach   kwiatów   mdlił   mnie.
      smell.NOM.M.SG flowers.GEN nauseated.M.SG me.ACC

      'The smell of the flowers made me nauseous.'

(18) Mdliło   mnie   od   zapachu   kwiatów.
      nauseated.N.SG me.ACC from smell.GEN.M.SG flowers.GEN

      'The smell of the flowers made me nauseous.'

Given that the accusative argument does not passivise here, only SUBJ needs to be marked, and only in case of (17), as (18) realises this argument as an arguably non-subject[9] prepositional phrase. Hence, the f-structures for these two examples could look as in (19)–(20) below.[10]

It should be noted that getting rid of particular grammatical functions is possible due to the widespread LFG use of Glue Semantics, which, as already observed in Dalrymple *et al.* 1993, pp. 13–14, and Kuhn 2001, § 1.3.3, makes PRED – and also the principles of Completeness and Coherence (Dalrymple 2001, pp. 35–39) – largely superfluous. In particular, Asudeh and Giorgolo 2012 propose to retain PRED but with values reflecting the predicate *sans* its valency; for example, the lexical entry for *mdliło* 'nauseated' as used in (18) will only contain the equation ($\uparrow$ PRED) = 'NAUSEATE' rather than ($\uparrow$ PRED) = 'NAUSEATE$\langle$OBJ$_{ACC}$, OBL$_{OD}\rangle$' or such.
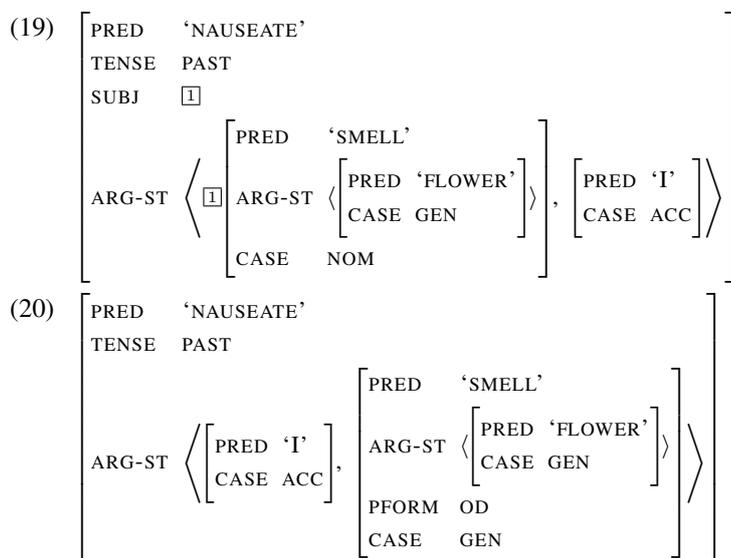
Conclusion

It is surprising how ill-defined, redundant and inconsistent the notion of grammatical functions – claimed to be fundamental in LFG – is on closer inspection. In this paper we propose to get back to the basic LFG assumptions and re-examine the need for a repertoire of grammatical functions as first-class theoretical citizens. We submit that an alternative approach – apparently successfully employed in HPSG – is entertained, namely, that all arguments are listed within a single attribute corresponding to HPSG's ARG-ST, and only some of them are singled out, perhaps in a language-dependent fashion.

---

[8]The value of CASE is not sufficient to distinguish subjects from other arguments in Polish.

[9]See Babby 2009, ch. 1 for a discussion of similar facts in Russian.

[10]The attribute PFORM is commonly used in implemented LFG/XLE grammars to indicate the form of a non-semantic ('case-marking') preposition.

(19)

$$
\begin{bmatrix}
\text{PRED} & \text{`NAUSEATE'} \\
\text{TENSE} & \text{PAST} \\
\text{SUBJ} & \boxed{1} \\
\text{ARG-ST} & \left\langle \boxed{1}
\begin{bmatrix}
\text{PRED} & \text{`SMELL'} \\
\text{ARG-ST} & \left\langle \begin{bmatrix} \text{PRED} & \text{`FLOWER'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \right\rangle \\
\text{CASE} & \text{NOM}
\end{bmatrix},
\begin{bmatrix} \text{PRED} & \text{`I'} \\ \text{CASE} & \text{ACC} \end{bmatrix}
\right\rangle
\end{bmatrix}
$$

(20)

$$
\begin{bmatrix}
\text{PRED} & \text{`NAUSEATE'} \\
\text{TENSE} & \text{PAST} \\
\text{ARG-ST} & \left\langle \begin{bmatrix} \text{PRED} & \text{`I'} \\ \text{CASE} & \text{ACC} \end{bmatrix},
\begin{bmatrix}
\text{PRED} & \text{`SMELL'} \\
\text{ARG-ST} & \left\langle \begin{bmatrix} \text{PRED} & \text{`FLOWER'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \right\rangle \\
\text{PFORM} & \text{OD} \\
\text{CASE} & \text{GEN}
\end{bmatrix}
\right\rangle
\end{bmatrix}
$$

References

Alsina, A., Mohanan, T., and Mohanan, K. (2005). How to get rid of the COMP. In M. Butt and T. H. King, editors, *The Proceedings of the LFG'05 Conference*, University of Bergen, Norway. CSLI Publications.

