



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	3 rd Nov, 04	0.1	Created
Kiran Khurshid	9 th Dec, 04	0.2	Added negative determiner
Zunaira Malik	9 th Aug, 05	0.3	New Release- removal of DETP

Rule ID: EGR258

Rule Syntax: Following is the constituent description of the rule.

DET -> [neg_det | art | gen_pro | dem | NPgen]

Rule Functional Description: Following are the functional specifications of the rule:

DET -> [neg_det: ^!=!, ^DTYPE = neg_det, ^DEF = NEG, ^NNUM = {SG, PL}; | art: ^!=!, ^DTYPE = art, ^NNUM = !NUM; | gen_pro: ^ GEN_PRO !=!, ^NNUM = {SG, PL}, ^DTYPE = gen_pro, ^DEF = POS; | dem: ^ DEM !=!, ^DTYPE = dem, ^NNUM = ! NUM; | NPgen: ^ GENITIVE = !, ^DTYPE = genitive, ^DEF = POS, ^NNUM = {SG, PL};].

Frequency: -

Description: This rule shows the functional and constituent structure of determiner phrase. The intuition for keeping the determiner as specifier is that they serve to “specify” the head noun rather than simply modify it like adj. [1, p. 101]. The determiner goes as DET feature structure in the f-structure which becomes the member of SPEC.

c-structure: These productions show that DETP (determiner phrase) can have DET (determiner). DET can be article (art), genitive pronoun (gen_pro), demonstrative (dem) or a possessive NP (NPgen).

f-structure: Everything in DET goes as a SPEC to the mother node i.e. DETP. DETP on the whole also becomes SPEC of NP. The feature DEF is set for each determiner, so that it can be unified to the DEF feature coming from the PREDET rule. This helps in allowing only definite determiners with pre-determiners (like quantifiers, multipliers and fractions).

Examples:

In the following examples the underlined part is the determiner phrase:

- 1) I have no book (negative determiner)
- 2) She earns three times my salary. (genitive pronoun)
- 3) Many of the books are missing. (article)
- 4) This room is rather a mess, isn't it? (demonstrative)
- 5) The girl's books are torn. (NPgen)



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

Rule Status: Active

Reference:

- [1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"
- [2] "A Lexicalized Tree Adjoining Grammar for English", Institute for Research in Cognitive Science, University of Pennsylvania
- [3] <http://webster.commnet.edu/grammar/determiners/determiners.htm>
- [4] B. A. Hockey and Heather Mateyak, "Determining Determiner Sequencing: A Syntactic Analysis for English" University of Pennsylvania.

Related Rules: EGR138

Related POS: EPOS108, EPOS107, EPOS115, EPOS121

Replaces: EGR130

Reason: New Release- removal of DETP

Replaced by: -

Reason: -



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

Analysis: Following is the in-depth analysis of the rule.

Analysis 1:

The determiner phrase is expected to come in the pre-nominal part of NPcommon in English. It can be split into two parts: the pre-determiner and the determiner.

As is obvious from their name, the pre-determiners occur prior to other determiners. They may contain a multiplier, a fraction, an adverb, a pre-determiner quantifier [3] or a pre-determiner number. For each of these, separate rules apply.

Analysis 2:

According to another analysis [2, p. 178] [4], determiner should be placed as the head of a phrase, whereas the noun should occur as its complement. This alternative DP-hypothesis considers the specifier as the head of the constituents NP, and that is in keeping with the requirements of X-bar theory [1, p. 101]. This theory states that sequences of determiners, as in the NPs *all her dogs* or *those five dogs*, are derived by multiple adjunctions of the determiner tree, with each tree anchored by one of the determiners in the sequence. The order in which the determiner trees can adjoin is controlled by the *features*. This analysis has identified nine features which are sufficient to order the determiners [4]. These features determine the ordering of multiple determiners in a phrase and their well-formedness (number etc.) with respect to each other.

Result: Analysis 2 was rejected because it does not give any hierarchy to the sub categories of the determiners. They may be placed in order according to this analysis, but the levels in which they can be subdivided according to intuition are not considered. Reason for this inflexible use of determiners is that the whole analysis is based on X-bar theory. This analysis also ignores the pre-determiners, hence making the determiners limited. Extending this analysis is a tedious task as all the nine features for well-formedness need to be understood fully and then have to be mapped for pre-determiners.

Finally, intuitively speaking, determiner should not be placed as the head of the noun phrase because it is “specifying” the head noun.

All the modern theories also negate the DP hypothesis [1, 101]. Therefore, Analysis 1 was finalized.

Future Work: Following issues need to be addressed, and constraints to be included respectively, after further analysis:

- 1) The list of intensifiers which can occur before determiners need to be analyzed.
- 2) The analysis of WH pronoun as pre-determiner is yet to be done
 - *Example:*
 - What an idiot he is.

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	28 th July, 04	0.1	Created
Zunaira Malik	20 th May, 05	0.2	Addition of mono and di-trans VPs

Rule ID: EGR201

Rule Syntax: Following is the constituent description of the rule.

VP-> [VPpred_main1|VPpred_main2] [(VP_mtrans) | (VP_dtrans)] VPpredlink (VPadjunct)

Rule Functional Description: Following are the functional specifications of the rules.

VP -> [VPpred_main1: ^ = !;|VPpred_main2: ^ = !;] [(VP_mtrans: ^=!;) | (VP_dtrans: ^=!;)]

VPpredlink: ^=!; (VPadjunct: ^=!;).

Frequency: -

Description: This production gives the detail of the VP production having a copula construction.

c-structure: When main verb of the form 'be' is used in a VP, it has a different analysis as compared to other main verbs because this verb shows linking of subject with another argument, which is giving information about the subject itself. VP consists of main verb which is followed by ADJP (adjectival phrase), NPmain or PPnmain and then it is followed by VPadjunct.

f-structure: The linking verb *be* can have APJP, NP or PP as its argument. All these arguments become PREDELINK in f-structure. _VERB_FORM (verb form) of verb in VPpred_main is always 'be'. VPpred_main contains sequence of auxiliaries followed by main verb having the form of 'be'. It is to be noted that PREDLINK analysis is always in the ACTIVE form of sentence and maximum one modal and two auxiliaries can precede the main verb. The CASE in NPmain is constrained to be ACC (accusative) and the NUM to be the same as that of SUBJ to avoid entries like **They are a good boy*

Examples:

- 1) I am good. (ADJP as PREDLINK)
- 2) I am a good girl. (NP as PREDLINK)
- 3) Cat is on the table. (PP as PREDLINK)
- 4) It is nice to meet you. (ADJP as PREDLINK, infinitive clause as adjectival complementation)
- 5) It is pleasant in the morning. (ADJP as PREDLINK, PP as ADJUNCT)
- 6) She painted the door green. (Object complementation)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR200, EGR202, EGR203, EGR204

Related POS: -

Replaces: EGR126

Reason: addition of di and mono-trans and VPadjunct

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule:

Analysis: Predicative construction involves a linking or *copular* verb which has a subject and another argument, as in the example below:

Example: i) The beacon is on the roof.

ii) The tractor is red.

The post-verbal argument can be of a number of categories, e.g., NP, PP, AP etc. Due to the semantic relationship between the subject and the phrase after linking verb, these verbs are given special sub-categorization frames. Traditionally, this has been done by having the post-verbal phrase 'be' an XCOMP whose subject is controlled by the linking verb's subject. However, a new analysis, termed the PREDLINK analysis, is used. Under both approaches, linking verbs may have their own c-structure category and their own VP rule which allows the post-verbal NP, AP, and PP to be assigned the appropriate grammatical function. [1, p. 69]

Result: The above analysis was finalized.

Future Work:

1) The sentence like *the chicken is cooked* will create ambiguity. It will be parsed having *is* as the main verb and *cooked* as an ADJP and will also be parsed through the production having *cooked* as the main verb. This kind of ambiguity will be catered in the later stage using some statistical method.

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	16 th September, 04	0.1	Created
Kiran Khurshid	9 th December, 04	0.2	Added negative particle
Zunaira Malik	20 th May, 05	0.3	Addition in Release 5.0
Zunaira Malik	10 th July, 05	0.4	Removal of VERB_FORM = 'be' constraint to allow verbs with <SUBJ,PREDLINK> subcat to work

Rule ID: EGR202

Rule Syntax: Following is the constituent description of the rule.

VPpred_main1 -> v (neg)

VPpred_main2 -> HelpVP v

VPpred_main3 -> (neg) v

Rule Functional Description: Following are the functional specifications of the rules.

- 1) VPpred_main1 -> v: ^ = !, !_MORPH_FORM =c {PRES,PAST}, !VOICE =c ACTIVE; (neg: ^=!;).
- 2) VPpred_main2 -> HelpVP: ^ = !, ^TNS_ASP = !_MAUX TNS_ASP, !_MAUX VOICE =c ACTIVE; v: ^ = !, !VOICE =c ACTIVE, !_MORPH_FORM =c ^_MAUX _ALLOWED_FORM;.
- 3) VPpred_main3 -> (neg: ^=!;) v: ^ = !, !VOICE =c ACTIVE, ^HelpVP _ALLOWED_FORM = !_MORPH_FORM ;.

Frequency: -

Description: This production gives the detail of the main part of VP (auxiliaries and main verb) in a copular construction.

c-structure: main verb phrase in copular construction has specific constraints which are different from normal constructions, hence separate productions are made.

f-structure: _VERB_FORM (verb form) of verb in VPpred_main is always 'be'. VPpred_main contains sequence of auxiliaries followed by main verb having the form of 'be' or verbs with subcat frames containing the PREDLINK option. It is to be noted that PREDLINK analysis is always in the ACTIVE form of sentence. Production 1 and 2 are used for declarative sentences and production 1 and 3 are used for interrogative.

Examples:

- 1) I am good. (for production 1)
- 2) She is not good. (for production 1)
- 3) I have been being a good girl. (for production 2)
- 4) She has not been being a good girl. (for production 2)
- 5) Are you being a good girl? (for production 3)
- 6) Are you a good girl? (for production 3)

- 7) Are you not being a good girl? (for production 3)
- 8) He feels good (for production 1)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR102

Related POS: EPOS106, EPOS120

Replaces: EGR103

Reason: removal of VERB_FORM = 'be' constraint to allow <SUBJ,PREDLINK> option for verbs like "feel" and "seems" to run

Replaced by: -

Analysis: Following is the in-depth analysis of the rule:

Analysis 1: Three productions are used to represent a verb phrase for copular constructions. The first two productions given above are used for declarative sentences, while the third is used only for interrogative sentences.

The first production allows only the present or past form of the main verb "be" in a copular construction.

Example: i) I am a good girl.
ii) She was different.

The negation of the above examples is possible by introducing a negative particle "not" after the main verb.

Example: i) I am not a good girl.
ii) She was not different.

The second production allows auxiliaries and modals before the main verb "be".

Example: i) She may have been good.

Similarly, the negation of this sentence can also be done by introducing "not" within the helping verb phrase. Note that this production uses the HelpVP rule, which allows the negative particle in itself.

Example: i) She may not have been good.

Finally, the third production allows the main verb "be" alone, after the subject of the interrogative sentence.

Example: i) Why is she being good?

From the above example, we see that the auxiliary occurring before the subject needs to agree with the auxiliary after the subject. For this reason the _ALLOWED_FORM of the HelpVP structure of the interrogative sentence is unified with the _MORPH_FORM of the second auxiliary. Furthermore, the verb "be" allowed in interrogatives can be other than present and past morphological forms.

For the negation within the copular construction of interrogatives, negative particle "not" is allowed immediately before the main verb in the third production.

Example: i) Why is she not being good?

Result: We decided on the above analysis.

Future Work: -



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Zunaira Malik	20 th May, 05	0.2	Updated name from VPact_mtrans to VP_mtrans. Does not go to VPact_main anymore.

Rule ID: EGR203

Rule Syntax: Following is the constituent description of the rule.

VP_mtrans -> NPmain

Rule Functional Description: Following are the functional specifications of the rule.

- 1) VP_mtrans -> NPmain: ^OBJ = !, ^OBJ CASE = ACC;.
- 2) VP_mtrans -> NPmain: ^OBJ2 = !, ^OBJ2 CASE = DAT, ^_FLAG = INTERROGATIVE;.

Frequency: -

Description: This rule shows the mono-transitive form of active VP

c-structure: This VP production is for the verb which requires a direct object only i.e. the mono transitive verb.

f-structure: The structure information of NP that is subcategorized by VP as an OBJ, is passed to S. It is to be noted that the OBJ case is accusative for declarative sentences. Interrogative sentences use the same rule as interrogative is just a transformed version of declarative type of sentences. But this rule is actually used for the di-transitive VP of interrogative sentences. Structurally the NP occurring right after verb in interrogative becomes OBJ2 and hence takes dative case.

Examples:

- 1) He went home (for production 1)
- 2) He is scolding him (for production 1)
- 3) What(OBJ) has Ahmad(SUBJ) been giving Rashid(OBJ2). (for production 2)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR205

Related POS: -

Replaces: EGR123

Reason: Rule updated. VPact_mtrans changes to VP_mtrans. No longer goes to VPact_main.

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis1: In English, position is a good indicator of object identification since the object must follow the verb and be adjacent to it. In English case marker feature of NP is used to restrict nouns which can come at the object position.

Example:

(1) They saw the box yesterday.

(2) *They saw yesterday the box.

In Example (1), 'box' is the OBJ of verb 'saw'. On the other hand, example (2) is considered ill formed as adverb intervened between the verb and its object. [1, p. 48]

Result: The analysis discussed above has been finalized.

Future Work: Further functional annotations might be added in the later stage.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Zunaira Malik	20 th May, 05	0.2	Change of name from VPact_dtrans to VP_dtrans. Does not go to VPact_main anymore.

Rule ID: EGR204

Rule Syntax: Following is the constituent description of the rule.

VP_dtrans -> NPmain NPmain

Rule Functional Description: Following are the functional specifications of the rule.

VP_dtrans -> NPmain :^OBJ2 = !, ^ OBJ2 CASE = DAT; NPmain :^OBJ = !, ^OBJ CASE = ACC, ~[^OBJ PRED =c 'pro'];.

Frequency: -

Description: This rule shows the ditransitive verb production of English Grammar.

c-structure: This VP production is for the verb in active form which requires a direct and an indirect object. Verb sub-categorization frame determines the transitivity of verb.

f-structure: The structure information of first NP is passed as OBJ2 structure and of second NP is passed as OBJ structure to S. It is to be noted that the direct object which comes in the second NP position cannot be pronoun, hence this is constrained.

