New LFG Grammatical Function (GF) for (free subject) verb MWEs

Stella Markantonatou and Niki Samaridi
marks@ilsp.gr, nsamaridi@gmail.com

Modern Greek (MG) and English free subject verb MWE (FSMWE) are "signified" by one or more fixed or semi-fixed strings (FSs) consisting of one or more words (in examples (2)-(7) FSs are in bold characters). FSs are part of the FSMWEs. FSs on their own or combined with free strings form FSMWE constituents that can be identified with a subset of (language dependent) constituency diagnostics. Certain constituency diagnostics seem to reflect semantic properties. For instance, in both MG and English, substitution of constituents containing FSs by personal pronouns can be employed as a diagnostic only when the FS has acquired independent "meaning" (see also the discussion on non/decomposable MWEs in Sag et al. (2001), Kay and Sag (2012)) because personal pronouns are used in discourse to keep track of entities. The FS in (1) can be used independently of the particular FSMWE--as can be seen by the variety of verbs with which it combines ('δίνω' (give)—also passive, 'ανάβω' (turn on)-- both causative and inchoative structures, 'παίρνω' (take)-also passive) and its frequent usage in titles (1a)-- and can be replaced by a personal pronoun (1). On the contrary, the FS in (2) has no such independence and can not be replaced by a personal pronoun but it is a constituent as indicated by word order permutations available in MG (4).

(1) έδωσε   **το πράσινο   φως**   για το Erasmus+ - **το**   έδωσε
    give.3rd.sg the green.neu.acc.sg light. neu.acc.sg   for Erasmus+   - it.acc.sg give
    'S/He gave the green light for Erasmus+.'

(1a) Πράσινο φως για το Erasmus+ από το Ευρωπαϊκό ...
    'Green light for Erasmus+ from the European…'

(2) έφαγαν   **την   σκόνη**   **του Διαμαντίδη** – *την   έφαγαν (see (4))
    eat.3rd.sg the.fem.acc.sg dust.fem.acc.sg the Diamantidis – she.fem.acc.sg eat.3rd.pl

We assume that FSs with no independent meaning ('non-independent FSs') have some kind of semantics. FSs like (3)-(7) denote no event or entity. On the other hand, however, they signify the respective FSMWEs and, in MG, they play a role in emphatic structures. (4) says that the Gasol Bros have been OVERTAKEN by Diamantidis. It is the fronting of the FS "τη_σκόνη" that ensures this interpretation; if the free NP "του Διαμαντίδη" was fronted the interpretation would be "DIAMANTIDIS overtook the Gasol Bros". Therefore, we assume that non-independent FSs have own semantics (Sag et al. (2001)) and, a PRED value is defined in their lexical entries within an LFG-oriented approach.

In LFG, constituents project syntactically relevant information on the f-structure. Our study of MG data shows that FSMWE constituents containing an non-independent FS allow for the following syntactic phenomena that are encoded on the f-structure: constituency, emphasis (3)-(4), binding phenomena (5), obligatory free NPs[+gen/+acc] (4) (they also give subject control phenomena but we will not take up this issue and its implications in this outline). Similar phenomena can be observed with English FSs. In (5) we have an obligatory free NP[+gen] (we consider 'of' as a marker of the genitive case) and in (6) binding phenomena.

(3) FS [την_τρίχα_τριχιά]

    [**την   τρίχα   τριχιά**]   κάνει   [Η Αίμη]
    the-fem.acc.sg hair. fem.acc.sg rope.fem.acc.sg make.3rd.sg the Emy
    'Emy EXAGGERATES things.'

(3a) Η Αίμη έκανε πάλι την τρίχα τριχιά -*Η Αίμη έκανε την τρίχα πάλι τριχιά
    'Emy EXAGGERATED things again.'

(4) FS [τη_σκόνη] plus Free NP[+gen]

    [**Τη   σκόνη** του   Διαμαντίδη]   έφαγαν   [οι   αφοί   Γκασόλ]
    the.acc dust.acc the.gen Diamantidis.gen eat.3rd.pl.past the.nom bros.nom Gasol

'Bros Gasol were OVERTAKEN by Diamantidis.'

(5) FS [leg] plus NP's

The students decided to pull [the teacher's **leg**] by telling… http://www.idiomeanings.com/idioms/pull-ones-leg/.

(5a) *I pulled it too.

?The teacher's leg was being pulled by the students.

(6) FS [στ'_αυγά] plus weak pronoun[+gen]

πες στην Ελένη$_j$ να κάτσει$_j$ [**στ' αυγά** της$_j$]

tell.imp to-the Eleni$_j$.fem.3rd.sg to sit$_j$.3rd.sg [on-the eggs hers$_j$.fem.3rd. sg]

'Tell to Eleni not to be involved.'

(7) FS [eyeteeth] plus Possessive pronoun

She would give [her **eyeteeth**] for this job. http://dictionary.reference.com/browse/eyeteeth

(7a) *She would give it for this job.

We consider the strings in square brackets in (5) and (7) as constituents because of the intervening free NP that is fully predicative and indicates a phrasal boundary. The indicated constituents do not allow for pronoun substitution as objects would typically do in English and are less liberal with passivisation (but see Kay and Sag (2012) for a more detailed discussion on passivisation).

Below we argue that the **constituents containing non-independent FSs** exemplified in (2)-(7) all have the same syntactic function that is different from that of the OBJect and the PCOMP therefore a new Grammatical Function is needed to encode it.

Why the same syntactic function—therefore one grammatical function: (i) they are phrasal structures headed by non-independent FSs (ii) they have own semantic contribution (iii) they are directly dependent on the verb--(6) contains a preposition ("σε" (to)) but the preposition is part of the FS, it is not predicative and there is no independent predicative preposition introducing the FS (iv) they allow for phenomena such as binding, modification ("I kicked the proverbial bucket") and an obligatory NP[+gen] that could be considered selected (v) they can not be replaced by personal pronouns.

Why not an OBJect (where the FS does not contain a P): (i) the FSs may be words (5), (7) or fragments of (probably different (3)) clauses (4), (6) and very often it is not clear to which Part of Speech they belong (3),(4),(6); it seems that one needs to postulate a "new" Part of Speech, let us call it 'FS', and talk about projections of the type FSP. Such projections will be typical of the new GF as NPs are typical of the SUBJ and OBJ GFs in both English and MG (ii) objects can be replaced by personal pronouns and, in MG, they allow for clitic doubling; as already discussed, none of the constituents containing non-independent FSs (3)-(7) allow for such phenomena (iii) they are more resistant to passivisation than objects in both languages.

Why not a PCOMP (where a P exists in the FS): "κάθομαι" (sit) (6) subcategorises for a prepositional complement, a PCOMP in LFG terms. However, the structure "στ' αυγά της" does not seem to contain a P predicate since no substitution of the PP with an adverb or substitution of the supposed NP complement of the P apply.

We define a new grammatical function, FIXCONST with value received from the projection of a phrasal constituent that is headed by a non-independent FS. In this way, rules that apply on objects such as substitution of an object XP by a personal pronoun will not apply on constituents headed by FSs.

References

Dalrymple, Mary. 2001. *Lexical Functional Grammar. Syntax and Semantics 34*. Academic Press

Kay, Paul and Ivan A. Sag. 2012. **A Lexical Theory of Phrasal Idioms** Unpublished manuscript, Stanford University, 2012 - icsi.berkeley.edu

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2001. Multi-word Expressions: A Pain in the Neck for NLP. Paper presented at *LinGO Working Papers*, Stanford University, CA.