

Urdu Ezafe and the Morphology-Syntax Interface

The purpose of this paper is two-fold: 1) introduce data from Urdu into the discussion of the well-known Persian *ezafe* construction (e.g., Ghomeshi 1997, Samvelian 2007); 2) continue the discussion on the morphology-syntax interface that was begun at the LFG00 conference (Sadler and Spencer 2000). While the Urdu *ezafe* construction is not as complex as its Persian counterpart, it is interesting in its own right and raises some of the same issues with respect to the interplay of phonology, morphology and syntax. In contrast to the conclusions as to the morphology-syntax interface arrived at in Samvelian (2007) for HPSG and in Luís and Otaguro (2005) for LFG, which both assume that the tension between lexical/affixal properties of phrasal affixes (Zwicky 1997) and their ability to take wide phrasal scope must be resolved by introducing new technology that maps between the morphology and the syntax, we demonstrate that LFG's projection architecture as it allows for a straightforward account of the properties of Urdu *ezafe* and phrasal affixes in general. In particular, we invoke postlexical prosodic phonology in order to ensure the correct placement of clitics, while accounting for their phrasal scope and lexical selectional properties in terms of c-structure representations and f-structure constraints, respectively.

Urdu is an Indo-Aryan South Asian language that is ultimately descended from (versions of) Sanskrit. However, in the course of its history, it has also undergone intensive contact with Arabic, English and Persian. As Persian was the language of the royal Moghul court for several centuries, poetry and the high literature in particular were highly influenced by Persian. That is, quite a bit of the vocabulary of "High Urdu" is due to Persian. Furthermore, the language borrowed pieces of morphology and syntactic constructions such as the Persian *ezafe*. The Persian *ezafe* construction, illustrated in (1) has been described and analyzed in some detail in Ghomeshi (1997) and, more recently, in Samvelian (2007).

- (1) in ketâb-e kohne-ye bi arzeš-e maryam
this book-Ezafe ancient-Ezafe without value-Ezafe Maryam
'this ancient worthless book of Maryam's' (Samvelian 2007:606)

The *ezafe* (or *zafat*, as it is referred to in the Urdu tradition) is generally realized as an *-e* in both Persian and Urdu and allows for further modification of the head noun. Several stackings of *ezafe* are possible, as illustrated in (1) for Persian and (2) for Urdu.

- (2) qüvat-e axuvat-e avam
strength-Ezafe brotherhood-Ezafe people
'The strength of the brotherhood of the people' (From the National Anthem of Pakistan)

There are interesting syntactic restrictions with respect to Persian *ezafe*, leading Ghomeshi (1997) to analyze it as a domain of X^0 projection (the *ezafe* finds its origins in an old relative pronoun/demonstrative, cf. Samvelian (2007:606) and references therein). With respect to Urdu, the *ezafe* construction functions comparatively simpler. There are two basic patterns: N-*ezafe* Adj ((3a)) and N-*ezafe* N ((3b)). In case of the N-*ezafe* N construction, the relationship is usually one of Possessee and Possessor or kinship. Note that this possessive relation is realized in the direct mirror-image of the more usual/colloquial genitive construction ((3c)). This demonstrates that the whole *ezafe* construction was borrowed from Persian, not just the *ezafe* morph by itself. Furthermore, *ezafe* shows a lexical restriction in that only combinations with "high language" Persian or Arabic-based words are allowed. So, *ezafe* in combination with the standard Urdu noun *pani* 'water' is disallowed, rather, the Persian-based noun *ab* 'water' must be used.

- (3) a. vazir-e azam
minister.M.Sg-Ezafe great
'the prime minister' (lit. the great minister)
- b. hükumat-e pakistan
government.F.Sg-Ezafe Pakistan.M.Sg
'The government of Pakistan.'

- c. pakistan=ki hukumat
 Pakistan.M.Sg=Gen.F.Sg government.F.Sg
 ‘The government of Pakistan.’