Examples:

- 1) She [gave Ahmad an apple]_{VP}.
- 2) * I gave a book him.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR205

Related POS: -

Replaces: EGR124

Reason: New release. Change of name from VPact_dtrans to VP_dtrans. Does not go to the rule VPact_main anymore.

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis : In English, position is a major indicator of direct and indirect object identification. If there are two NPs following a verb, then the immediate one is considered as indirect object. Passivization test is not really good when it comes to identifying direct and indirect object as both the objects has the ability to passivize. In English case marker feature of noun is used to restrict nouns which can come at the direct and indirect object position.

Example:

(1) She gave him the box.

In the above example, him is OBJ2 and box is OBJ of verb 'give'. [1, p.49]

Result: The above analysis has been finalized.

Future Work: Further functional annotations might be added in the later stage.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	16 th September 2004	0.1	Created
Kiran Khurshid	22 nd December 2004	0.2	Introduced coordinate correlatives
Zunaira Malik	20 th May, 05	0.3	Addition in Release 5.0
Zunaira Malik	13 th June, 05	0.4	Addition of coord_conj OR rule and correlative without constraint pertaining to AND, addition of POSTNOMP
Zunaira Malik	15 th July, 05	0.5	Addition of "BUT" NPmain production with neg and addition of (neg) to "AND" production
Zunaira Malik	27 th July, 05	0.6	Change of NUM = SG,PL in NPmain BUT and AND-explanation in future work

Rule ID: EGR205

Rule Syntax:

NPmain -> NP [comma NP]* coord_conj NP (POSTNOMP).

NPmain -> correlative NP coord_conj NP (POSTNOMP).

NPmain -> NP

NPmain -> NP [comma NP]* coord_conj (POSTNOMP).

NPmain -> correlative NP coord_conj NP.

NPmain -> (neg) NP [comma NP]* coord_conj (neg) NP (POSTNOMP).

Rule Functional Description:

- 1) NPmain -> (neg: ^=!;)NP: ! \$ ^; [comma:; NP: ! \$ ^;]* coord_conj: ^ CONJ_FORM = !CONJ_FORM, !CONJ_FORM =c 'AND', ^ NUM = {SG,PL}; (neg: ^=!;) NP: ! \$ ^, ^ PERS = ! PERS; (POSTNOMP: ^=!;).
- 2) NPmain -> correlative: ^=!; NP: ! \$ ^; coord_conj: ^ CONJ_FORM = !CONJ_FORM, !CONJ_FORM =c 'AND', ^CONJ_FORM = ^_ALLOWED_CONJ, ^ NUM = PL; NP: ! \$ ^, ^ PERS = ! PERS; (POSTNOMP: ^=!;).
- 3) NPmain -> NP: ^=!; .
- 4) NPmain -> NP: ! \$ ^; [comma:; NP: ! \$ ^;]* coord_conj: ^ CONJ_FORM = !CONJ_FORM, !CONJ_FORM =c 'OR'; NP: ! \$ ^, ^ NUM = ! NUM, ^ PERS = ! PERS; (POSTNOMP: ^=!;).
- 5) NPmain -> correlative: ^=!; NP: ! \$ ^; coord_conj: ^ CONJ_FORM = !CONJ_FORM, ^CONJ_FORM = ^_ALLOWED_CONJ; NP: ! \$ ^, ^ NUM = ! NUM, ^ PERS = ! PERS; .
- 6) NPmain -> (neg: ^=!;) NP: ! \$ ^; [comma:; NP: ! \$ ^;]* coord_conj: ^ CONJ_FORM = !CONJ_FORM, !CONJ_FORM =c 'BUT', ^ NUM = {SG,PL}; (neg: ^=!;) NP: ! \$ ^, ^ PERS = ! PERS; (POSTNOMP: ^=!;).



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



Frequency: -

Description: This rule is used for the combination of noun phrase using the coordinate conjunctions “and” and “or” and the correlative conjunctions either...or, neither...nor and both...and.

c-structure: A main noun phrase can either contain only one noun phrase or more than one noun combined by coordinate conjunctions.

f-structure: Production 1 represents the combination of noun phrases with “and”. Production 2 combines noun phrases with both...and. Production 3 shows the combination of noun phrases with “or” and production 4 shows combines noun phrases using either...or or neither...nor. A single noun phrase is handled by Production 5. The coordinate conjunction is identified by the value of CONJ_FORM. Production 7 relates to the POSTNOMP with “in the room” as an ADJUNCT for “girl and boy”.

Example:

- 1) Girl and boy are going to school. (Production 1)
- 2) Both the girl and boy are working. (Production 2)
- 3) Ahmad or Aslam is doing the work. (Production 3)
- 4) Either Ahmad or Aslam is doing the work. (Production 4)
- 5) Neither Ahmad nor Aslam is doing the work. (Production 4)
- 6) A good girl is working. (Production 5)
- 7) Not cats but dogs are good (Production 6)
- 8) I have a cat but not a dog (Production 6)
- 9) She is kiran and not Shanza (Production 1)
- 10) She is not a doctor and not a patient. (Production 1)

Rule Status: Active

Reference:

- (1) Miriam Butt, Tracy Holloway King, “A Grammar Writer’s Cookbook”
- (2) Quirk et al, “A Comprehensive Grammar of the English Language”

Related Rules: EGR118

Related POS: EPOS017, EPOS122

Replaces: - EGR111

Reason: -Addition in Release 5.0

Replaced by: -



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

Analysis:

Analysis 1: The rules for subject phrases or clauses which are coordinated with (either....) "or" may be understood by the following examples:

- Example:* i) Either the Mayor or her deputy (is/*are) bound to come.
ii) What I say or what I think (is/*are) no business of yours.
iii) Either the strikers or the bosses (*has/have) misunderstood the claim.
iv) Either your brakes or your eyesight (is/?are) at fault.
v) Either your eyesight or your brakes (are/?is) at fault [2, p 762].

Here "?" means that native speakers are not sure about these constructions.

The pairs *both...and*, *either...or*, and *neither...nor* are termed CORRELATIVES, composed of an endorsing item and a coordinator. [2, p 936] (See POS document for "correlative" for more details)

Analysis 2: The c- and f-structures for NP coordination remain basically identical to those of same category coordination. The simplified rule for NP coordination is shown below. As with basic same category coordination, additions must be made for paired conjuncts and for more than two conjuncts [1, p 144].

$$\text{NPCOORD (CAT)} = \text{CAT CONJnp CAT}$$

NP coordination often involves number, person and gender mismatches between the individual conjuncts and the entire coordinated NP [p 143, 1]. The general rule in English, French and German for determining the number, gender and person of a coordinated NP is as follows:

- If any conjunct is plural, the entire NP is plural
- If any conjunct is masculine, the entire NP is masculine
- If any conjunct is first person, the entire NP is first person
- If there is no first person conjunct and any conjunct is second person, the entire NP is second person; otherwise the NP is third person [p 145, 1]

Result: We decided on analysis 2 for the formation of c- and f- structures. However, the agreement after the NP coordination was taken from analysis 1.

Future Work:

1. We need to cater "she is Kiran not Shanza" where not is a coordinate conjunction.
2. change of NUM to {SG, PL} in coord conj productions BUT and AND. There is a difficult situation which arises when negative marker (not) is used with these conjunctions in NPmain.

Situation 1: agreement with non-neg noun

Not cats but a dog is good.
Not a cat but dogs are good.
Dogs but not a cat are good.
A cat but not dogs is good.

If either one of the value of NUM is kept instead of both for e.g. PL is kept, then *She is Shanza and not Kiran* does not parse but, *They are Shanza and Kiran* (and all its neg versions) parse.

We need to see in detail what we have to do to model this situation. Commas create even more problems.

A cat and a dog is good – **agreement with the noun that occurs first**
Cats but not a dog are good.
A cat but not a dog is good.



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

But with commas

A cat, a dog and a lion are good. -Num is plural due to multiple nouns before comma

A cat, a dog and lions are good.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Zunaira Malik	20 th May, 05	0.2	goes to VPoblig and VPpassoblig and VPadjunct

Rule ID: EGR206

Rule Syntax: Following is the constituent description of the rule.

VP-> VPact_main (VPoblig) (VPadjunct)

VP-> VPpass (VPobligpass)(VPadjunct)

Rule Functional Description: Following are the functional specifications of the rules.

1) VP -> VPact_main: ^ = !, ^_MORPH_FORM = {PRES,PAST,BARE,PROG,PERF}; (VPoblig: ^ = !;) (VPadjunct: ^ = !;).

2) VP-> VPpass: ^ = !; (VPobligpass: ^ = !;)(VPadjunct: ^ = !;).

Frequency: -

Description: This production gives the detail of the Verb Phrase with all possible phrases contained in it.

c-structure: VP can have an active Verb Phrase (VPact) followed by a variety of Prepositional Phrases. Note that VPact can be intransitive or transitive having NPs contained in it. VP can have a passive Verb Phrase (VPpass) followed by a variety of prepositional phrases.

f-structure: Depending on the sub-categorization frame of VP, PP can become OBL (oblique), XCOMP (complement) or ADJUNCT of the mother node. In case of PPxcompmain, the subject of the mother node will become the subject of XCOMP of the mother node. In case of passive Verb Phrase (VPpass) PPs can be followed by the VP. The di-transitive verb in passive construction changes its sub-categorization frame to (SUBJ, OBJ, OBL). Hence the first PP should have a PFORM value as 'to', as given in production 3.

Examples:

Active Form:

- 1) She went to school. (PP as OBL)
- 2) He relied on this book. (PP with lack of semantic force and acting as OBJ of verb)
- 3) The driver wants to give Ahmad a book. (XCOMP)
- 4) The driver might have been trying to fix the problem. (XCOMP)
- 5) She slept in the morning. (PP as ADJUNCT)

Passive Form:

- 1) She was examined by the doctor. (PP as ADJUNCT)
- 2) The book is being given to him by me. (multiple PPs with first PP as an OBL and next PP as an ADJUNCT)



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

Rule Status: Active

Reference:

- [1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"
- [2] Correspondence with Miriam Butt, Universität Konstanz, Germany.

Related Rules: EGR200, EGR207, EGR211, EGR212

Related POS:-

Replaces: -

Reason: - goes to VPoblig and VPpassoblig

Replaced by: EGR019

Reason: More constraints added



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

Analysis: Following is the detailed analysis of the rule:

Analysis:

In the European languages, motion verbs such as "go" and "come" are generally taken to be intransitive, and then one would go for the ADJUNCT analysis in sentences like *He went to eat*. However, the verb 'find' as in *He finds a book to read* is a different matter. If one tried the XCOMP analysis here, one would immediately ask oneself what the SUBJ of the embedded XCOMP might be: according to Bresnan's theory of control, it should be the "book", however this is not the case. If one thinks about what this sentence is doing it seems that "to read" is in fact modifying the book. So, the "to read" should probably be analyzed as an ADJUNCT. [2]

In passive constructions, the demoted agent is often assumed to be an OBL with a restricted form, e.g., with the preposition *by* in English. This analysis has the advantage that it captures the fact that the OBL plays a special role with respect to the verb, i.e., that it is the logical subject. However, this analysis also has the disadvantage that the OBL is not obligatory since the demoted subject need not to be overtly expressed. This means that every passive occurring with an argument of the appropriate form will have two analyses, one in which the argument is the OBL and the one in which it is an adjunct. In our grammars passives do not sub-categorize for an OBL, meaning that the demoted agent is always an adjunct; the semantics can then determine whether a member of the adjunct set is the agent of the verb. [1, p. 58]

Result: The above analysis has been finalized for the rule. Intransitive verbs do not sub-categorize infinitival clause such as *to read* and hence becomes an ADJUNCT in f-structure. If verb sub-categorizes infinitival clause, it becomes XCOMP in f-structure.

Future Work:

- 1) Sentences like *He went for enjoying the concert*, is yet to be analyzed.

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	20 th May, 05	0.1	Created
Zunaira Malik	23 rd July, 05	0.2	Addition of optional PPn and NPmain within square brackets

Rule ID: EGR207

Rule Syntax: Following is the constituent description of the rule

VPoblig -> [(VP_mtrans) | (VP_dtrans)] [(adv) |(PPnmain) | (PPn_objmain)| (VPinfmain)]

VPobligpass -> [(PPn)| (NPmain)] [(adv)|(PPnmain) | (p) | (VPinfmain)]

Rule Functional Description: Following are the functional specifications of the rules.

VPoblig -> [(VP_mtrans: ^=!;) | (VP_dtrans: ^=!;)] [(adv: ^OBL = !;) |(PPnmain: ^OBL=!;) | (PPn_objmain: ^OBJ = !;)] (VPinfmain: ^XCOMP = !, ^XCOMP SUBJ = NULL, ^_FLAG = ^XCOMP _FLAG;)].

VPobligpass ->[(PPn: ^OBL = !, PFORM =c 'to;)](NPmain: ^OBJ = !;)] [(adv: ^OBL = !;)|(PPnmain: ^OBL=!;) | (p: ^ = !;) | (VPinfmain: ^XCOMP = !, ^XCOMP SUBJ PRED = 'pro', ^XCOMP SUBJ PRONTYPE = NULL, ^_FLAG = ^XCOMP _FLAG;)].

Frequency: -

Description: This production gives the detail of the VP adjunct production

c-structure: The production show the elements that VPs can take as obligatories in either active or passive form.

f-structure: in case of actives, production 1 chooses between a mono-trans or di-trans verb and then accordingly chooses the elements following the verb or required by its SUBCAT frame which maybe an ADV, a PP phrase, a PP phrase with an object (he relied on the book) or an XCOMP. For Passives, production 2 chooses either the “to” form preposition type phrase or an NP for cases like “The book was given me”. These maybe combined with ADVs, a PP phrase, a preposition or an XCOMP.

Examples:

He ate food quickly. (production 1)
He gave me the book in the morning. (production 1)
He relied on the book. (production 1)
He gave me the book to read. (production 1)
The book was given me. (production 2)
The book was given to me. (production 2)
He was induced to read. (production 2)
The book was given quickly. (production 2)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR203, EGR204, EGR208, EGR209, EGR213

Related POS: - EPOS200, EPOS201

Replaces:

Reason:

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule:

Analysis: Predicative construction involves a linking or *copular* verb which has a subject and another argument, as in the example below:

Example: i) The beacon is on the roof.

ii) The tractor is red.

The post-verbal argument can be of a number of categories, e.g., NP, PP, AP etc. Due to the semantic relationship between the subject and the phrase after linking verb, these verbs are given special sub-categorization frames. Traditionally, this has been done by having the post-verbal phrase 'be' an XCOMP whose subject is controlled by the linking verb's subject. However, a new analysis, termed the PREDLINK analysis, is used. Under both approaches, linking verbs may have their own c-structure category and their own VP rule which allows the post-verbal NP, AP, and PP to be assigned the appropriate grammatical function. [1, p. 69]

Result: The above analysis was finalized.

