Urdu Ezafe and the Morphology-Syntax Interface

Tina Bögel, Miriam Butt, Sebastian Sulger

Universität Konstanz

LFG 2008, Sydney
Aim of Paper

1. Introduce hitherto undiscussed phenomenon of Urdu ezafe.
Aim of Paper

1. Introduce hitherto undiscussed phenomenon of Urdu *ezafe*.
2. Discuss and illustrate issues of the morphology-syntax-prosody interface with respect to *clitics* and *phrasal affixes* in general.
Aim of Paper

1. Introduce hitherto undiscussed phenomenon of Urdu *ezafe*.
2. Discuss and illustrate issues of the morphology-syntax-prosody interface with respect to *clitics* and *phrasal affixes* in general.
3. Propose that there is no real distinction between phrasal affixes and clitics — that is, phrasal affixes are just those clitics which are on their way towards *reanalysis* as pieces of inflectional and derivational morphology.
Aim of Paper

1. Introduce hitherto undiscussed phenomenon of Urdu *ezafe*.
2. Discuss and illustrate issues of the morphology-syntax-prosody interface with respect to *clitics* and *phrasal affixes* in general.
3. Propose that there is no real distinction between phrasal affixes and clitics — that is, phrasal affixes are just those clitics which are on their way towards *reanalysis* as pieces of inflectional and derivational morphology.
4. Illustrate our proposals for the morphology-syntax-prosody interface in general via an implementation of Urdu *ezafe* in particular.
Urdu contains the following constructions, which are generally written as one or two words in the grammar books (??).

(1) 

hukuumat-e-paakistaan  
government-Ez-Pakistan  
‘the government of Pakistan’  
Schmidt (1999:246)  
Urdu

(2) 

jahr-e  bagdaad  
city-Ez Bagdad  
‘the city of Bagdad’  
Platts (1909:62)  
Urdu
Usual Urdu NP Syntax

This construction is strange since Urdu NP syntax usually conforms to the head-final pattern that is (almost) pervasive in this SOV language.

(3)

eek laal gaarii
one red car.F.Sg
‘one/a red car’

(4)

paakistaan=kii hukuumat
Pakistan=Gen.F.Sg government.F.Sg
‘Pakistan’s government’

Schmidt (1999:246) Urdu

Urdu Ezafe

hukuumat-e-paakistaan
government-Ez-Pakistan
**Question:** What should the analysis be?

- Is the -e a morphological affix?
  - Is it a phonological “linker” (like the German -s/-en- within compounds)?
  - Is it a clitic like the most of the case markers in Urdu (?), including the genitive?

- What is the constituent structure?

**Example:**

```
Urdu Ezafe

hukuumat-e-paakistaan
government-Ez-Pakistan
‘the government of Pakistan’
```
Urdu Ezafe — A Borrowing from Persian

- The Urdu construction in fact has been borrowed from Persian (e.g., Platts 1906), where it is known as the Ezafe (from Arabic *iḍāfa* ‘adjunction/addition’) construction.
- The Urdu term: *izaafat*.
- The Ezafe construction tends to be part of the high/literary language since Persian was the language of the Mughal court and mostly influenced the language of the courtiers and poets.
- It remains productive today, but modern speakers show a tendency to leave it out (Schmidt 1999:247).
- In the Urdu Arabic-based script there are four different diacritic ways of rendering the Ezafe -e — and like most vowel diacritics is often left out.
Urdu Ezafe — A Borrowing from Persian

- If the Urdu version was directly borrowed from Persian, then let us see what we can learn from existing analyses of the Persian Ezafe.
- Preview of Findings:
  - The Persian construction is more complex.
  - However, the Urdu version functions very similarly.
  - In both Urdu and Persian the ezafe -e turns out to be a clitic.

Question to be Answered: How is this clitic -e best analyzed with respect to the morphology-syntax-prosody interface?
The Persian Ezafe Construction

The Persian Ezafe (y)-e construction is quite well known (???????).