In a recent paper, Samvelian (2007) revisits Ghomeshi’s original analysis and comes to the conclusion that *ezafe* should instead be recognized as a phrasal affix along the lines of Zwicky (1997). We find Samvelian’s arguments for the status of *ezafe* as a phrasal affix convincing for Persian and show that the Urdu *ezafe* also cannot be analyzed as inflectional morphology, but has retained its status as a phrasal affix. This is demonstrated by examples as in (4), where the *ezafe* takes scope over a coordinated structure.

- (4) [ye mal o daulat]=e dunya
 this material and wealth=Ezafe world
 ‘This material and wealth of the world.’ (in *zarb-e-kaleem* by Muhammad Iqbal)

However, with respect to the formal analysis of *ezafe*, we differ from Samvelian, who adopts Zwicky’s perspective and seeks to resolve the tension between the **affixal** nature of *ezafe* (it is not “promiscuous” with respect to what it attaches to, e.g., in Urdu it is restricted to certain types of nouns) and its **phrasal** nature (takes scope over phrases, unlike what affixes can usually do) within an HPSG analysis by means of affixation via a lexical rule in combination with an EDGE feature that percolates through the tree, ensuring the right placement of the *ezafe*. Samvelian rejects an Anderson-type postlexical approach (Anderson 2005) on the grounds that this approach does not allow access to individual word-level properties. This is also essentially the problem that Luís and Otoguro (2005) see within an LFG approach (though they discuss it with respect to violations of lexical integrity), leading them to propose a new type of mapping within the morphology-syntax interface. In contrast, we show that the standard assumptions about the architecture of LFG allow a straightforward and elegant account of Urdu *ezafe* and phrasal affixes in general. We present an implementation of Urdu *ezafe* which uses an integrated finite-state morphological analyzer as described in Butt et al. (1999) and Kaplan et al. (2004). The *ezafe* is treated as a lexical item in that it has its own lexical entry and is realized as a leaf c-structure. Its idiosyncratic affixal properties are realized via constraints that are percolated/checked with the f-structure representation. Finally, its phrasal properties are realized by a combination of c-structure representation, whereby the scope of the *ezafe* is realized in terms of c-structure scope over constituents, but the phonological/prosodic attachment to just one member of the constituent is realized postlexically in terms of the p(rosodic)-projection proposed by Butt and King (1998). That is, the *ezafe* is analyzed as being part of the same prosodic word as the preceding noun.

Based on our analysis and implementation of Urdu *ezafe*, we thus argue that there is no need to augment the existing architecture (and, in particular, the morphology-syntax interface) of LFG. Rather, if one analyzes/situates the right aspects of phenomena involving phrasal affixes at the right levels of representations, the existing architecture proves more than adequate.

References

- Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics*. Oxford University Press.
- Butt, Miriam, Tracy H. King, Maria-Eugenia Niño and Frederique Segond. 1999. *A Grammar Writer’s Cookbook*. CSLI Publications.
- Butt, Miriam and Tracy H. King. 1998. Interfacing Phonology with LFG. *On-line Proceedings of LFG98*.
- Luís, Ana R. and Ryo Otoguro. 2005. Morphological and Syntactic Well-formedness: The Case of European Portuguese Clitics. *On-line Proceedings of the LFG05 Conference*.
- Ghomeshi, Jila. 1997. Non-projecting nouns and the Ezafe construction in Persian. *Natural Language and Linguistic Theory* 15(4):729–788.
- Kaplan, Ron, John Maxwell III, Tracy H. King, and Richard Crouch. 2004. Integrating finite-state technology with deep LFG grammars. In *Proceedings of ESSLLI, Workshop on Combining Shallow and Deep Processing for NLP*.
- Sadler, Louisa and Andrew Spencer. 2000. Workshop on Morphology and LFG. *On-line Proceedings of LFG00*.
- Samvelian, Pollet. 2007. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43(3):605–645.
- Zwicky, Arnold. 1987. Suppressing the Z’s. *Journal of Linguistics* 23(1):133–148.