Future Work:

1) The sentence like *the chicken is cooked* will create ambiguity. It will be parsed having *is* as the main verb and *cooked* as an ADJP and will also be parsed through the production having *cooked* as the main verb. This kind of ambiguity will be catered in the later stage using some statistical method.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	16 th September 2004	0.1	Created
Kiran Khurshid	24 th December 2004	0.2	Added coordinate correlative
Zunaira Malik	20 th May, 2005	0.3	New release

Rule ID: EGR208

Rule Syntax:

PPnmain -> PPn [comma PPn]* coord_conj PPn
 PPnmain -> correlative PPn coord_conj PPn
 PPnmain -> PPn

PPn_objmain -> PPn_obj [comma PPn_obj]* coord_conj PPn_obj
 PPn_objmain -> correlative PPn_obj coord_conj PPn_obj
 PPn_objmain -> PPn_obj

Rule Functional Description:

- 1) PPnmain -> PPn: ! \$ ^, ^ PFORM = ! PFORM; [comma;; PPn: ! \$ ^;]* coord_conj: ^ CONJ_FORM = ! CONJ_FORM; PPn: ! \$ ^;.
- 2) PPnmain -> correlative: ^!=; PPn: ! \$ ^, ^ PFORM = ! PFORM; coord_conj: ^ CONJ_FORM = ! CONJ_FORM, ^CONJ_FORM = ^_ALLOWED_CONJ; PPn: ! \$ ^;.
- 3) PPnmain -> PPn: ^ = !;.
- 4) PPn_objmain -> PPn_obj: ! \$ ^; [comma;; PPn_obj: ! \$ ^;]* coord_conj: ^ CONJ_FORM = ! CONJ_FORM; PPn_obj: ! \$ ^;.
- 5) PPn_objmain -> correlative: ^!=; PPn_obj: ! \$ ^; coord_conj: ^ CONJ_FORM = ! CONJ_FORM, ^CONJ_FORM = ^_ALLOWED_CONJ; PPn_obj: ! \$ ^;.
- 6) PPn_objmain -> PPn_obj: ^ = !;.

Frequency: -

Description: This rule shows the coordination of prepositional phrase.

c-structure: These productions show coordinate conjunction and correlative conjunctions at PP level.

f-structure: Use of coord_conj (coordinate conjunction) introduces a set containing two or more conjunct elements along with some features like CONJ_FORM, NUM, etc. In all the PPnmain productions, PFORM has been placed at an upper level, but this has only been done for the PP phrase occurring before the conjunction. This is because there is a conflict problem which will arise if PFORMS for all the PPs separated by conjunctions are placed above. Moreover, right now the parser does not create a set of elements that are passed above with the same attribute name. We need to pass all PFORMS for all PPs above but will do it after enhancement of the parser. Right now, the parser only checks the PFORM for the PP phrase occurring before the conjunction and parses the rest with respect to the first PP, whether it is an ADJUNCT or an OBL, the subsequent PPs parse depending on the first PP.



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

These features are generalized to the whole set. The endorsing item of the correlative occurs before the first prepositional phrase and the coordinator occurring after it has to agree with it.

Example:

- 1) She came by air or by rail. (Production 1)
- 2) She came either by air or by rail. (Production 2)
- 3) She relied on the documentaries and on the book. (Production 4)
- 4) She relied either on the documentaries or on the book. (Production 5)

Rule Status: Active

Reference:

- (1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR213

Related POS: EPOS117, EPOS122

Replaces: EGR144

Reason: same rule copied for new release

Replaced by: -



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

Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT} ([\text{COMMA CAT}] + (\text{COMMA})) \text{ CONJ CAT}$$

Here, the CAT (category) is PP (preposition phrase).

The prepositional phrase can combine using coordinate conjunctions. This is shown in the examples given below:

Example: i) I put the toffees in this box and in that box.

ii) Did you put the toffees in this box or in that box?

iii) The toffees are in the box yet on the table.

However, there are other kinds of prepositional phrases too, which are not adjuncts, but are objects of the verb. These are named PPn_obj. The coordination of these prepositional phrases is handled by separate rules. The following example represents this:

Example: i) She relies on the book and on his instructions.

Coordination of PP might be possible using correlative conjunctions. In that case there is an endorsing item at the beginning of the phrase which takes only a particular coordinator with it. For more information refer to the POS document for correlative.

Example: i) I put the toffees either in this box or in that box.

ii) I put the toffees neither in this box nor in that box.

iii) I put the toffees both in this box and in that box.

iv) She relies neither on the book nor on his instructions.

v) She relies both on the book and on his instructions.

vi) She relies either on the book or on his instructions.

Result: We decided on the above analysis.

Future Work:

- Negation is not yet handled. Hence, the following cases will be dealt with later:
 - *Example:* i) She came by rail not by air.
- Need to add PFORM passing notation in all PPmain productions separated by conjunctions. Need to add them to a single set of PFORMs at the conjunction set level. Right now, the parser does not allow the formation of a PFORM set, containing multiple PFORMS from the lower set of OBL or ADJUNCTs.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	16 th September 2004	0.1	Created
Kiran Khurshid	24 th December 2004	0.2	Added correlative conjunctions
Zunaira Malik	20 th May, 2005	0.3	New Release

Rule ID: EGR209

Rule Syntax:

VPinfmain -> VPinf [comma VPinf]* coord_conj VPinf

VPinfmain -> correlative VPinf coord_conj VPinf

VPinfmain -> VPinf

Rule Functional Description:

- 1) VPinfmain -> VPinf: ! \$ ^; [comma; VPinf: ! \$ ^;]* coord_conj: ^ CONJ_FORM = ! CONJ_FORM; VPinf: ! \$ ^;.
- 2) VPinfmain -> correlative: ^=!; VPinf: ! \$ ^; coord_conj: ^ CONJ_FORM = ! CONJ_FORM, ^CONJ_FORM = ^_ALLOWED_CONJ; VPinf: ! \$ ^;.
- 3) VPinfmain -> VPinf: ^ = !;.

Frequency: -

Description: This rule combines infinitive verb phrases through coordinate conjunctions and correlatives.

c-structure: An infinitive verb phrase can combine with another infinitive verb phrase using a coordinate conjunction and correlatives.

f-structure: The infinitive verb phrases occurring before the coordinate conjunction can be more than one, separated by a comma in between. However, in the case of correlatives no comma is currently allowed.

Example:

1. She wants to run or to drive the tractor. (Production 1)
2. She wants either to run or to drive the tractor. (Production 2)
3. She wants neither to run nor to drive the tractor. (Production 2)
4. She wants to run and to get to school. (Production 1)
5. She wants both to run and to get to school. (Production 2)
6. She doesn't wants to run but to drive the tractor. (Production 1)



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

Rule Status: Active

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR106

Related POS: EPOS117, EPOS122

Replaces: EGR145

Reason: same rule copied for new release

Replaced by: -



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

Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT ([COMMA CAT]+ (COMMA)) CONJ CAT}$$

Here, the CAT (category) is VPinf (infinitive verb phrase). This rule combines infinitive verb phrases with coordinate conjunctions.

Example: i) She wants to hide or to leave unnoticed.
ii) She wants to sit and to eat slowly.
iii) She doesn't want to run but to drive home.

Coordination of VPinf might be possible using correlative conjunctions. In that case there is an endorsing item at the beginning of the phrase which takes only a particular coordinator with it. For more information refer to the POS document for correlative.

Example: i) She wants either to run or to drive the tractor.
ii) She wants neither to run nor to drive the tractor.
iii) She wants both to run and to get to school.

Result: We decided on the above analysis.

Future Work:



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	8 th Feb 05	0.1	Created
Zunaira Malik	20 th May, 05	0.2	New Release
Zunaira Malik	13 th June, 05	0.3	Addition of ADVP before VP

Rule ID: EGR210

Rule Syntax: Following is the constituent description of the rule.
VPparticiple -> (ADVP) VP.

Rule Functional Description: Following are the functional specifications of the rule.

VPparticiple -> (ADVP: ! \$ ^ADJUNCT ADV, ! ADV_TYPE =c S_MOD;) VP: ^ = !, ^SUBJ = NULL, !_FLAG =c PARTICIPLE ;.

Frequency: -

Description: This rule shows the nonfinite participle verb phrases.

c-structure: A nonfinite participle verb phrase is an active or a passive verb phrase with single verb (in –ed or –ing form).

f-structure: A flag is passed on from the VPact or VPpass if a verb occurs without an auxiliary in –ing or –ed (passive) form. An adverb phrase is also passed as a set if the ADV_TYPE and the S_MOD do not match.

Examples:

- 1) The letter written by me is placed on the table. (VPparticiple: -ed participle)
- 2) The girl writing the letter has left. (VPparticiple: -ing participle)
- 3) Eating a hearty breakfast, we prepared for our long journey.
- 4) I saw them shooting at him.
- 5) I saw the tower climbed by a student
- 6) I saw the tower being climbed by a student (rare, not modeled yet)
- 7) Her aunt having left the room, I asked Ann for some personal help (rare, not modeled yet)
- 8) Four car bombs (accurately) targeting Iraqi security forces killed 13 people.

Rule Status: Active

Reference:

- (1) Quirk et al, "A Comprehensive Grammar of the English Language"
- (2) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules:

Related POS: -

Replaces: -

Reason: - addition of optional ADVP before VP

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis1: The nonfinite clause may be with or without a subject. The classes of nonfinite verb phrase serve to distinguish four structural classes of nonfinite verb clauses. [p. 993, 1] This rule is used to represent the participle nonfinite verb clauses:

- 1) –ing participle [p. 993, 1]:
without subject: Leaving the room, he tripped over the mat.
With subject: Her aunt having left the room, I asked Ann for some personal help.
- 2) –ed participle [p. 993, 1]:
without subject: Covered with confusion, they apologized abjectly.
with subject: The discussion completed, the chairman adjourned the meeting for half an hour.

Passive and present/progressive participles can be used as clausal adverbials with or without subordinating conjunctions, as [p. 36, 2]:

Example: i) The driver caught sight of the dog, turning the corner.
ii) Turning the wheel, press the break gently.

Again, as with infinitival adjuncts, the nonovert PRO subject of the participial clause is introduced by a rule in the c-structure. This then accounts for the instances in which the subject of the participial is not necessarily related to that of the matrix clause, as in example i). [p. 36, 2]

Result: -

Future Work:

- The subject of the participle verb phrase will be assigned later.
- Passive participle does not work properly yet. Debug later.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Kiran Khurshid	16 th Feb, 05	0.2	Added VPpass for participles
Zunaira Malik	20 th May, 05	0.3	New Release. Removal of VPact

Rule ID: EGR211

Rule Syntax: Following is the constituent description of the rule.

VPpass -> HelpVP (ADVP) v (ADVP)

VPpass -> (ADVP) v (ADVP)

Rule Functional Description: Following are the functional specifications of the rule.

VPpass -> HelpVP: ^=!, !_MAUX VOICE =c PASSIVE, ^TNS_ASP = !_MAUX TNS_ASP;
(ADVP:! \$ ^ADJUNCT ADV,!ADV_TYPE =c V_MOD ;) v:^=!, !VOICE =c PASSIVE; (ADVP:! \$
^ADJUNCT ADV,!ADV_TYPE =c V_MOD ;).

VPpass -> (ADVP:! \$ ^ADJUNCT ADV,!ADV_TYPE =c V_MOD ;) v:^=!, !VOICE =c PASSIVE, ^_FLAG =
PARTICIPLE; (ADVP:! \$ ^ADJUNCT ADV,!ADV_TYPE =c V_MOD;).

Frequency: -

Description: This rule shows the VP active and passive form production of English Grammar.

c-structure: VP in active can be intransitive, transitive or ditransitive.

VP in passive form is always transitive, hence is not further sub categorized into intransitive and ditransitive forms. VP passive production can have a verb (v) which is preceded by one or many helping verbs (modal and auxiliaries). It is to be noted that the VP in passive form is always preceded by a Helping Verb phrase. However, if it's not, it is supposed to be a passive participle.

f-structure: In VPact production, the whole structure of its daughter nodes is passed to the mother node.

In VPpass production, everything in helping verb and main verb is passed to the mother node. A passive voice constraint is added to block the active form of verb to pass through this production. ADVP can optionally modify the verb.

If the passive verb occurs alone it is a passive participle verb phrase, hence a flag named _FLAG is set to the value PARTICIPLE which is later checked in the rule of VPparticiple, in order to parse the sentence. [3, Pg. 153]

Examples:

- 1) He has been examined carefully. (PASSIVE)
- 2) The letter written by me is placed on the table. (passive participle)



Rule Status: Active

Reference:

[1] James Allen, "Natural Language Understanding"

[2] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

[3] Quirk et al, "A Comprehensive Grammar of the English Language"

Related Rules: EGR214, EGR215

Related POS: EPOS106

Replaces: EGR002

Reason: some constraints added.

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule.

Analysis 1:

In the case of English and French, all transitive verbs which have an OBJ as the second argument may be passivized. Here (and generally passivization involved demotion of a SUBJ to either an unexpressed argument or an ADJUNCT or OBL. Within our grammars, passivization is accomplished by means of a lexical rule as:
PASS (SCHEMATA) = SCHEMATA

(^OBJ) -> (^SUBJ)

(^SUBJ) -> (NULL)

(^PASSIVE) =c +

Example:

[Doctor]_{SUBJ} is examining [her]_{OBJ}.

[She]_{SUBJ} is being examined.

The lexical rule is called by verbs that can passivize, e.g., by transitive verbs. The lexical rule has the OBJ become the SUBJ and the old SUBJ become NULL. In addition, it requires passive form of the verb to be used. [2, p. 58]

Analysis 2:

In the passive form the noun phrase that would easily be in the object position is used in the subject position. Tense is still carried by the initial verb in the verb group. Also, even though the first noun phrase semantically seems to be the object of the verb in passive sentence, it is syntactically the subject. This can be seen by checking the pronoun forms. For example, *I was hit* is correct, not **Me was hit*. Furthermore, the tense and number agreement is between the verb and the syntactic subject. Thus you say *I was hit by them*, not **I were hit by them*. [1, p. 30]

Result: Analysis 1 and 2 approach is used in conjunction. Currently the parser does not provide the functionality to handle passive schemata through lexical rules.

Future Work: Further functional annotations might be added in the later stage. The rule might get revised to be modeled through analysis 1.

"Doctor having examined a patient, gave the medicine" type constructions are called active participles and will be modeled later on.