\(5\)

\[
\text{[in ketâb]-e [kohne]-ye [bi arzeš]-e maryam}
\]
\[
\text{this book-Ez ancient-Ez without value-Ez Maryam}
\]
\[
\text{‘this ancient worthless book of Maryam’s’} \quad \text{Samvelian (2007:606)}
\]

- The Ezafe licenses a modifier of a head noun to the right (NPs are generally head-final, just as in Urdu).
- Prosodically, the Ezafe attaches to a nominal host to its left: nouns, adjectives and some PPs.
- Historically it is derived from a relative pronoun \textit{hya}. 
The Ezafe Construction

The Ezafe Construction

- Ghomeshi:
The Ezafe Construction

- Ghomeshi:
  - Provides an X-bar account which involves non-projecting heads.
The Ezafe Construction


- Ghomeshi:
  - Provides an X-bar account which involves non-projecting heads.
  - The Ezafe itself is inserted at PF as a linker for non-projecting heads and is only triggered by lexical heads.

Ghomeshi:
- Provides an X-bar account which involves non-projecting heads.
- The Ezafe itself is inserted at PF as a linker for non-projecting heads and is only triggered by lexical heads.

Samvelian (2007) reexamines the standard analysis, provides new empirical facts and proposes an alternative analysis couched within HPSG.
The Ezafe Construction


- Ghomeshi:
  - Provides an X-bar account which involves non-projecting heads.
  - The Ezafe itself is inserted at PF as a linker for non-projecting heads and is only triggered by lexical heads.

- Samvelian (2007) reexamines the standard analysis, provides new empirical facts and proposes an alternative analysis couched within HPSG.

- For the purposes of our analysis, we take Samvelian (2007) as a point of departure.
Properties of Persian Ezafe:

- Previous claim: attaches only to heads, not phrases.
- Samvelian shows that Ezafe can attach to phrases as well — the Ezafe-licensed modifiers of the head noun can be phrasal.

(6)
mojgân-e [az rimel sangin]-e maryam
eyelid.Pl-Ez of mascara sangin-e Maryam
‘Maryam’s mascara-laden eyelids’ Samvelian (2007:635) Persian
Samvelian’s HPSG Analysis:

- Concludes that Ezafe is a *phrasal affix* (and therefore part of word-level morphology).

- Ezafe is introduced via lexical rules: 1) a *word-to-word* type; 2) *plain-word* to *phrasal-affix-word*.

- Placement at the right edge of a constituent is due to an EDGE constraint (Miller 1992).

- Ezafe introduces/licenses a Dependent (DEP) to the right of the head/phrase — when the DEP has been licensed, the EZ feature can be reset to [–EZ].

- When a dependent is found, the [+DEP] feature can be reset to [–DEP].
Persian Ezafe — Samvelian

↑ = Lexical Rule Application

\[
\begin{aligned}
&\overline{N}[-EZ, -DEP] \\
&\quad | \\
&\quad \overline{N}[-EZ, +DEP] \\
&\quad \quad | \\
&\quad \overline{N}[+EZ] \\
&\quad NP[-EZ] \\
&\quad Maryam
\end{aligned}
\]

\[
\begin{aligned}
&N[+DEP] \\
&\quad mojgân-e
\end{aligned}
\]

\[
\begin{aligned}
&AP[+EZ] \\
&\quad A[+EZ] \\
&\quad sangin-e \\
&\quad A[-EZ] \\
&\quad sangin
\end{aligned}
\]

\[
\begin{aligned}
&P \\
&az \\
&NP[-EZ] \\
&rimel
\end{aligned}
\]

\[
\begin{aligned}
&N[-EZ] \\
&mojgân
\end{aligned}
\]
Properties of Urdu Ezafe: Headedness

- As in Persian, the syntactic head is the noun to the very left.
- Evidence comes from verb agreement.
- In (7) and (8) the adjective agrees with the leftmost noun.

(7) \[\text{vaadii} = \text{e sind}\^h \quad \text{bahut bar-ii} \quad \text{hai}\]
valley.F.Sg=Ez Indus.M.Sg very big-F.Sg be.Pres.3.Sg
‘The Indus valley is very big.’

