



Center for Research in Urdu Language Processing National University of Computer and Emerging Sciences, Lahore Pakistan

## **Revision History:**

| Name          | Change Date  | Version | Description of Changes   |
|---------------|--------------|---------|--|
| Tafseer Ahmed | 5 Oct, 2004  | 4.1.0.1 | Created  |
| Sara Hussain  | 16 Dec, 2004 | 4.1.0.2 | comments added for<br>compatibility with grammar<br>file and better understanding            |
| Aasim Ali     | 29-Jun-2005  | 5.1.0.1 | Removed redundant rule for coordinate conjuntion of 2 <sup>nd</sup> & 3 <sup>rd</sup> Person |
| Aasim Ali     | 18-Jul-2005  | 5.1.0.2 | Removed constraint for<br>PERSon   |
| Aasim Ali     | 19-Jul-2005  | 5.1.0.3 | Changed ~AUR to YA   |
| Aasim Ali     | 28-Jul-05    | 5.1.0.4 | Removed redundant checks<br>on Form, Case, Num, Gend<br>etc.                                 |

# Rule ID: UGR124

**Rule Syntax:** Following is the constituent description of the rule. KPmain -> KP

KPmain -> KP [[comma KP]]\* coord\_conj KP

**Rule Functional Description:** Following are the functional specifications of the rule. KPmain -> KP:^=!; .

//AUR Conjuncted KPmain -> KP: ! \$ ^; [[comma:; KP: ! \$ ^;]]\* coord\_conj:^=!, !CONJ\_FORM =c AUR; KP: ! \$ ^; .

//YA Conjuncted KPmain -> KP: ! \$ ^; [[comma:; KP: ! \$ ^;]]\* coord\_conj:^=!, !CONJ\_FORM =c YA; KP: ! \$ ^; .

## **Frequency:**

Description: This rule represents the co-ordination of Case Phrase.

*c-structure*: It consists of a case phrase or a list of comma separated case phrases.

f-structure: A set of case phrases having co-ordination type and agreement features at the root.

**Examples:** میں یا وہ مجھے، تمہیں یا اسے تم نے اور عمران نے





## Center for Research in Urdu Language Processing National University of Computer and Emerging Sciences, Lahore Pakistan

Rule Status: Under Process

**Reference:** 

Related Rules: UGR002

Related POS: UPOS121, UPOS122

Replaces: -

Reason: -

Replaced by: -

**Analysis:** Following is the analysis of the rule.

*Analysis :* Case Phrase can be coordinated by using coordination conjunction. So a Case Phrase(KP) in sentence can be a single case phrase or a list of comma separated case phrase having a coordinating conjunction before last case phrase.

There are two issues in the coordination of KP. There is a favorable order of KPs depending upon their types. For example, مجهے، تمہیں یا اسے یا مجھے is more acceptable than تمہیں ، اسے یا مجھے . The natural order is that 1<sup>st</sup> person will be followed by 2<sup>nd</sup> Person and that will be followed by third person.

The other issue is the features of coordinated KP. A case phrase(KP) has person, number, gender, form, case and respect features. The verbal phrase agrees with all these features(except case). So the features for coordinated KP as a whole should be known.

If coordination is made by using AUR, then it shows a collection which results in plural number. For example, میں اور تم آئے۔

Similarly, we can say that respect feature will be usual because usual repect is equivalent to plural in Urdu syntax. The Gender of AUR coordinated KP is masculine, because coordination of masculine KP and feminine KP results in masculine KP(that is default gender of Urdu.) For example,

Co-ordination of KP has a condition that KPs of same form and case can be coordinated. For example, following are not valid

\*میں نے اور اسے \*تم اور اس

Rules for these features are given in following table.

| Conjunction Type | Aur       | Others( e.g. ya) |
|------------------|-----------|------------------|
| Number           | Plural    | Of last KP       |
| Gender           | Masculine | Of last KP       |

EGD Urdu Localization Project - Grammar Rule Document





#### Center for Research in Urdu Language Processing National University of Computer and Emerging Sciences, Lahore Pakistan

| Person  | 1st then 2 <sup>nd</sup> then 3rd | Of last KP          |
|---------|-----------------------------------|---------------------|
| Form    | All KPs should have               | All KPs should have |
|         | same.                             | same.               |
| Respect | Usual                             | Of last KP          |
| Case    | All KPs should have               | All KPs should have |
|         | same.                             | same.               |

The rules for coordination like above are not present in grammar books. The rules are extracted from a small survey from native speakers. The rules for AUR coordination seems to be well established among these people. But for others, the rules are confusing and more than one rules can exist. For example,

لڑکا یا لڑکی آئی۔ لڑکی یا لڑکے آئے۔

In both of the above examples, verb is agreeing with the last KP. But there are examples that are confusing fro the speakers. For example, one can predict both 'aaya' and 'aayen' in following sentence. الأخيان يا لأكا آيا۔

The sentence is also confusing, because it seems that Urdu prefers singular phrases before plural ones. So a more natural version of above sentence in Urdu will be لرحا يا لرحكا يا لرحكا يا لرحكا يا الركيان

The sentence also show agreement of verb with last KP.

Constraint for person has been removed keeping in view to prefer the order of the original sentence in English.

All checks related to Form, Case, Num, Gend etc have been removed because they will be coming from English and will fill the Urdu / Mapping f-structure accordingly.

Result: We decided on analysis.

**Future Work:** The Gender rule for AUR conjunction is an over-simplification, that says that gender of aur conjuncted phrase is always Masculine. The rule is true for masc-masc and fem-masc conjunction, but it is not true for fem-fem conjunction. New rules should be written to successfully parse sentences like للأكيال اور استانيان آئيں۔