[3, Pg.153]

"Patient being examined by the doctor, died immediately" are called passive participles and will also be modeled later on. [3, Pg.153]



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Kiran Khurshid	8 th Dec, 04	0.2	Added rule for clausal negation
Zunaira Malik	20 th May, 05	0.3	New Release
Zunaira Malik	23 rd June, 05	0.3	Addition of coord-conj and ADVP
Zunaira Malik	7 th July, 05	0.4	Addition of coordconj with ADVs along with V_PREMOD and POSTMOD

Rule ID: EGR212

Rule Syntax: Following is the constituent description of the rule.

VPact_main -> [HelpVP (ADVPmain_PREMOD) v (ADVPmain_POSTMOD)

VPact_main -> (ADVP) v (ADVP)

VPact_main -> aux neg v (ADVP)

Rule Functional Description: Following are the functional specifications of the rule.

- 1) VPact_main -> HelpVP: ^ = !, ^_MAUX VOICE =c ACTIVE, ^TNS_ASP = !_MAUX TNS_ASP; (ADVPmain_PREMOD: ! \$ ^); v: ^ = !, !_MORPH_FORM =c ^_MAUX_ALLOWED_FORM, !VOICE =c ACTIVE; (ADVPmain_POSTMOD: ! \$ ^);
- 2) VPact_main -> (ADVP: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;) v: ^ = !, ^_AUX1_MORPH_FORM = !_MORPH_FORM, ^_MORPH_FORM = {PRES, PAST, TO_INF}, !VOICE =c ACTIVE, ^_FLAG = DECL; (ADVP: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;)
- 3) VPact_main -> (adv: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;) v: ^ = !, ^_AUX1_MORPH_FORM = !_MORPH_FORM, !VOICE =c ACTIVE, ^_MORPH_FORM = {BARE, PERF, PROG}, ^_FLAG = INTERROGATIVE, ^HelpVP_ALLOWED_FORM = !_MORPH_FORM; (ADVP: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;)
- 4) VPact_main -> (ADVP: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;) v: ^ = !, ^_AUX1_MORPH_FORM = !_MORPH_FORM, !VOICE =c ACTIVE, ^_MORPH_FORM = {PRES, PAST}, ^_FLAG = INTERROGATIVE; (ADVP: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;)
- 5) VPact_main -> aux: ^TNS_ASP TENSE = !TNS_ASP TENSE, !_AUX_FORM =c 'do'; neg: ^ = !; v: ^ = !, ^_AUX1_MORPH_FORM = !_MORPH_FORM, ^_MORPH_FORM = BARE, !VOICE =c ACTIVE, ^_FLAG = DECL; (ADVP: ! \$ ^ADJUNCT, !ADV_TYPE =c V_MOD;)
- 6) VPact_main -> (ADVP: ! \$ ^ADJUNCT ADV, !ADV_TYPE =c V_MOD;) v: ^ = !, !_MORPH_FORM =c PROG, !VOICE =c ACTIVE, ^TNS_ASP TENSE = PRES, ^FLAG = HYPERLINK; (ADVP: ! \$ ^ADJUNCT ADV, !ADV_TYPE =c V_MOD;)

Frequency: -

Description: This rule shows the VP production in active form of English Grammar.

c-structure: This VP production is for the case where verb (v) is preceded by one or many helping verbs (modal and auxiliaries) or is a verb alone. This VP makes a sentence in active voice. It is to be noted that this production is



used by all forms of VP (intransitive, transitive and ditransitive). These productions are also used for both declarative and interrogative sentence types. Adverbial Phrase can optionally occur before or after main verb. Moreover, the adverbial phrases occurring before or after can contain multiple adverbs conjoined with the help of conjunctions like AND. The occurrences of these conjoined adverbs before or after the verb may cause different translations which is why they have been divided into V_PREMOD and V_POSTMOD to show whether the ADV phrases occurred before or after the verb.

f-structure: The information contained in HelpVPact (Active Helping VP) and verb is passed to the mother node. _MORPH_FORM feature contained in verb is for checking well-formedness. As the VP production is for active form hence VOICE of verb is checked. Production 1 is for main verb with helping verbs used for both interrogative and declarative sentences. It includes option ADVPmain_PREMOD and ADVPmain_POSTMOD productions to cater to ADVs separated by the conjunction AND occurring before and after a verb. This distinction has been made to help translation into Urdu which might be different for the PREMOD and POSTMOD forms. Production 2 is for a VP having only main verb and is used only for declarative sentences. Production 3 and Production 4 are for interrogative sentences having just main verb. The auxiliary of such productions occurs in the beginning. Production 5 represents the negative verb phrase which contains no helping verb phrase. In such a verb phrase the verb is in BARE form while the tense aspect is contained in the auxiliary "do" occurring before the negative particle "not".

Examples:

- 1) She is working. (for production 1)
- 2) What have you been writing (for production 1)
- 3) He slept. (for production 2)
- 4) She works. (for production 2)
- 5) What have you seen? (for production 4)
- 6) What are you writing? (for production 4)
- 7) She does not read slowly. (for production 5)
- 8) They do not work. (for production 5)
- 9) She is working quickly and neatly (for production 1 with coord_conj in V_POSTMOD)
- 10) She is finally working. (V_PREMOD production 1 without conjunction)
- 11) She is quickly and swiftly eating her soup. (V_PREMOD with conjunction)

Rule Status: Active

Reference: -

Related Rules: EGR214, EGR215, EGR235

Related POS: EPOS101, EPOS106, EPOS120

Replaces: EGR125

Reason: New release

Replaced by: -

Reason: -



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



Analysis: Please refer to the related rules and related POS documents (in references) for details.

Result: -

Future Work: -



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Zunaira Malik	20 th May, 05	0.2	New Release
Zunaira Malik	16 th July, 05	0.3	Addition of rule for correlative “from---to”
Shanza Nayyer	25 th July, 05	0.4	Addition of adv to PPn

Rule ID: EGR213

Rule Syntax: Following is the constituent description of the rule.

PPn -> (adv) p NPmain

PPn_obj -> p NPmain

PPn -> p NPmain p NP

Rule Functional Description: Following are the functional specifications of the rules.

- 1) PPn -> (adv: ^ADJUNCT ADV =!, !ADV_TYPE =c P_MOD, ~[!SEM_TYPE =c {COMPARITIVE, SUPERLATIVE}]); p: ^ = !; NPmain: ^OBJ = !, ^OBJ CASE = ACC;.
- 2) PPn_obj -> p: ^ PFROM= ! PFROM; NPmain: ^ = !, ^CASE = ACC;.
- 3) PPn -> p: ^ _ALLOWED_PFORM = ! _ALLOWED_PFORM , ^ CORR_CONJ = ! PFORM ; NPmain: ! \$ ^ ; p: ^CONJ_FORM = ! PFORM, ^CONJ_FORM = ^ _ALLOWED_PFORM; NP: ! \$ ^;.

Frequency: -

Description: This production gives the detail of the Preposition Phrase.

c-structure: Prepositional phrase has preposition as the head of PP followed by its dependent noun phrase.

f-structure: If preposition has the semantic value (as in, *put the box on the table*), p subcategorizes NP (production 1). Note that an adverb can also be added (with ADV_TYPE as P_MOD) to intensify the prepositional phrase. If preposition does not have the semantic value, NP becomes the OBJ of the verb and p merely acts like a connector (as in, *he relies on this book*) (production 2).

In case of production 3, the preposition from contained an ALLOWED PFORM constraint which asks for the preposition “to”. It is modeled so the F-structure looks like the F-structure for correlatives. Between “from” and “to” NP’s can exist.

Examples:

- 1) She puts the book on the table (for production 1)
- 2) She is sitting right behind you (for production 1 with adv)
- 3) She relied on him (for production 2)
- 4) She works from Sunday to Monday. (for production 3)



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

Rule Status: Active

Reference:

- [1] "A Grammar Writer's Cookbook", Miriam Butt, Tracy Holloway King.
- [2] Correspondence with Miriam Butt, Universität Konstanz, Germany.

Related Rules: EGR205

Related POS: EPOS016

Replaces: EGR119

Reason: PPxcomp analyzed separately in the rule of VPinf (EGR106), PCASE changed to PFORM

Replaced by: -

Reason: -



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

Analysis: Following is an in-depth analysis of the rule.

Analysis: Prepositions are usually subdivided into two major classes: semantic and non-semantic.

Semantic prepositions usually give rise to adjunct PPs:

Example: i) The book is on the table.

The most common type of PP involves a preposition which has a clear semantic content of its own, such as the locative *on*, *in*, *under* etc., the instrumental *with*, and the directional *into*, *onto* etc. These prepositions are endowed with a PRED value and a sub-categorization frame which indicates that preposition requires an object.

In certain other constructions, a particular preposition is required by the verb for its particular meanings:

Example: i) I referred to the book.

In these cases, the preposition makes no or very little semantic contribution of its own. As such these PPs are treated as arguments of the verb. [1, pp. 125 – 129]

For English it is safe to treat all PPs as obliques and to use the OBJ2 for 'the book' in *She gave him the book*. [2]

Refer to [1, p. 50] for details on Obliques:

Example: i) She gave the book to him.

Result: If verb sub-categorizes a PP, it becomes an OBL, else it becomes ADJUNCT in the f-structure.

Future Work:

-



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	27 th July, 04	0.1	Created
Kiran Khurshid	9 th Dec, 04	0.2	Added negation
Zunaira Malik	20 th May, 05	0.3	New Release

Rule ID: EGR214

Rule Syntax: Following is the constituent description of the rule.

HelpVP -> AuxP

HelpVP -> modal (neg)

HelpVP -> modal (neg) AuxP

Rule Functional Description: Following are the functional specifications of the rules.

- 1) HelpVP -> AuxP: ^_MAUX = !_MAUX;
- 2) HelpVP -> modal: ^_MAUX = !, ^MODAL_STRUCT = !MODAL_STRUCT, ^_MAUX VOICE = ACTIVE, ^_FLAG = {INTERROGATIVE, DECL};(neg: ^=!;).
- 3) HelpVP -> modal: ^ = !; (neg: ^=!;) AuxP: ^=!, !_AUX1 _MORPH_FORM =c ^_ALLOWED_FORM, ^_MAUX TNS_ASP = ^TNS_ASP;

Frequency: -

Description: This rule shows the functional and constituent structure of Helping Verbs. It gives the detail of the Helping Verb Phrase in active and passive form with all possible forms of auxiliaries.

c-structure: A Helping Verb Phrase can have an AuxP alone, a modal alone or a modal preceded by an AuxP. AuxP consists of at least one and at most three auxiliaries. For details see EGR036. The first production is for AuxP alone, second is for modal alone and third is having a modal and at least one auxiliary in a helping verb phrase.

f-structure: _MAUX is the garbage structure, used to check well-formedness of the helping verb phrase and in turn of the VPact_main. It contains all the relevant features, such as VOICE and TNS_ASP, which are required to make these checks. _MAUX is passed the AuxP and modals.

Examples:

- 1) She is / has been working hard. (For production 1)
- 2) A car may hit her. (For production 2)
- 2) A car may not hit her. (For production 2)
- 3) The boy may have been sleeping in the car. (For production 3)
- 3) The boy may not have been sleeping in the car. (For production 3)



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

Rule Status: Active

Reference: -

[1] Randolph Quirk et al, "A Comprehensive Grammar of English Language"

Related Rules: EGR216

Related POS: EPOS114, EPOS120

Replaces: EGR100

Reason: New Release

Replaced by: -

Reason: -



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

Analysis: Following is the in-depth analysis of the rule.

Analysis: Previously, four rules (EGR017, EGR020, EGR021 and EGR022) were being used to represent the auxiliaries and modals which can come before the main verb. Now this rule has been made to have a generic rule which represents all possible combinations of modals and auxiliaries.

The possible combinations of modals and auxiliaries are taken from [1, p. 159]. They are also listed below:

	ACTIVE	PASSIVE
Present:	Sees	Is seen
Past:	Saw	Was seen
Modal:	May see	May be seen
Perfective:	Has seen	Has been seen
Progressive:	Is seeing	Is being seen
Modal + Perfective:	May have seen	May have been seen
Modal + Progressive:	May be seeing	May be being seen
Perfective + Progressive:	Has been seeing	Has been being seen
Modal + Perfective + Progressive:	May have been seeing	May have been being seen

Using the above table, we can also devise the following combinations of modals and auxiliaries including negation:

	ACTIVE	PASSIVE
Present:	does not see	Is not seen
Past:	did not see	Was not seen
Modal:	May not see	May not be seen
Perfective:	Has not seen	Has not been seen
Progressive:	Is not seeing	Is not being seen
Modal + Perfective:	May not have seen	May not have been seen
Modal + Progressive:	May not be seeing	May not be being seen
Perfective + Progressive:	Has not been seeing	Has not been being seen
Modal + Perfective + Progressive:	May not have been seeing	May not have been being seen

Result: We decided on the above analysis.

Future Work:



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	6 th September, 04	0.1	Created
Shanza Nayyer	25 th October, 04	0.2	Fixed after discussion with Dr. Miriam Butt
Zunaira Malik	20 th May, 05	0.3	New Release

Rule ID: EGR215

Rule Syntax: Following is the constituent description of the rule.

ADVP -> (adv)⁰⁻² adv

Rule Functional Description: Following are the functional specifications of the rule.

ADVP -> [adv: ! \$ ^SPEC , !ADV_TYPE =c ADV_MOD;]#0#2 adv: ^=!,;

Frequency: -

Description: This rule represents the adverbial phrase.

c-structure: An adverbial phrase may contain an adverb modified by the adverb(s) itself.

f-structure: The whole structure of the head adverb is made part of the set of adverbs. If the head adverb is modified by the adverb itself, the modifier adverb is made part of set of specifiers. ADV_TYPE is introduced to make sure that the right kind of adverb is used, as adverb can be a verb, sentence, adjective or adverb modifier. Cause of the diversity of adverb, SEM_TYPE feature is introduced in the lexicon to convey the semantics of different adverbs. For the details of adverb type, see the analysis part of the document.

Examples:

- 1) She is playing pretty well. (adverb modifier)
- 2) This work is almost impossible. (adjective modifier)
- 3) She is playing really well. (verb modifier)
- 4) Surprisingly she reached home in time. (sentence modifier)

Rule Status: Active

Reference:

[1] Pam Peters, "The Cambridge Guide to English Language"

Related Rules: -

Related POS: EPOS102

Replaces: -EGR105

Reason: - New Release

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis:

Adverbs are the most varied class of English words, with a variety of syntactic roles. Following are the types of adverbs:

SUBJUNCT: (adjective, adverb): they typically modify other adverbs or adjectives. They moderate the force of various kinds of word. Many such as *really, relatively, too, very*, modify adjectives and other adverbs. Some such as *almost, quite, rather* can modify verbs as well. Subjuncts of both kinds have the effect of either softening or intensifying the words they modify, hence the two major groups:

(downtoners)	fairly	rather	somewhat
(intensifiers)	extremely	most	so

A special group of restrictive subjuncts serve to spotlight others and to narrow the focus of the sentence. They include adverbs such as *especially, even, only*.