(8) \[\text{daryaa} = \text{e sind}\^h \quad \text{bahut bar-aa} \quad \text{hai}\]
valley.M.Sg=Ez Indus.M.Sg very big-M.Sg be.Pres.3.Sg
‘The Indus valley is very big.’
Properties of Urdu Ezafe: Constituency

Unlike Persian:

- One does not have the interesting recursive property of Persian — as far as we have been able to determine, there is only ever one Urdu Ezafe complement.
- Complements of Urdu Ezafe cannot be phrasal.
- But the head of the Ezafe construction can be phrasal.

(9)  
\[\text{[ye \ bār-aa diivaan]}=e \quad \text{aam}\]  
this big-M.Sg hall of audience=Ez public  
‘this public hall of audience’

(10)  
\[\text{[har bār-e diivaan]}=e \quad \text{aam]}=mē\]  
ev every big-Obl hall of audience=Ez public=in  
‘in every big public hall of audience’
Properties of Urdu Ezafe: Constituency

We assume that the Ezafe forms a constituent with the complement it introduces — motivated by its historical source as a Relative Clause.

Phrase-Structural Analysis

\[
\text{NPez} \rightarrow \text{NP EzP} \\
\text{EzP} \rightarrow \text{Ez} \{N | A\}
\]

However, we could just as well assume a ternary-branching tree:

\[
\text{NPez} \rightarrow \text{NP Ez} \{N | A\}.
\]
Properties of Urdu Ezafe: Status of Ezafe

- As with case clitics, the Urdu Ezafe can take scope over a coordination.

(11)  
[maal or daulat]=ko kumaa-o  
material and wealth=Acc earn-imp.Rude  
‘Earn/gather material and wealth!’

(12)  
[maal or daulat]=e dunyaa  
material and wealth=Ez world  
‘the material and wealth of the world’ (from Ghalib)

- In Urdu, morphological inflections are not able to do this.

- We therefore treat the Urdu Ezafe as a **clitic**.
LFG Analysis of Urdu Ezafe

LFG is an inherently modular theory of grammar.

- The Principle of Lexical Integrity means that Word Formation is taken to be the domain of Morphology.
- Anything beyond the Word Level is the provenance of Syntax (build phrases out of words).
- Postlexical and Prosodic Phonology interacts with Syntax.
- Lexical Phonology interacts with the Morphology.
LFG Analysis of Urdu Ezafe

- *Clitics* are “little words” — attaching a clitic to a word is **not** the business of Morphology.

- However, *clitics* tend to not be able to form prosodic words on their own, which is why they are phrased together with another prosodic word as part of the prosodic phrasing (prosodic phonology).

- The introduction/placement of “special” clitics (cf. ?) like the Ezafe, which have a given syntactic distribution, is the business of Syntax.

Phrase-Structural Analysis

\[
\text{NPez} \rightarrow \text{NP EzP} \\
\text{EzP} \rightarrow \text{EZ} \{ \text{N} \mid \text{A} \}
\]
LFG Analysis of Urdu Ezafe

- Lexical Entry for Ezafe looks like this:

**Lexical Entry for Ezafe**

\[ e \text{ EZ}^* (\hat{\text{MOD MOD-TYPE}}) = \text{ezafe} \]

- Feature MOD-TYPE signals the quality of the modification

- Phrase-Structural Analysis makes sure that a modifying Noun or Adjective follows after the Ezafe:

\[ \text{EzP} \rightarrow \text{EZ} \{ \text{N} | \text{A} \} \]
LFG Analysis of Urdu Ezafe

- Morphology Output of Ezafe and Nouns:

$\text{xfst[1]}$: up \{sher|panjAb\}
\{sher|panjAb\} + Noun + Unmarked + Masc + Sg + Nom
→ Nouns are present in Analyzer

$\% \text{ analyze-string } \{e\}$
analyzing \{e\}
e " + Token"
→ Ezafe is not present in Analyzer (included directly in LFG Grammar)
LFG Analysis of Urdu Ezafe

- C-Structure Analysis:

```
CS 1: NP
     | NPez
     |       EzP
     |        N
     |         EZ
     |          N
     |           sher e panjAb
```