Asudeh, A. and Giorgolo, G. (2012). Flexible composition for optional and derived arguments. In M. Butt and T. H. King, editors, *The Proceedings of the LFG'12 Conference*, pages 64–84, Stanford, CA. CSLI Publications.

Asudeh, A., Giorgolo, G., and Toivonen, I. (2014). Meaning and valency. In Butt and King (2014), pages 68–88.

Babby, L. H. (2009). *The Syntax of Argument Structure*. Cambridge University Press, Cambridge.

Bresnan, J., Asudeh, A., Toivonen, I., and Wechsler, S. (2015). *Lexical-Functional Syntax*. Blackwell Textbooks in Linguistics. Wiley-Blackwell, 2nd edition.

Butt, M. and King, T. H., editors (2006). *The Proceedings of the LFG'06 Conference*, University of Konstanz, Germany. CSLI Publications.

Butt, M. and King, T. H., editors (2014). *The Proceedings of the LFG'14 Conference*, Stanford, CA. CSLI Publications.

Dalrymple, M. (2001). *Lexical-Functional Grammar*. Academic Press.

Dalrymple, M. and Kaplan, R. M. (2000). Feature indeterminacy and feature resolution. *Language*, **76**(4), 759–798.

Dalrymple, M. and Lødrup, H. (2000). The grammatical functions of complement clauses. In M. Butt and T. H. King, editors, *The Proceedings of the LFG'00 Conference*, University of California, Berkeley. CSLI Publications.

Dalrymple, M. and Nikolaeva, I. (2011). *Objects and Information Structure*. Cambridge University Press.

Dalrymple, M., Hinrichs, A., Lamping, J., and Saraswat, V. (1993). The resource logic of complex predicate interpretation. In *Proceedings of ROCLING 1993*, pages 3–21.

Dalrymple, M., King, T. H., and Sadler, L. (2009). Indeterminacy by underspecification. *Journal of Linguistics*, **45**, 31–68.

Falk, Y. N. (2006). On the representation of case and agreement. In Butt and King (2006).

Hall, B. C. (1965). *Subject and Object in Modern English*. Ph.D. dissertation, Massachusetts Institute of Technology.

Heinz, W. and Matiasek, J. (1994). Argument structure and case assignment in German. In J. Nerbonne, K. Netter, and C. Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, number 46 in CSLI Lecture Notes, pages 199–236. CSLI Publications, Stanford, CA.

Kallas, K. (1993). *Składnia współczesnych polskich konstrukcji współrzędnych*. Wydawnictwo Uniwersytetu Mikołaja Kopernika, Toruń.

King, T. H. (1995). *Configuring Topic and Focus in Russian*. Dissertations in Linguistics Series. CSLI Publications.

Koenig, J.-P., Mauner, G., and Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, **89**, 67–103.

Kuhn, J. (2001). Resource sensitivity in the syntax-semantics interface: Evidence from the German Split NP construction. In D. Meurers and T. Kiss, editors, *Constraint-Based Approaches to Germanic Syntax*, pages 177–215. CSLI Publications, Stanford, CA.

McConnell-Ginet, S. (1982). Adverbs and logical form: A linguistically realistic theory. *Language*, **58**(1), 144–184.

Nordlinger, R. (1998). *Constructive Case: Evidence from Australian Languages*. CSLI Publications, Stanford, CA.

Patejuk, A. (2015). *Unlike Coordination in Polish: An LFG Account*. Ph.D. dissertation, Institute of Polish Language, Polish Academy of Sciences, Cracow.

Patejuk, A. and Przepiórkowski, A. (2014). Control into selected conjuncts. In Butt and King (2014), pages 448–460.

Sag, I. A. (2007). Remarks on locality. In S. Müller, editor, *Proceedings of the HPSG 2007 Conference*, pages 394–414, Stanford, CA. CSLI Publications.

Sag, I. A., Gazdar, G., Wasow, T., and Weisler, S. (1985). Coordination and how to distinguish categories. *Natural Language and Linguistic Theory*, **3**, 117–171.