CONJUNCT(sentence, clause): they forge a semantic link between a sentence and the one before it. They play a cohesive role between separate sentences, or clauses. They include words like *also, however, therefore*, and thus express logical relationships such as addition, contrast and causation.

DISJUNCT (clause, sentence, v(sometimes)): modifies whole clauses or sentences. They affect the interpretation of the whole clause or sentence, either as judgements of the likelihood of something happening (*maybe, possibly, surely, probably*); or as expressions of attitude towards the event (*fortunately, mercifully, regrettably, worryingly*).

They can be moved around within it:

Fortunately the letter got there in time.

The letter *fortunately* got there in time.

The letter got there in time *fortunately*.

They can also be used for emphasis.

ADJUNCT (sentence, verb) : adverbs which detail the circumstances of the verb are these days often called adjuncts, to indicate that they connect with the core of the clause without being part of it. They add detail to whatever action the verb itself describes. They may specify the time or place of the action, the manner in which it took place, or its extent.

Example

Time: tonight, tomorrow, soon, then

Place: abroad, upstairs

Manner: well, thoughtfully, energetically

Extent: largely, totally, partly, thoroughly, mostly

NEGATIVE: *not*, the negative adverb, is treated separately from other adverbs in modern English grammars. This is because of its affinity with negative words of other kinds, such as determiners and pronouns (neither, no, none). *Not* has wide-ranging powers within sentences, to modify a word(verb, adjective or another adverb), a phrase, or a whole clause.

COMPOUND ADVERBS: there are also compound adverbs, for example *downtown* and *indoors*. Many adverbs are phrases:

Straight away

In no way

To the bottom

A little bit

Without a care in the world

Adverbial ideas can be expressed through several kinds of clauses.



Adverbs also allows degrees of comparison. Those consisting of one syllable, e.g. *fast, hard, soon*, make their comparative and superlative forms with inflections in the same way as adjectives: *sooner, soonest etc.* Adverbs formed with *-ly* enlist the help of *more* and *most*, as in *more energetically, most energetically*. [1, pg. 18]

Result: Cause of complexity of modeling the above analysis, the adverb is simply checked for its use as sentence, adverb, verb or adjective modifier. The above analysis is used to add some semantics to the adverbs. Hence such properties are added to ADV_TYPE feature of an adverb in the lexicon.

Future Work:

1. Comparatives in adverbs will be covered later on.
2. Compound adverbs having especially phrasal adverbs will be covered in future.
3. Negative adverb will be catered.
4. Semantics of adverbs will be analyzed in detail.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	27 th July, 04	1.0	Created
Zunaira Malik	20 th May, 05	2.0	New Release

Rule ID: EGR216

Rule Syntax: Following is the constituent description of the rule.

AuxP -> aux (neg)

AuxP -> aux (neg) aux

AuxP -> aux (neg) aux aux

Rule Functional Description: Following are the functional specifications of the rules.

AuxP -> aux: ^_AUX1 = !, ^_MAUX = !, ^_MAUX _MORPH_FORM = {BARE, PROG, PERF, PRES, PAST}, ^_AUX1 _MORPH_FORM = {BARE, PROG, PERF, PRES, PAST}, ^_FLAG = {DECL, INTERROGATIVE}, ^_MAUX TNS_ASP _JUNK = 'interrogative'; (neg: ^=!).

AuxP -> aux: ^_AUX1 = !, ^_AUX1 TNS_ASP _JUNK = 'two_aux_in_auxp', ^_AUX1 _MORPH_FORM = {BARE, PROG, PERF, PRES, PAST}; (neg: ^=!) aux: ^_MAUX = !, !_MORPH_FORM = c ^_AUX1 _ALLOWED_FORM, ^_MAUX TNS_ASP = ^_AUX1 TNS_ASP;.

AuxP -> aux: ^_AUX1 = !, ^_AUX1 _MORPH_FORM = {BARE, PROG, PERF, PRES, PAST}; (neg: ^=!) aux: ^_AUX2 = !, !_MORPH_FORM = c ^_AUX1 _ALLOWED_FORM; aux: ^_MAUX = !, !_MORPH_FORM = c ^_AUX2 _ALLOWED_FORM, ^_MAUX TNS_ASP = ^_AUX1 TNS_ASP;.

Frequency: -

Description: This rule shows the functional and constituent structure of auxiliary phrase.

c-structure: The rule shows that auxiliary phrase (AuxP) consists of one to three auxiliaries. This AuxP is used by both passive and active forms of sentences as well as with and without modals.

f-structure: The features of auxiliary are simply passed to AuxP as the auxiliaries are checked based on what was preceded by them. _MORPH_FORM is used to check well-formed ness and to make sure that the right auxiliary is used at a particular position. In case there is more than one auxiliary in AuxP, features of the auxiliaries are distinguished by grouping them in _AUX1, _AUX2 and _AUX3 structure. These AUX structures are also used for well-formed ness.

Examples:

- 1) She was driving. (from production 1)
- 2) She was not driving. (from production 1)
- 3) She has been eating her dinner. (from production 2)
- 4) She has not been eating her dinner. (from production 2)
- 5) She has been being tested. (from production 3)
- 6) She has not been being tested. (from production 3)



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

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules:

Related POS: EPOS101, EPOS120

Replaces: EGR101

Reason: New Release

Replaced by: -

Reason: -



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

Analysis:

Analysis1:

According to a previous analysis given in [1, p.62] auxiliaries have been treated as raising verbs. Under this view, auxiliaries are predicates which take a SUBJ and XCOMP complement. Their subject is identified with the subject of their XCOMP. The f-structure correspondingly reflects as many levels of embedding as there are auxiliaries in the c-structure. In particular, in the f-structure, the top level predicate does not correspond to what is intuitively the main predicate of the sentence. Hence in the following example, 'has' is the main predicate of the sentence:

Example:

She has appeared.

Analysis2:

Auxiliary can simply be considered as feature carrying elements which contribute tense/aspect information to the clause, but which do not have a PRED or sub-categorization frame. This analysis provides a flat f-structure. [1, p.64]

Result: Analysis 2 was finalized as it is more close to the intuition as the main verb is the main PRED of the sentence. In Analysis 1, auxiliary does not come out naturally as the main predicate of the sentence. The f-structure reflects as many levels of embedding as there are auxiliaries in the c-structure.

Future Work:



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	6 th September 2004	0.1	Created
Kiran Khurshid	22 nd September 2004	1.0	Approved by SA
Zunaira Malik	20 th May, 05	1.1	Addition of VPoblig

Rule ID: EGR217

Rule Syntax:

VPinf ->inf:: [VPact_main | VPpass] (VPoblig)

Rule Functional Description:

VPinf ->inf:: [VPact_main: ^!=!, ^_AUX1 _MORPH_FORM = TO_INF, ^SUBJ = NULL; | VPpass: ^!=!, ^_AUX1 _MORPH_FORM = TO_INF, ^SUBJ = NULL;] (VPoblig: ^!=!).

Frequency: -

Description: This rule is used to represent the to-infinitive phrase which can occur in a sentence.

c-structure: An infinitive verb phrase can contain "to" followed by an active verb phrase or a passive verb phrase.

f-structure: All feature values beginning with "_" are used to check the wellformedness of the phrase. The _MORPH_FORM of the verb phrases should be TO_INF in order to occur in the infinitive verb phrase.

Example:

- (1) You need to run to catch the bus
- (2) We are glad to have invited you.
- (3) I'd like to be working.
- (4) I'd hate to be questioned about it.
- (5) I'm glad to have been working.
- (6) He's said to have been invited.
- (7) I expect to be being interviewed then.



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

Rule Status: Active

Reference:

- (1) Pam Peters, "The Cambridge Guide to English Usage"
- (2) Quirk et al, "A Comprehensive Grammar of the English Language"
- (3) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR211, EGR207, EGR212

Related POS: EPOS119

Replaces: EGR106

Reason: New rule to define infinitive verb phrase. It is also renamed so that its name is more representative.

Replaced by:



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

Analysis:

Analysis: The basic nonfinite forms of verbs such as (to) ask, (to) go, (to) decide are called **infinitives** [1]. There are 5 basic forms of full verbs given on [pp. 96-97, 2]. The base form of the verb has two nonfinite forms: 1) the bare infinitive and 2) the *to*-infinitive. This rule is used to parse the *to*-infinitive phrases.

If we relate the structure of the nonfinite verb phrase to that of the finite verb phrase, we can tabulate the eight possible combinations as follows [pg. 153, 2]:

Type of verb phrase	Infinitive	Type
Simple	to examine	
Complex	to have examined	B
	to be examining	C
	to have been examining	D
	to have been examined	BC
	to be being examined	CD
	to have been being examined (rare)	BCD

Where Type B, C and D mean [pg. 152, 2]:

Type B: perfective aux (HAVE) + -ed participle

Type C: progressive aux (BE) + -ing participle

Type D: passive aux (BE) + -ed participle

“inf” is used to represent the “to” in the *to*-infinitive clause. It can either be followed by an active verb phrase or a passive verb phrase. The MORPH_FORM of these verb phrases must be TO_INF which means that the verbs of these clauses must be in the *to*-infinitive form (the nonfinite base form).

The infinitive clause is contained in XCOMP at the place from where it is called. An XCOMP is a complement whose subject is obligatorily functionally controlled from outside the clause as in [pg. 51, 3]:

Example: The woman wants to drive the tractor.

XCOMP = to drive the tractor.

Result: We decided on the above analysis.

Future Work:

- (1) Only some verbs allow infinitives with auxiliaries with them. This will be looked into, later.
- (2) The SUBJ of the VPinf will point to the subject of the main clause when the parser will have implemented it. For the time being the SUBJ is kept NULL.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	4 th Nov, 04	0.1	Created
Zunaira Malik	13 th June, 05	0.2	Change in rule
Zunaira Malik	25 th July, 05	0.3	Addition of second NUMBERP rule for 'of' phrases "hundreds of people" and GEN constraint in Cardinals

Rule ID: EGR218

Rule Syntax: Following is the constituent description of the rule.

NUMBERP -> (adj) [ord | card | [ord card]].

NUMBERP -> card p.

Rule Functional Description: Following are the functional specifications of the rule.

NUMBERP -> (adj: ^ADJUNCT = !;) [ord: ^=!, ^NUM = SG; | card: ^=!, ~[!CTYPE =c GEN]; | [ord: ^SPEC=!, !ORD_FORM =c 'first'; card: ^ = !, !NUM =c PL, ~[!CTYPE =c GEN];]].

NUMBERP -> card: ^=!, !CTYPE =c GEN; p: ! PFORM =c 'of', ^PFORM = 'of';.

Frequency: -

Description: This rule shows the functional and constituent structure of number phrase.

c-structure: Number phrase can consist of ordinal, cardinal, or an adjective. Number phrase can also consist of an ordinal followed by a cardinal.

f-structure: In the structure of the NUMBERP (number phrase), the NUM(number) of the ordinal if it comes alone will always be singular. The number (NUM) of the cardinal when it comes alone in the NUMBERP (number phrase) will always be plural. The ord (ordinal) will become a SPEC of the mother node if it is followed by the card (cardinal). In this case, the ORD_FORM (ordinal form) in the mother node must be 'first' . an adjective can also occur before the ordinal cardinal or ordinal or cardinal. The CTYPE must be GEN only in case of "hundreds of people" where the NUM is PL for the cardinal. The constraint is put to disallow "hundred of people" or "hundreds people" to be parsed.

Example:

- 1) The good book's first page is missing.
- 2) This table is three times the two big tables' width
- 3) The first three girls are very efficient.
- 4) *The second one book is missing.
- 5) Top ten books in the list
- 6) Hundreds of people



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

Rule Status: Active

References:

- [1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook".
- [2] James Allen, "Natural Language Understanding".
- [3] B.A. Hockey and Heather Mateyak, "Determining Determiner Sequencing: A syntactic Analysis for English"

Related Rules:

Related POS: EPOS112, EPOS113

Replaces: EGR133

Reason: Addition of second NUMBERP rule for 'of' phrases "hundreds of people" and GEN constraint in Cardinals

Replaced by: -



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

Analysis: Following is the in-depth analysis of the rule:

Analysis1:

Ordinals as in the example (1) behave just like simple attributive adjectives with regard to inflection.

Example:

1) The third tractor

However, in English these have a different c-structure distribution and so are assigned the syntactic category NUMBERP. They are also distinguished at f-structure in terms of the ATYPE that is assigned: ORDINAL.

Cardinals, on the other hand, do not inflect, but do require the noun to be plural (unless the cardinal is *one*), and exhibit a slightly different syntactic pattern at c-structure, as show by the contrast given in the example (2).

Example:

- 2) a. the three brown dogs
b. * the brown three dogs

Cardinals are thus also introduced by a special rule at c-structure (NUMBERP), and are distinguished at f-structure by being assigned ATYPE cardinal. [1, pp. 108 – 109]

Analysis2:

A simple noun phrase may have at most one determiner, one ordinal and one cardinal. It is possible to have all three, as in *the first three contestants*. [2, p. 27]

Analysis3:

McCawley gives as an argument “that numerals can be preceded by definite determiners whereas a determiner cannot in general be preceded by other determiner”. However many determiners can be preceded by other determiners as in *the many problems*, so this does not seem like very strong evidence for classifying cardinal numbers as adjectives.

Result:

NUMBERP in analysis 1 is used in c-structure to differentiate in from Adjectival Phrase. In f-structure it is represented as specifier in contrast to analysis 1, in which it is added as part of adjunct. This is supported in [3]. The sequencing of ordinals cardinals is as in analysis 2.

Future Work: -



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	1 st Nov, 04	0.1	Created
Zunaira Malik	13 th June, 05	0.2	Modified
Shanza Nayyer	28 th June, 05	0.3	
Zunaira Malik	27 th July, 05	0.4	Addition of VPinfmain in Pro-1

Rule ID: EGR219

Rule Syntax: Following is the constituent description of the rule.

ADJP -> (adv) adj.(VPinfmain)

ADJP -> adj.

ADJP -> adv adj.

ADJP -> adv adj.

Rule Functional Description: Following are the functional specifications of the rule.