- Ezafe forms Ezafe Phrase together with modifying Noun
LFG Analysis of Urdu Ezafe

■ F-Structure Analysis:

"sher e panjAb"

\[
\begin{array}{l}
\text{PRED } 'sher' \\
\quad \quad \quad \quad \text{PRED } 'panjAb' \\
\quad \quad \quad \quad \text{MOD } \{ \text{NTYPE } \text{NSEM } \text{COMMON count} \} \\
\quad \quad \quad \quad \text{NSEM } \text{common} \\
\quad \quad \quad \quad \text{GEND } \text{masc}, \text{MOD-TYPE } \text{ezafe}, \text{NUM } \text{sg}, \text{PERS } 3 \\
\text{CHECK } [\text{EZAFE }] \\
\text{NTYPE } \text{NSEM } \text{COMMON count} \\
\text{NSEM } \text{common} \\
1 \text{GEND } \text{masc}, \text{NUM } \text{sg}, \text{PERS } 3
\end{array}
\]

■ P-(Prosodic-)Structure Analysis:

\[
\begin{array}{l}
83 \{1 \text{CL-FORM } \text{ezafe}, \text{DOMAIN } \text{P-WORD}, \text{P-FORM } 'sher' \} \\
129 \{2 \text{DOMAIN } \text{P-WORD}, \text{P-FORM } \text{panjAb} \} \\
135 \text{DOMAIN } \text{P-PHRASE}
\end{array}
\]
LFG Analysis of Urdu Ezafe

- We differentiate Ezafe Construction and "standard" possessive NP Construction involving SPEC [POSS]:

- SPEC [POSS] Analysis:

```
"pAkisAn kI hukUmat"

PRED  'hukUmat'
NTYPE [NSEM [COMMON count]
NSYN common

SPEC  POSS
NTYPE [NSEM [PROPER [LOCATION-TYPE country, PROPER-TYPE location]]
NSYN proper
SEM-PROP [SPECIFIC +]
1 CASE gen, NUM sg, PERS 3
15 GEND fem, NUM sg, PERS 3
```
LFG Analysis of Urdu Ezafe

Ezafe Construction:

"hukUmat e pAkistAn"

```
PRED  'hukUmat'
  | PRED  'pAkistAn'
  | NTYPE  [NSEM [PROPER [LOCATION-TYPE country, PROPER-TYPE location]]]
  | SEM-PROP [SPECIFIC +]
  | 21 MOD-TYPE ezafe, NUM sg, PERS 3
CHECK [EZAFE +]
NTYPE [NSEM [COMMON count]]
  | NSYN common
  | 1 GEND fem, NUM sg, PERS 3
```
Analysis of Urdu Ezafe

Should do phrasal affix? — Why we do not need to. Many fine distinctions between types of clitics, but basically don't need them, since it is all about their function and distribution and this can be modeled as part of the syntax, so don’t need to be looking for fine tests which tend to be confusing (even Samvelian doesn’t manage).
Recall that Samvelian analyzes Persian Ezafe as a *phrasal affix*. 
Persian Ezafe As as Phrasal Affix

- Recall that Samvelian analyzes Persian Ezafe as a *phrasal affix*.
- Should the Urdu Ezafe be also analyzed that way?
Recall that Samvelian analyzes Persian Ezafe as a *phrasal affix*.

Should the Urdu Ezafe be also analyzed that way?

Using the definition in ?, in principle *yes*. 
Phrasal Affixes According to Anderson

Anderson’s approach:

- Assumes a Lexicalist Hypothesis (?) — all clitics are introduced postlexically (NB: contra Samvelian).
- Assumes Prosodic Phonology (e.g., ????) — clitics which cannot be prosodic words on their own have to be incorporated into the prosodic word or prosodic phrase of a host.
- Sees *phrasal affixes* as a type of clitic which show restrictions as to which hosts they will appear with (these are Zwicky’s *special clitics*).
- *Phrasal Affixes* are the “morphology of phrases”.
- Both derivational (e.g., possibly Ezafe) and inflectional (e.g., determiner, possessor) type of Phrasal Affixes can be identified, just as in morphology.
Phrasal Affixes and Historical Change

