

Particle Verbs in Computational LFGs: Issues from English, German, and Hungarian

Martin Forst, Tracy H. King, and Tibor Laczkó

In a number of languages, especially Germanic and Finno-Ugric, there are classes of verbs commonly called “particle verbs” (Ackerman 1983, Piñón 1992, Lüdeling 2001, Toivonen 2001). Particle verbs are verbs whose meaning and argument structure depends on the combination of a (base) verb and a particle. Often the meaning and argument structure of a particle verb is not compositional, i.e. it is not predictable from the combination of its components, but it must be listed in the lexicon. An example of a meaning expressed by a particle verb in English, German, and Hungarian is *He gave up the fight.* = *Er gab den Kampf auf.* = *Ő fel#adta a küzdelmet.* However, particle verbs can also be compositional (*push them up/in/out; push up/in/out the boxes*) and highly productive, which is a challenge for the coverage of computational grammars (Villavicencio 2003).

In this paper, we present the, so far undocumented, ways in which particle verbs are implemented in two relatively mature computational grammars, the English and the German *ParGram* LFGs (Butt et al. 2002), and we will address the issues that arise with respect to particle verbs in the development of a computational LFG for Hungarian. We will see that considerations concerning the *ParGram* LFG implementation of productive Hungarian particle + verb combinations raise questions as to the current treatment in the other two grammars. In addition, a set of Hungarian particles exhibit inflectional properties as well; we will also outline an LFG analysis of this phenomenon.

Particle verbs — syntactic or morphological objects? English particle verbs are typically analyzed in such a way that the two components are separately inserted in their respective syntactic positions, which is not surprising given that particles are always written as separate words and short NPs can intervene between base verbs and particles. In German and Hungarian, however, particle + verb combinations are generally spelled as a single word when the particle immediately precedes the verb (although a certain variation with respect to the spelling as one or two words can be observed with semantically compositional particle + verb combinations), and this order is in a way the default order, since only clearly definable conditions (V1 and V2 in German; negation, imperatives, etc. in Hungarian) cause particles to appear in positions other than the immediately preverbal one. In addition, there are verbs in German and Hungarian that do not exist on their own, but only in combination with particles; examples of such verbs are *aus#flippen* ‘to flip/freak out’ (German) and *be#fejez* ‘to finish’ (Hungarian). As a result, there is substantial controversy in the linguistic literature concerning the status of particle + verb combinations as syntactic or morphological objects. We will argue for a uniformly syntactic treatment of particles across the three LFG implementations and offer analyses that nevertheless capture the lexical properties of particle verbs in a principled manner.

Current Implementations in the English and German *ParGram* LFGs As verb particles are always spelled as separate words in English, particle verbs receive a syntactic analysis in the English *ParGram* LFG. The lexical entries of verb particles contribute a feature called PRT-FORM, which simply records the form of the respective particle, and the lexical entries of base verbs introduce the semantic form of the particle verb with its argument structure. Finally, the lemma of the base verb and the form of the particle are concatenated via an implementational device (CONCAT) so that the combination of the two, rather than just the lemma of the base verb, is the PRED of the respective f-structure. All particle verbs are listed with their argument structures in the verb lexicon of the grammar, and they appear under the corresponding base verb, but restricted to co-occurring with the appropriate particle. Below are the lexical entries involved in the analysis of our English example sentence to illustrate this treatment, as well as the f-structure associated with it. This analysis captures the syntactico-semantic facts in that the PRED reflects the potentially idiosyncratic particle verb meaning and the corresponding argument structure. However, it does not allow the system to construct productive particle verbs on the fly nor does it differentiate between compositional and non-compositional particle verbs.

<pre>give V (^ PRED) = '%NewPred<(^ SUBJ)(^ OBJ)>' (^ PRT-FORM) = c up @ (CONCAT %stem # (^ PRT-FORM) %NewPred).</pre>	$\left[\begin{array}{l} \text{PRED } \text{'give\#up<[1:he], [99:fight]>'} \\ \text{SUBJ } 1 \left[\text{PRED } \text{'he'} \right] \\ \text{OBJ } 99 \left[\begin{array}{l} \text{PRED } \text{'fight'} \\ \text{SPEC } \left[\text{DET } \left[\text{PRED } \text{'the'} \right] \right] \end{array} \right] \end{array} \right]$
<pre>up PART (^ PRT-FORM) = up.</pre>	

In German V1 and V2 clauses, particle verbs are spelled as separate words. In these contexts, the German *ParGram* LFG thus treats them in the same way as its English counterpart. In verb-final clauses and in headed VPs, however, particle verbs are usually spelled as single words. Compare, e.g., *Er lud seine Kusine ein.* ‘He invited his cousin.’ and *Er wird seine Kusine einladen.* ‘He will invite his cousin.’