- ADJP -> (adv: ^ADJUNCT ADV =!, !ADV_TYPE =c ADJ_MOD, ~[!SEM_TYPE =c {COMPARITIVE,SUPERLATIVE}]); adj: ^ =!, ^ADJ_FORM = 1; (VPinfmain: ^XCOMP =!, ^XCOMP SUBJ PRED = 'pro', ^XCOMP SUBJ PRONTYPE = NULL, ^_FLAG = ^XCOMP _FLAG);.
- ADJP -> adj: ^ =!, !ADJ_FORM =c {1,2,3};.
- ADJP -> adv: ^ADJUNCT ADV =!, !ADV_TYPE =c ADJ_MOD, !SEM_TYPE =c COMPARITIVE; adj: ^ =!, ^ADJ_FORM = 2; .
- ADJP -> adv: ^ADJUNCT ADV =!, !ADV_TYPE =c ADJ_MOD, !SEM_TYPE =c SUPERLATIVE; adj: ^ =!, ^ADJ_FORM = 3; .

Frequency: -

Description: This rule represents the adjective phrase.

c-structure: An adjective phrase may contain one or more adjective preceded by an adverb. The adverb is optional if the adjective form is 1 otherwise it is compulsory.

f-structure: The whole structure of an adjective is made a part of the adjective phrase set and the adv (adverb) is passed on to the ADJUNCT of the mother node (ADJP) as an ADV. The adverb of first production can be any adjective modifying adverb which is not in comparative or superlative form. This is done to avoid entries like 'more beautiful' to pass from production 1. It is to be noted that the adjectives having no intrinsically contained comparative and superlative forms such as 'beautiful', they do not have ADJ_FORM feature contained in them. ADJ_FORM feature to them is assigned through rules. Hence ADJ_FORM =1 is assigned to 'beautiful' in case of production1, ADJ_FORM =2 is assigned to 'more beautiful' in case of production 3 and ADJ_FORM =3 is assigned to 'most beautiful' in case of production 4. Production 2 is to parse adjective having intrinsically contained comparative and superlative forms such as 'quicker' and 'quickest' Production 1 contains a VPinfmain optional production at the end to cater to adjectives taking adjectival complements of the XCOMP form.

Examples:



- 1) She is a very good girl. (Production 1)
- 2) She is a beautiful girl. She is a quick girl. (Production 1)
- 3) He is the quicker/quickest dog. (Production 2)
- 4) The more beautiful girl. (Production 3)
- 5) The most beautiful girl (Production 4)
- 6) She is likely to run (Production 1- ADJ COMP of XCOMP form)

Rule Status: Active

Reference:

[1] <http://webster.commnet.edu/grammar/adjectives.htm>

[2] Randolph Quirk, "A Comprehensive Grammar of the English Language"

Related Rules: EGR105

Related POS: EPOS116

Replaces: EGR114

Reason: one rule divided in to 4 different ones catering to different adverbial forms that can occur with the adjective.

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis: Adjectives are words that describe or modify another person or thing in the sentence. [1] They are pre-nominals which may be preceded by an adverbial phrase. Only the adverb intensifiers are allowed before the adjectives. In such a case the adverbial phrase is made a part of the ADJUNCT mother node (ADJP). If more than one adjective occur, they all are made a part of the adjective phrase structure's set.

Example: i) She is a very mean girl.

For more details about adjectives see EPOS014.

Result: The above analysis was finalized for the rule.

Future Work:

1. The adjective phrase with conjunction is not allowed yet. It will be done later.
2. Adjectival Complementation will be catered.

Example

The tractor is good (to drive). (optional complementation with drive having tractor as OBJ and null SUBJ)

They are busy (preparing the lunch). (optional complementation)

He is likely to resign. (obligatory complementation) – **already done!**

All sales are subject to test. (obligatory complementation) [2, pg. 67]



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	22 nd Oct, 04	0.1	Created
Kiran Khurshid	7 th Feb, 05	0.2	Allowed multiple PPnmain and VPparticiple in post nominal
Kiran Khurshid	15 th Feb, 05	0.3	Allowed parenthesis
Zunaira Malik	13 th June, 05	0.4	Addition of ADJ_CompP, CompP and NPmain
Shanza Nayyer	29 th June, 05	0.5	ADJ
Zunaira Malik	25 th July, 05	0.6	Change of NPmain to NPnoun

Rule ID: EGR220

Rule Syntax: Following is the constituent description of the rule.

POSTNOMP -> (comma) [PAREN | VPparticiple | PPnmain | Rel_CI | ADJ_CompP | CompP | NPnoun].

Rule Functional Description: Following are the functional specifications of the rule.

POSTNOMP -> (comma;) [PAREN: ^ADJUNCT-PAREN= !; | VPparticiple: ! \$ ^ADJUNCT PARTICIPLE; | PPnmain: ! \$ ^ADJUNCT PREP; | Rel_CI: ^ADJUNCT REL_CL= !; | ADJ_CompP: ^ADJUNCT ADJ = !; | CompP: ^ADJUNCT ADJ COMPARITIVE = !, ^ADJUNCT ADJ ADJ_FORM = 2 ; | CompP: ^SPEC POST-DET QUANT COMPARITIVE = !, ^SPEC POST-DET QUANT SEM_TYPE = COMPARITIVE ; | NPnoun: ! \$ ^ADJUNCT RESTRICTIVE_NP, ^NUM = !NUM ; | VPinfmain: ^XCOMP = !, ^XCOMP SUBJ = NULL, ^_FLAG = ^XCOMP _FLAG;].

Frequency: -

Description: This rule shows the phrase/ clause that occurs after an NP and functions as a modifier of NP.

c-structure: Noun Phrase can be followed by a prepositional phrase, a relative clause, a participle verb phrase, a parenthetical phrase, an adjective comparative phrase, an adjectival comparative, a comparative quantifier or an NPnoun.

f-structure: The parenthetical phrase is made a part of the ADJUNCT-PAREN structure according to the Cookbook analysis [1, p. 148]. All of the productions are made into sets as they can occur one or many times. VPparticiple can also come as an adjunct to NP, providing some detail of the noun, e.g. 'she being the eldest has to be most responsible'.

PPnmain can also occur as a modifier of head noun, e.g. 'Ahmad, from Lahore is applying for visa'.

A relative clause is the clause adding description to the noun and hence is added as an ADJUNCT.

ADJ_CompP is used to show an adjective comparison phrase and it goes as an ADJUNCT to NP, .e.g. 'more beautiful than her'.

A CompP production which is part of ADJP is made an ADJUNCT in f-structure and checks for presence of comparative form of adjective in earlier part of ADJP, e.g. 'more amazing work than I thought'. CompP can also be part of the QUANTP as in 'more work than I wanted', so the production checks for the SEM_TYPE of quant to be COMPARATIVE.



In case of the NPnoun, example 7 type occurrences are handled as restrictive NPs. This will later be in a comma separated form to avoid over-generation e.g. 'John, the vice president is resigning'.

Examples:

- 1) The girl whose books I'm reading is good. (Relative Clause)
- 2) The glass on the table in the car is broken. (multiple PPn)
- 3) The letter written by me is placed on the table. (VPparticiple: -ed participle)
- 4) The girl writing the letter has left. (VPparticiple: -ing participle)
- 5) The girl more beautiful than me is in the contest. (ADJ_CompP with more as ADV)
- 6) The girl quicker than me is in the race. (ADJ_CompP with comparative form of an ADJ)
- 7) The more beautiful girl than me. (CompP)
- 8) John, the Vice President is resigning. (NPmain)

Rule Status: Active

Reference:

- (1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR113, EGR144, EGR148

Related POS: -

Replaces: EGR112

Reason: - change of NPmain to NPnoun

Replaced by: -

Reason: -



Analysis: After parsing the sample corpus it was found that multiple prepositional phrases are allowed after each noun.

Kindly refer to the related rules documents for more details.

Result: -

Future Work:

- The rule for PPnmain further has NPmain allowed in it. Hence for NPs containing more than one PPs there should be over-generation. However due to a bug in the parser only one f-structure is formed. This will be seen into later.
- More post-nominals might be added in the future.
- Once the comma starts working in the parser, a (comma) optional phrase can also occur before the POSTNOMP (John, the vice president, father of two, and CEO of the company came to lunch). It is commonly seen that NPs occurring post-nominally are preceded by comma, hence for NPs comma will be made mandatory to avoid over-generation.
- The checking of comparative form in CompP is done till yet through assignment operation due to parser problem and hence is over-generating some redundant f-structures. This will be changed to constraint operation (=c) operation in the later stage.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	13 th June, 05	0.1	Created

Rule ID: EGR221

Rule Syntax: Following is the constituent description of the rule.

S -> NPmain.

S -> PPnmain.

S -> NPmain.

Rule Functional Description: Following are the functional specifications of the rule.

S -> NPmain : ^=!, ^FLAG = HYPERLINK ;.

S -> PPnmain: ^=!, ^FLAG = HYPERLINK ;.

S -> NPmain : ^SUBJ=!, ^FLAG = HYPERLINK; VPinf: ^=!, ^TNS_ASP TENSE= FUTURE;.

Frequency: -

Description: This rule shows the root level production of a sentence in hyperlink or heading form.

c-structure: A sentence in a heading may contain an NPmain, a PPnmain or an NPmain with a future tense/aspect tense.

f-structure: A hyperlink flag is used to indicate that the sentence is a hyperlink or a heading. NPmain and PPnmain are passed on as they once the flag is detected, while the NPmain with a future tense passes on to SUBJ and VPinf passes on as it is.

Examples:

- 1) efficiently enabling mass access to information, living with HIV, facing a pension crunch
- 2) Hurriyat leaders to get warm welcome.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

[2] <http://englishplus.com/grammar/glossary.htm>

Related Rules: EGR205, EGR144

Related POS:

Replaces:

Reason:

Replaced by: -



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



Analysis:

Result:

Future Work:



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	3 rd Nov, 04	0.1	Created
Kiran Khurshid	15 th Feb, 05	0.2	Added compound noun (noun before noun)
Zunaira Malik	13 th June, 05	0.3	Change of ADJP to ADJPmain
Zunaira Malik	20 th June, 05	0.4	Addition of NCOUNT constraint in PREDET and assignment of value to NCOUNT, change of NUM to be assigned NNUM value and removal of DEF in DET
Zunaira Malik	25 th July, 05	0.5	Addition of NNUM constraint to POSTDET for quantifier and number phrase agreements

Rule ID: EGR222

Rule Syntax: Following is the constituent description of the rule.

PRENOMP -> (PREDET) (DET) (POSTDET) (ADJPmain) [n]*.

Rule Functional Description: Following are the functional specifications of the rule.

PRENOMP -> (PREDET : ^SPEC PRE-DET =!, ^NCOUNT = ! NCOUNT, ^NUM = ! NNUM;) (DET: ^SPEC DET = !, ^NCOUNT =

{POS, NEG}, ^NUM = ! NNUM;) (POSTDET : ^SPEC POST-DET =!, ^NUM = ! NNUM;) (ADJPmain: ^ADJUNCT ADJ = !,

^ADJUNCT ADJ ADJ_TYPE = ATTRIBUTIVE ;) [n: ! \$ ^ADJUNCT MOD_N;]*.

Frequency: -

Description: This rule shows the functional and constituent structure of pre-nominal. This production further expands to give all possible specifiers and modifiers which precedes and follows noun.

c-structure: The production caters the modifiers and specifiers before and after noun.

The detail of PREDET, DET, POSTDET and ADJPmain will be discussed in their relevant documents.

f-structure: In the structure of the PRENOM, in case of the DETP (determiner phrase), the DET will be copied to the mother node as the SPEC of the mother node, NCOUNT as {POS, NEG} and NUM as NNUM will be copied to the mother node. PREDET and POSTDET are already made a SPEC of their particular type. PREDET causes NCOUNT value to travel up from below to enforce a constraint of similar NCOUNT from bottom to mother node,



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

NUM also passes to check for a PL object in case of the quantifier “few” or “many”, which can only take plural subjects..

To prevent parsing of

Double the cat - where ‘cat’ has an NCOUNT = POS which needs to be compared with the NCOUNT = NEG value of “double” and fail.

The ADJECTIVE in the ADJPmain will be copied to the mother node as an ADJUNCT of the mother node. NUM and NCOUNT in the DET will be unified in the mother node (it is to enforce as constraint on NUM and NCOUNT). A noun may also occur before a noun. This forms a compound noun. The noun becomes a part of the set ADJUNCT MOD_N.

Examples:

- 1) Some of those 12 girls were dishonest.
- 2) My first daughters’ son is well-mannered.
- 3) Few good men attended the conference.
- 4) The good grammar writer John is working.
- 5) Dogs voice concern over Iraq Oil
- 6) Double the fun
- 7) Hundreds of people

Rule Status: Active

Reference:

[1] James Allen, “Natural Language Understanding”

[2] Miriam Butt, Tracy Holloway King, “A Grammar Writer’s Cookbook”

Related Rules: EGR130, EGR131, EGR132, EGR114

Related POS: EPOS105

Replaces: EGR129

Reason: Addition of NNUM constraint to POSTDET for quantifier and number phrase agreements

Replaced by: -

Reason: -



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

Analysis: Following is the in-depth analysis of the rule.

Analysis 1:

In addition to head, a noun phrase may contain *specifiers* and *qualifiers* preceding the head. The qualifiers further describe the general class of objects identified by the head, while the specifiers indicate how many such objects are being described, as well as how the objects being described relate to the speaker and hearer. Specifiers are constructed out of *ordinals*, *cardinals* and determiners. Determiners can be sub-divided into the classes: articles, demonstratives, possessives, wh-determiners, quantifying determiners.

The qualifiers in a noun phrase occur after the specifiers (if any) and before the head. They consist of adjectives and noun being used as modifiers. [1, p. 26]

For detail of noun and adjective see the relevant POS files.

Analysis 2:

Articles, quantifiers and pronominal genitives pattern similarly in English in that they appear in the first position in an NP: they precede any modifiers as in the example below, and cannot be preceded by another article or quantifier which modifies the noun.

Example: i) The/ a/ every/ Kim's small dog barks.

ii)* The a/ every/ Kim's small dog barks.

The intuition that has guided most of the modern syntactic approaches to these constructions is that they serve to "specify" the head noun rather than simply "modify" it.

Articles, quantifiers and pronominal genitives are treated uniformly in all three grammars in that they are represented under SPEC feature in the f-structure. [2, p. 101]

Titles like *Professor John Smith* and names like *linguistics department* are treated as N-N sequence and are parsed by a specialized subset of c-structure rules within the NP rule system. While German is well-known for its lexical noun compounding, English employs a non-lexical compounding strategy and French uses PPs. [2, p. 90]

Example: i) hydraulic oil filter.