- We basically agree with Anderson’s approach.
- Except for one point: we see the “morphology of phrases” (=shape of phrase) as being essentially the domain of syntax (contra Anderson).
- Anderson argues strongly against a syntactic approach to clitics.
- However, our approach does not presuppose movement for positioning of clitics, which is what he objects to most (and proposes an OT analysis in terms of alignment instead).
- Our analysis of Urdu Ezafe would therefore seem to be entirely in line with ?. 
Postlexical vs. Word Level Analysis

- Samvelian analyzes Persian Ezafe as a Phrasal Affix.
- However, she explicitly rejects Anderson’s view that phrasal affixes come into play postlexically.
- Samvelian instead analyzes Persian Ezafe as part of word level morphology (following proposals by ??).
- Why?
Samvelian’s Reasoning

Samvelian has three reasons for her design decision:

1. Inserting items postlexically is supposed to leave them with no access to word-level properties of their hosts (which they need in order to determine whether they are compatible with them).

2. Persian Ezafe shows Haplology with two other (types of) clitics: personal pronouns and the definite marker =i (cf. ?).

3. Anderson assumes a clear cut distinction word-level affixes and phrasal affixes, but Persian Ezafe would seem to be both at once.
Reason 1: Access to Word Level Properties in the Syntax

- Samvelian’s Reason 1 is the product of fallacious reasoning about the architecture of grammars.
- MORE HERE
Reason 2: Haplology

- Miller (1992) proposes that haplology be used as a test for whether something should be dealt with in the lexicon as an affix or not (based on argumentation about the English possessive).
- Samvelian (2007:627) takes this to mean that one has to establish the following:
  
  (a) they are in complementary distribution (i.e., in competition) when adjoined to the right edge of the same constituent;
  
  (b) any sequence containing two or more of these enclitics is excluded, even when their scope is in the same constituent

- She looks at the distribution of the definite =i vs. personal pronoun clitics vs. ezafe and finds:
Reason 2: Haplology

However:

- As Samvelian also shows, \( =i \) and ezafe have overlapping functions (\( =i \) can take over the role of ezafe) — their cooccurrence could therefore be ruled out on functional grounds.
- Ezafe does not seem to attach to pronouns anyway and so this combination could therefore be prohibited for independent reasons.

(13)
\[
*\text{qahreman-e } [\text{r\^{a}nde } \text{\v{s}ode } \text{az } \text{mihan-\'{a}s}] \text{-e}
\]
\[
\text{hero-Ez drive.PastP become.PastP from homeland-Poss.3.Sg}
\]

‘the hero, (who is) driven away from his homeland’

(14)
\[
\text{qahreman-e } [\text{az } \text{mihan-\'{a}s } \text{r\^{a}nde } \text{\v{s}ode}] \text{-e}
\]
\[
\text{hero-Ez drive.PastP become.PastP from homeland-Poss.3.Sg}
\]
Reason 2: Haplology

- It is also not clear that being in complementary distribution or disallowing for a sequence of clitics speaks against a postlexical analysis.
  - Syntactic elements like indefinite vs. definite determiners or past vs. present tense auxiliaries tend to be in complementary distribution.
  - If one has clitics which fulfill a certain function or license modifiers (like the ezafe does), then it is not clear that they should be stackable with other clitics — this would depend on the syntactic construction and what the syntax of the language is able to license.

- Conclusion: The Haplology Criterion/Test does not work as a strong argument in favor of a lexical treatment.
Reason 3: Dual Nature of Ezafe?

- Persian Ezafe can attach to either a word (the head noun) or a phrase.
- Urdu Ezafe can attach to a word (the head noun) or a phrase (containing the head noun).
- If one assumes that the Ezafe has a certain position at phrase-structure, as per the Urdu analysis, then it simply prosodically attaches to the host to the left of it — no special rules need to be formulated and Ezafe is not any different than other phrasal affixes (or “special clitics”).
why one could do Persian Ezafe the Urdu way
Further Issues/Phenomena to Address

- EP clitics
- Kurdish ezafe
Thanks go to Tafseer Ahmed for help with the data.