The finite-state morphology currently used by the German *ParGram* LFG outputs analyses like the following for forms of particle verbs:

```
einlud
ein#laden +V .13 .Sg .Past .Ind
```

The hash mark indicates the boundary between the particle and the base verb and thus potentially disambiguates analyses involving a separable verb particle from analyses involving homophonous non-separable verb prefixes; however the entire lemma is still a single unit. As a result, the grammar must analyze spelled-together particle verbs as morphological objects, and the lexical information for the particle verb *ein#laden* must be listed both under the base verb lemma (as in the English *ParGram* LFG), e.g. *laden*, and under the particle verb lemma, e.g. *ein#laden*. In order to allow for a uniformly syntactic analysis of particle verbs like the one in the English grammar, the analysis produced by the finite-state morphology would have to separate the particle from the verb, as done, e.g., by SMOR, a morphology developed at the IMS of Stuttgart University:

`ein <VPART> laden <+V> <13> <Sg> <Past> <Ind>`

As in the English grammar, all particle verbs must currently be listed with their argument structures in the German verb lexicon, so that the system exhibits the same limitations with respect to productively formed combinations. The CONCAT template makes it possible to project analogous f-structures regardless of whether a given particle verb is spelled together or as separate words.

Compositional and Productively Formed Particle Verbs As pointed out already, the implemented analyses do not differentiate compositional particle + verb combinations from idiomatic particle verbs. This is a problem for the coverage of computational grammars because new combinations inevitably show up in corpus texts and because the regular character of these combinations is not captured. E.g., the particles *hinterher* (German) and *rá* (Hungarian) can basically combine with any motion verb and (optionally) introduce an OBJ_θ/OBL_θ, as the following sentences exemplify:

Lauf	dem	Glück	nicht	länger	hinterher!	Mari	rá-lépett	a	doboz-ok-ra.
run.IMP.2SG	the-DAT	happiness	not	longer	after	Mari.NOM	onto step.PAST.3SG	the	box-PL-onto
'Don't run after happiness any longer!'						'Mari stepped onto the boxes.'			

This behavior can be analyzed by means of a lexical entry for the particle where, rather than a PRT-FORM feature, it contributes a PRED that subcategorizes for the argument it introduces, and a predicate composition rule involving restriction similar to the one proposed for Urdu causatives by Butt et al. (2003). Other productively used particles fill argument slots of the base verb or simply contribute aspectual information. We will provide fully worked-out and implemented analyses at the conference.

Hungarian Inflected Preverbs In addition to the uninflected particles found in Germanic, Hungarian has inflected preverbs: when certain particle verbs take a pronominal argument, their preverbs are inflected for the person and number features of this argument. The pronominal argument is not overtly realized in that case.

Mari	rá-juk	lépett.
Mari	onto-3PL	step.PAST.3SG
'Mari stepped onto them.'		

We propose that these inflected particles can be straightforwardly treated in LFG by dint of an analysis whereby the inflected particle provides the PRED 'pro', as well as number and person information, for the particle verb's argument (e.g. *rájuk* (↑ OBL PRED)='pro', (↑ OBL PERS)=3, (↑ OBL NUM)=pl).

Conclusions We propose that the implemented LFG analysis of particle verbs for English and German is appropriate and feasible for non-compositional particle constructions in Hungarian. In addition, LFG, and its implementation via XLE, allows for a straight-forward analysis of Hungarian inflected preverbs. A closer examination of the productive, compositional Hungarian particle constructions has resulted in a re-examination of the analysis of such constructions in English and especially German. An orthogonal issue to those addressed here is that of how particle verbs participate in derivational morphology and how best to implement this (e.g. English *bystanders*, German *Einladung* 'invitation'); we leave this area for future work.

References

- Ackerman, F. (1983)** *Miscreant morphemes: phrasal predicates in Ugric*. PhD thesis, UC Berkeley. **Booij, G. (1990)** The boundary between morphology and syntax: separable complex verbs in Dutch. In *Yearbook of Morphology 3*. **Butt et al. (2002)** The Parallel Grammar Project. In *Proceedings of the COLING 2002 Workshop on Grammar Engineering and Evaluation*. **Butt et al. (2003)** Complex Predication via Restriction. In *Proceedings of the LFG 2003 Conference*. **Dehé, N. et al. (2002)** *Verb-Particle Explorations*. Mouton de Gruyter. **Lüdeling, A. (2001)** *On Particle Verbs and Similar Constructions in German*. CSLI Publications. **Piñón, C. J. (1992)** The preverb problem in German and Hungarian. In *Proceedings of BLS, 395-408*. **Toivonen, I. (2001)** *The Phrase Structure of Non-Projecting Words*. PhD thesis, Stanford University. **Villavicencio, A. (2003)** Verb-Particle Constructions and Lexical Resources. In *Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*.