In English and French, these nonlexical compounds are dealt with by means of a special c-structure rule which analyses N-N sequence that are not titles or names as compounds in the f-structure. [2, p. 90]

While the rules necessary to treat most N-N sequence found in the languages are in place, a problem remains. Almost any noun can be turned into a "title", as in *grammar writer John*, or a name. Hand coding each nominal lexical entry for precise information is unfeasible, and loosening the c-structure rules to allow for new creations of N-N sequences can lead to overgeneration. Thus, while the grammars can parse most N-N sequences, the issue has not yet been completely resolved. [2, p. 92]

According to the Cookbook analysis [2] the compound noun is made a part of the structure named COMPOUND within an NP, but as we analyze the noun occurring before a noun as a modifier noun only hence we place it within the set ADJUNCT in the set named MOD_N.

Result: The overall idea of pronominal is taken from the above analyses but the ordering of pronominal in it is not very sophisticated. Hence the different kinds of pre-nominals formed in the above productions are subject to revision.

Future Work:

- 1) There will be more possibilities of sequencing of pre-nominals which will be accounted for in later stage.
- 2) The agreement of PRENOMP with head noun will be checked.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	1 st Nov, 04	0.1	Created
Zunaira Malik	13 th June, 05	0.2	Addition of another POSTNOMP to prevent over-generation
Zunaira Malik	25 th July, 05	0.3	Addition of DEMONST constraint

Rule ID: EGR223

Rule Syntax: Following is the constituent description of the rule.

NPnoun -> (PRENOMP) [n] (POSTNOMP)(POSTNOMP).

Rule Functional Description: Following are the functional specifications of the rule.

NPnoun -> (PRENOMP: ^ = !;) [n: ^ = !, ^PERS = 3, !PRONTYPE = c DEMONST; | card: ^ = !, ^PERS = 3;]
(POSTNOMP: ^ = !;)(POSTNOMP: ^ = !;).

Frequency: -

Description: This rule shows the production for proper and common noun.

c-structure: This production shows the syntax of proper and common noun. It was noted that proper noun can be preceded by all possible modifiers and specifiers as that of a common noun depending on the context in which proper noun is used. Some noun modifiers can even occur after the head noun.

POSTNOMP can traverse for zero or many times (this will be implemented once comma is implemented in the scanner, to avoid generating too many structures.

example: grammar will do the following (in future):

-NP acting postnominally is specifically separated by commas (the girl, amna is sleeping/ the girl in bed, amna, is sleeping)

f-structure: The PRENOMP becomes SPEC of the mother node while noun becomes the head. PRENOMP contains the pre-nominal modifiers. Some modifiers follow the head noun and are added to POSTNOMP.

Examples:

- 1) There are three Alis in this class; I want to meet all of them.
- 2) Urdu is a rich language.
- 3) Some of these many Johns are honest.
- 4) The good John worked and the bad John played.
- 5) One fourth of the Punjab was taken by the foreigners.
- 6) A good girl went to school.
- 7) John's good son is driving.
- 8) The girl which I saw was beautiful.
- 9) The girl on the table is very naughty.
- 10) One of the girls is evil.
- 11) Those of us who went.



Rule Status: Active

Reference:

- [1] <http://webster.commnet.edu/grammar/determiners/determiners.htm>
- [2] Discussion meeting with Dr. Shahbaz Arif (Head of English Language and Linguistics, Punjab University)
- [3] W. John Hutchins, Harold L. Somers, "An Introduction to Machine Translation"

Related Rules: EGR129, EGR112

Related POS: EPOS105

Replaces: EGR120

Reason: Addition of DEMONST constraint

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis 1:

Definite article can be used with some proper nouns. Like it is not used with the names of languages

Example: He was learning Chinese.

But when the word “Chinese” refers to the people, the definite article might come into play.

Example: The Chinese are hoping to get the next Olympics. [1]

Analysis 2:

Everything that is possible with noun is possible with proper noun (except indefinite article (a,an)), with certain contextual constraints. When there is a group of same proper nouns, it actually starts behaving like a common noun. We cannot use modifiers/specifiers with proper noun without contextual aspect involved in it. [2]

Example:

1) *This Hamid went to school. Vs This Hamid went to school, not that one.

2) *My Pakistan is a good country Vs Pakistan and India, both have Punjab province; Pakistan's Punjab is better.

Result: Both analyses are used. Currently proper noun is considered without the semantic features, which confirm the use of particular pre-nominal before pronoun.

Future Work:

- 1) The properties of proper noun that allows/disallows specific specifier/modifier, will be analyzed in later stage. This is a semantic level check and currently the grammar development is kept specific to syntactic considerations.
- 2) Noun as noun modifier will be added.
Example: Oil breaks
- 3) Noun partitives will be analyzed.
Example: A certain kind of machine
- 4) Till yet PRENOMP is placed flatly with head n and post-nominal modifiers. Considering the intuition behind X-bar theory (complement more strongly attached to the head) [3, p. 35], the level of POSTNOMP and PRENOMP might be revised.
- 5) POSTNOMP can traverse for zero or many times (this will be implemented once comma is implemented in the scanner, to avoid generating too many structures.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	13 th June, 05	0.1	Created

Rule ID: EGR224

Rule Syntax: Following is the constituent description of the rule.

VPact_main -> (ADVP) v (ADVP)

Rule Functional Description: Following are the functional specifications of the rule.

1) VPact_main -> (ADVP:! \$ ^ADJUNCT ADV,!ADV_TYPE =c V_MOD;) v: ^ = !, !_MORPH_FORM =c PROG, !VOICE =c ACTIVE , ^TNS_ASP TENSE = PRES, ^FLAG = HYPERLINK ; (ADVP:! \$ ^ADJUNCT ADV,!ADV_TYPE =c V_MOD;).

Frequency: -

Description: This rule shows the VP production in active form of English Grammar with respect to hyperlinks and headlines.

c-structure: The productions consists of an optional ADVP followed by a verb and then another optional ADVP.

f-structure: The Adverb in the ADVP is taken up as an ADJUNCT after comparing the ADV_TYPE and the V_MOD. A hyperlink flag is also added to detect that the verb phrase is a hyperlink or headline.

Examples:

- 1) Government Working to Build Future

Rule Status: Active

Reference: -

Related Rules:- EGR215

Related POS: EPOS003

Replaces:

Reason:

Replaced by: -

Reason: -



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



Analysis: Please refer to the related rules and related POS documents (in references) for details.

Result: -

Future Work: -



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	13 th June, 05	0.1	Created

Rule ID: EGR225

Rule Syntax: Following is the constituent description of the rule.
VPpass -> v (ADVP).

Rule Functional Description: Following are the functional specifications of the rule.
VPpass -> v: ^ = !, !VOICE =c PASSIVE, ^FLAG = HYPERLINK ; (ADVP:! \$ ^ADJUNCT ADV,!ADV_TYPE =c V_MOD;).

Frequency: -

Description: This rule shows the VP passive form production of English Grammar with respect to headlines and hyperlinks.

c-structure: the rules consists of a verb followed by an optional ADVP.

f-structure: A hyperlink flag is present and the ADV in the ADVP is added to an ADJUNCT set after comparing ADV_TYPE with V_MOD.

Examples:

1) Thirteen killed in bomb attacks.

Rule Status: Active

Reference:

[1] James Allen, "Natural Language Understanding"

[2] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR215

Related POS: EPOS003

Replaces:

Reason:

Replaced by: -

Reason: -

Analysis:

Result:

Future Work:



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





Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	13 th June, 05	0.1	Created

Rule ID: EGR226

Rule Syntax: Following is the constituent description of the rule.
Sdecl -> VPparticiple VPmain (ADVP).

Rule Functional Description: Following are the functional specifications of the rule.
Sdecl -> VPparticiple: ^SUBJ = !; VPmain: ^!= !; (ADVP: ! \$ ^ADJUNCT ADV, ! ADV_TYPE =c S_MOD;).

Frequency: -

Description: This rule shows the sentence level production of declarative sentence of English Grammar.

c-structure: This rule (declarative sentence) consists of a participle Verb Phrase (VPparticiple) and a Verb Phrase (VP). Adverbial Phrase (ADVP) can optionally come as sentence modifier.

f-structure: The structure contained in VP participle is equal to the SUBJ structure of mother node S. All that is contained in VP is equivalent to the mother node S. ADVP becomes ADJUNCT.

Examples:

1) Smoking is bad for health.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR210, EGR146, EGR215

Related POS:

Replaces:

Reason:

Replaced by: -

Reason: -



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



Analysis

Result:

Future Work:



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	13 th June, 05	0.1	Created

Rule ID: EGR227

Rule Syntax: Following is the constituent description of the rule.

ADJ_compP -> adj.

ADJ_compP -> adv adj.

Rule Functional Description: Following are the functional specifications of the rule.

1) ADJ_compP -> adj: ^ = !, !ADJ_FORM = c 2; CompP: ^COMPARITIVE = !;

2) ADJ_compP -> adv: ^ADJUNCT ADV = !, !ADV_TYPE = c ADJ_MOD, !SEM_TYPE = c COMPARITIVE; adj: ^ = !, ^ADJ_FORM = 2; CompP: ^COMPARITIVE = !;

Frequency: -

Description: This rule shows the Adjective Comparative Phrase production in English Grammar.

c-structure: The first production is for adjectives having intrinsically contained comparative form. The second production is for the adjective showing comparative sense through comparative adverb.

f-structure: The first rule takes the comparative adjective having intrinsically contained comparative sense. In the second one, the adverb is adding comparative sense to the adjective and is taken as an ADJUNCT of ADJ in f-structure. CompP is added to f-structure as COMPARITIVE structure.

Examples:

- 1) The cat quicker than the dog.
- 2) The cat more beautiful than the dog.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook" ,

[2] Quirk, Greenbaum, Leech, Svartvik, " A Comprehensive Grammar of the English Language"

Related Rules:

Related POS: EPOS200, EPOS014

Replaces:

Reason:

Replaced by: -

Reason: -



Analysis:

“more” is the comparative of the quantifiers “many” and “much”.

Morton has more expensive clothes than I have.

More modifies expensive here making the adjectival phrase a comparative one. This results in an adjectival comparative phrase.

Morton has [more expensive clothes than I have]. “more” cannot be used with definite NPs in a comparative construction. For e.g.

Sally has the better radio than Daniel (has). * [2, pg. 1135]

“more..than” is considered to be a comparative correlative since the second endorses the meaning of the first.

A cat more beautiful.... Although sounds alright but seems incomplete without a “than” clause.

A cat more beautiful than the dog. [2, pg. 999]

According to the cookbook by Miriam Butt, adjectival comparatives are analyzed in two main parts:

- 1) degree (more as a comparative)
- 2) comparative phrase it can license (“than a truck”)

The degree is treated like as an adjunct like all other adjectives, but a special one with special syntactic properties. It is called the ADJUNCT_COMP (indicates that it is an adjunct that specifies a degree of comparison). The comparative phrase is analyzed as an OBL argument of the adjective called the OBL_COMP.

“than” is also treated as a special type of conjunction as it may head complete clauses.

Example:

1) *It is more comfortable than a tractor.* [Adj_comp Phrase]

more comfortable [AP]

more [comparative] , *comfortable* [adjective]

than a tractor [obl_comp]

than [conj_comp]

[1, pg. 118, 119, 120, 121, 122]

Result: -

- 1) a beautiful cat (ADJ_FORM =1
- 2) a more beautiful cat (ADJ_FORM =2 due to presence of more) (comparative form)
- 3) a more beautiful cat than the dog.

More beautiful cat [ADJP]

Than the dog [COMP Phrase]

Than [comp_conj]

More beautiful cat than the dog [ADJ_COMP phrase]



Future Work: -

- 1) We shall cater to “times” phrases that use more than, for e.g. *five times more people came to the demonstration than (did) last time.*
- 2) Identification of “than” as a preposition, when it usually introduces Noun phrases.
For e.g. *I weigh more than 200 pounds.*
It goes faster than 100 miles per hour.
- 3) Identification of “than” as a relative pronoun, where there is an omission of the subject in the comparative clause.
For e.g. *you spent more money than was intended to be spent.*
- 4) Identification of “more than” as a subject pronoun.
For e.g. *more than a 1000 inhabitants have signed the petition.*
More than one member has protested against the proposal.
- 5) Usage as an adverb.
For e.g. *he likes her more than me.*
John , more than his brothers, was responsible for the loss.
She'll enjoy it more than last year.
- 6) Identification of “more of a “ phrases.
For e.g. *he is more of a fool than I thought he was.* (“more” here is used as an adverb)
- 7) Identification of “no more ...than” phrases.
For e.g. *I have no more money than you have.*
Rachel is no more competent than Saul (is).



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	13 th June, 05	0.1	Created
Shanza Nayyer	8 th Aug, 05	0.2	Future Work added

Rule ID: EGR228

Rule Syntax: Following is the constituent description of the rule.
CompP -> comp_conj [NPmain | Sdecl].

Rule Functional Description: Following are the functional specifications of the rule.

CompP -> comp_conj: ^=!; [NPmain: ^=!; | Sdecl: ^=!, ^OBJ PRED = 'pro', ^OBJ PRONTYPE = NULL;].

Frequency: -

Description: This rule shows the CompP production in English Grammar.

c-structure: The CompP phrase can take a comparative conjunction followed by an NP or a declarative sentence.

f-structure:

The production takes comp_conj followed by NP or declarative sentence. It is to be noted that Sdecl takes null object.

Examples:

- 1) He is stronger than me.
- 2) It is more than I wanted.

Rule Status: Active

Reference: -

Related Rules: EGR205, EGR139

Related POS: EPOS202

Replaces:

Reason:

Replaced by: -

Reason: -



Analysis: Please refer to the related rules and related POS documents (in references) for details.

Result: -

Future Work:

1. Comparison in NUMBERP, NP and ADVP needs to be handle the following comparitives:

Example:

He worked more cleverly than I though

More than five books are missing

He did more than my expectations

2. ADJ_CompP needs to handle the following forms of comparative phrase:

Only more than ordinary people can take this risk

She is a more than ordinary



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	28 th Oct, 04	0.1	Created
Zunaira Malik	14 th June, 05	0.2	Addition of VP Participle for hyperlinks and headings

Rule ID: EGR229

Rule Syntax: Following is the constituent description of the rule.
NP -> [NPnoun | NPpron | VPparticiple].

Rule Functional Description: Following are the functional specifications of the rule.
NP -> [NPnoun: ^=!; | NPpron: ^=!; | VPparticiple: ^=!;].

Frequency: -

Description: This rule shows different types of noun phrases that can be made.

c-structure: NP can be a noun phrase having common noun, proper noun, simple pronoun or a VP participle in case of hyperlinks or headlines. The details of these noun phrases are in their relevant grammar rule files.

- NPnoun is a noun phrase containing common and proper noun along with its all possible modifiers or specifiers.
- NPpron is the noun phrase having pronoun.
- VP participle is a Noun Phrase having an –ing form of the verb beings used as a noun (Deverbal nouns). Usually occurs inside a PP.

f-structure: The structure information of different NPs goes to the mother NP node which becomes the SUBJ or OBJ of the f-structure.

Examples:

- 1) This good girl is walking. (NPnoun)
- 2) John ate an apple. (NPnoun)
- 3) He is a good boy. (NPpron)
- 4) On visiting this site. (VP participle)
- 5) Smoking is bad for health. (VP Participle)
- 6) Driving is good for us. (VP Participle)

Rule Status: Active

Reference: - A Grammar Writer's Cookbook, Miriam Butt [1]

Related Rules: EGR120, EGR121, EGR210

Related POS: -

Replaces: EGR118

Reason: New release. VP participle added to cater to de-verbal nouns (verbs in –ing form)

Replaced by: -

Reason: -



Analysis: -

GERUNDS (De-verbal nouns or verbs in participle form acting as nouns)

They are analyzed externally as NPs, in that they receive case, are assigned person and number for verb agreement, and may appear in NP positions in the c-structure. Internally however, they are clausal.

- a. *[Driving the tractor] is good.*
- b. *[His/John's driving] the tractor amazes me.*

They either occur with a null pronominal subject (a) or a genitive pronoun or possessive subject NP.[1]

Driving the tractor is good.

According to the c-structure:

[Driving the tractor] (NP gerund or VP)

[Driving] (verb)

[the tractor] (NP)

[is good] (VP copular)

[is] (V copular)

[good] [AdjP]

Result: -

Future Work:

- 1) Other kinds of NPs like partitives etc. will be added to the noun phrase in later stage.
- 2) Gerunds will be able to take the possessive form when being used as possessive NPs.

Example:

NPpartitive: A kind of these machines are in my factory



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	21 st April, 04	0.1	Created
Kiran Khurshid	27 th April, 04	0.2	Modified after discussion in meeting
Zunaira Malik	20 th June, 05	0.3	Problem with prep constraint for 'of'

Rule ID: EGR230

Rule Syntax: Following is the constituent description of the rule.

FRACTIONP -> card ord

The following does not work

FRACTIONP -> card ord p

Rule Functional Description: Following are the functional specifications of the rule.

Allows parsing of :

He ate two third the pie

FRACTIONP -> card: ^ CARD=!, ord: ^ ORD=!, ~[!ORD_FORM =c 'first' | !ORD_FORM =c 'second'];.

The following one does not work due to problems with the parser, might be a conflict with the genitive 'of' phrase , for e.g.

Advice of the parents

FRACTIONP -> card: ^!=!, ord: ^!=!, ~[ORD_FORM =c 'first' | ORD_FORM =c 'second']; p: !PCASE =c 'of', ^PCASE = 'of';.

Frequency: -

Description: This rule shows the functional and constituent structure of fraction.

c-structure: This production gives the order in which fraction can occur before determiner. The rule at work here is that a cardinal number (e.g. two, three) is followed by an ordinal number (e.g. fifth, sixth) which may or may not, in turn be followed by the preposition "of".

f-structure: Here the constraint which is applied on ordinals is that the ordinal *first and second* should not follow any cardinal number.

It is also noted that if the fraction is followed by a preposition, than it is always 'of' hence all other prepositions are constrained. Also only PCASE is passed to the mother node and all other feature values of preposition like sub-categorization frame etc. is filtered out. This is done cause 'of' here is used as a connector to the fraction, not as the head of the phrase. However, this constraint creates a problem, maybe because the parser conflicts with genitive 'of' phrase in case, for e.g.:

Cup of tea



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



Examples:

- 1) Charlie ate two-fourth of the pie //problem
- 2) * Charlie ate two-first of the pie //problem
- 3) He completed two-sixth the work //works with omission of prepositional phrase constraint

Rule Status: Active

Reference: -

Related Rules: EGR010

Related POS: EPOS009, EPOS010, EPOS016

Replaces: - EGR135

Reason: - previous one does not work with the 'of' phrase constraint so it was removed to allow functioning without the 'of' phrase which is optional.

Replaced by: -



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

Analysis: -

Result: -

Future Work: Following issues need to be addressed, and constraints to be included respectively, after further analysis:

- 1) The ordinal number must be less than the cardinal number of a fraction. This relationship needs to be implemented in the rule. This will be dealt with while considering the semantics of the sentences.
- 2) Fractions cannot occur with words like girl, boy, etc. (e.g. *one fourth of a girl), hence some constraints at the semantic level will be introduced to differentiate between the “animated” and “unanimated” nouns, occurring after fractions.
- 3) Fraction phrase like 1/100th etc. still need to be analyzed.
- 4) Problem of ‘of’ phrase in case of “ he ate two-third of the pie” needs to be solved by resolving the conflict with the genitive ‘of’ phrase in “son of John”.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	4 th Nov, 04	0.1	Created
Zunaira Malik	20 th June, 05	0.2	Problem with one single rule for PRENOMP so divided into two

Rule ID: EGR231

Rule Syntax: Following is the constituent description of the rule.

The following does not work for PRENOMP

It parses “*times a day*” but not “*two times a day*” when it should do the opposite.

MULTP -> [mult | card mult].

It works if divided into two separate rules

MULTP -> mult

MULTP -> card mult

Rule Functional Description: Following are the functional specifications of the rule.

MULTP -> [mult: ^ !=!, ~[! MULT =c 'times']; | card: ^ !=!, mult: ^ = !, ! MULT =c 'times';].

MULTP -> mult: ^ !=!, ~[! MULT_FORM =c 'times'];.

MULTP -> card: ^ !=!, !NUM =c PL; mult: ^ = !, ! MULT_FORM =c 'times';.

Frequency: -

Description: This production shows the MULTP (multiplier phrase) that can occur before determiner.

c-structure: Multiplier can have cardinal number (two, three, etc.) followed by the multiplier ‘times’. Multipliers such as *once*, *twice* etc. can also come before determiner.

f-structure: The structure information of multiplier goes as the SPEC to the DETP. Multiplier can have a plural cardinal number (two, three, etc.) followed by the multiplier ‘times’. It is to be noted that cardinal numbers include all numbers except *one*. *One* is the only cardinal having NUM as singular and will be catered in the later stage.

Examples:

- 1) He gave me double the amount I asked for.
- 2) Twice his salary is my expenditure.
- 3) This van holds three times the passengers as that sports car



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

Rule Status: Active

Reference:

[1] <http://webster.commnet.edu/grammar/determiners/determiners.htm>

Related Rules:

Related POS: EPOS111, EPOS112

Replaces: EGR134

Reason: problem with one single rule for PRENOMP so divided into two

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule.

Analysis: The multipliers precede plural count nouns and mass nouns and occur with singular count nouns denoting number or amount [1]. They are somehow multiplying the quantity of the coming noun. Some multipliers (such as twice, or thrice) can occur alone before the determiner, while in a special case of the multiplier “time”/“times” they demand that a cardinal number precedes them.

Example:

My wife is making twice my salary

Result: The above analysis was finalized.

Future Work: Following issues need to be addressed, and constraints to be included respectively, after further analysis:

- 1) The usage of the multipliers “time” and “times” need to be specified.
- 2) The determiner can be an article, pronoun genitive or a demonstrative. But with certain multipliers indefinite article is not allowed. This constraint needs to be looked into.

*Example: *Twice a total Vs Twice the total*

- 3) need to figure out why one single rule does not work for PRENOMP. It parses “times a day” but not “two times a day”



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	3 rd Nov, 04	0.1	Created
Zunaira Malik	20 th June, 05	0.2	Change in MULTP of NCOUNT, NUM to NNUM, Change in FRACTIONP by removal of SPEC set name, change in adv of ADV_TYPE and NCOUNT, NUM to NNUM

Rule ID: EGR232

Rule Syntax: Following is the constituent description of the rule.
PREDET -> [MULTP | FRACTIONP | adv | PREDETQUANTP]

Rule Functional Description: Following are the functional specifications of the rule:

PREDET -> [PREDETQUANTP: ^QUANT = !, ^DEF = POS, ^NUM = ! NUM; | MULTP: ^MULT = !, ^NCOUNT = ! NCOUNT, ^NNUM = {SG}; | FRACTIONP: ^FRACTION = !, ^DEF = POS, ^NCOUNT = {POS, NEG}, ^NNUM = SG; | adv: ^ADV = !, ^ADV_TYPE = c N_MOD, ^DEF = NEG, ^NCOUNT = {POS, NEG}, ^NNUM = SG;].

Frequency: -

Description: This rule shows the functional and constituent structure of pre- determiner phrase.

c-structure: Basically DETP phrase is divided into three parts PREDET, DETP and POSTDET. PREDET can have MULTP (multiplier phrase), FRACTIONP (Fraction phrase), adv(adverb) or PREDETQUANT (pre-determiner quantifier). Addition in NUM of PREDETQUANTP to check for PL subjects in case of “few” and “many” which can only take plural subjects.

f-structure: Everything in PREDET goes as a SPEC to the mother node i.e. PRENOMP.

Examples:

In the following examples the underlined part is the pre-determiner phrase:

- 1) She spent three times the amount. (MULTP)
- 2) Give me two sixth the pie. (FRACTIONP)
- 3) Many of the books are missing. (PREDETQUANTP)
- 4) This room is quite a mess, isn't it? (adv)



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

Rule Status: Active

Reference:

- [1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"
- [2] "A Lexicalized Tree Adjoining Grammar for English", Institute for Research in Cognitive Science, University of Pennsylvania
- [3] <http://webster.commnet.edu/grammar/determiners/determiners.htm>
- [4] B. A. Hockey and Heather Mateyak, "Determining Determiner Sequencing: A Syntactic Analysis for English" University of Pennsylvania.

Related Rules: EGR134, EGR135, EGR136

Related POS: EPOS102

Replaces: EGR131

Reason: change in MULTP, FRACTIONP and adv

Replaced by: -

Reason: -



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

Analysis: Following is the in-depth analysis of the rule.

Analysis 1: Check the DETP reference document for detailed analysis

Result: -

Future Work: Following issues need to be addressed, and constraints to be included respectively, after further analysis:

Constituency tests will be applied to check of validity of PREDET being flatly placed with DETP.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	22 nd Oct, 04	0.1	Created
Shanza Nayyer	26 th Oct, 04	1.0	Rule finalized after discussion with SA
Shanza Nayyer	2 nd Nov, 04	1.1	Added a new rule to relative clause
Zunaira Malik	22 nd June, 05	1.2	change of relative pronoun occurrence to optional in production 2 and change of ADV_TYPE in production 3 to SEM_TYPE

Rule ID: EGR233

Rule Syntax: Following is the constituent description of the rule.

Rel_CI -> pro VPmain

Rel_CI -> (pro) Sdecl

Rel_CI -> adv Sdecl

Rel_CI -> PP Sdecl

Rule Functional Description: Following are the functional specifications of the rule.

- 1) Rel_CI -> pro: ^SUBJ = !, ! PRONTYPE =c RELATIVE, !CASE =c NOM; VPmain: ^!=!;
- 2) Rel_CI -> (pro: ^OBJ = !, ! PRONTYPE =c RELATIVE, !CASE =c ACC;) Sdecl: ^!=!;
- 3) Rel_CI -> adv: ^ADJUNCT ADV = !, !SEM_TYPE =c INTERROG; Sdecl: ^!=!;
- 4) Rel_CI -> [PPn: ^OBL = !, !OBJ PRONTYPE =c RELATIVE, !OBJ CASE =c ACC; |PPn_obj: ^OBJ = !, !] Sdecl: ^!=!;

Frequency: -

Description: This rule shows the relative clause that modifies NP.

c-structure: Relative Clause is a clause having a complete sentence modifying NP. The sentence can be transitive or intransitive. Relative clause is introduced by relative pronoun (optional in some cases).

f-structure: If a relative pronoun is followed by a VP, relative pronoun itself is acting as the SUBJ of the clause as given in production 1. If relative pronoun is followed by an S having a mono transitive verb, the relative pronoun acts as the OBJ as given in production 2. The relative pronoun in this case is optional as 'whom', 'that' and 'which' can be omitted. If an adverb introduced the relative clause, it acts as an adjunct in the f-structure as given in production 3. When PP precedes S production, PP acts an oblique/ object depending on the semantics of PP and the sub-categorization frame of verb as given in production 4.

Examples:

- 1) The bank that has the best interest rates. (modeled by rule 1)
- 2) The key (which) I borrowed is broken. (modeled by rule 2)
- 3) The year when she graduated was lucky for her (modeled by rule 3)
- 4) The girl to whom I gave a book is cooperative (modeled by rule 4)
- 5) the girl on whom I relied betrayed me (modeled by rule 4)



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



Rule Status: Active

Reference: [1] Pam Peters, "The Cambridge English Guide to Usage"

Related Rules: EGR119, EGR139, EGR146

Related POS: EPOS103, EPOS015

Replaces: - EGR113

Reason: - change of relative pronoun occurrence to optional in production 2 and change of ADV_TYPE in production 3 to SEM_TYPE

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis1: Relative clause serves either to define, or to describe and evaluate the noun to which it is attached. It stands next to it, even if it delays the predicate of the main clause.

For example:

The radio that we bought at the market has never given any trouble

Relative clauses are often introduced by of the *relative pronouns* such as *that, which, who etc.* In certain stylistic and grammatical circumstances there may be no pronoun at all, the so-called “zero relative”.

Except in the most formal style, the pronoun can be omitted from relative clauses of which it's the *object* as given below.

The radio we bought at the market has never given any trouble

But when the relative pronoun is the *subject* of the relative clause. Its almost always expressed, whatever the style.

The radio that came from the market has never given any trouble

Some relative clauses are linked to the main clause by adverbs such as *when, where, why*:

I remember the time when we made marshmallow

You remember the place where we went

They remember the reason why we looked so strange

The wh-adverbs act as relaters of the second clause to a noun of time, place or reason in the main clause. In less formal styles, the relative adverbs can be replaced by *that*, as in

I remember the time that we made marshmallow

And even omitted altogether:

I remember the time we made marshmallow

Sentence relative which relate to the whole preceding clause not to any noun in it:

They wanted to go home by ferry, which I thought was a good idea

Sentence relatives are always prefaced by *which*.

[1, p. 468]

Result: The above analysis was finalized.

1. Linking the complementizer (relative pronoun) of relative clause to the head of NP will be done once functional uncertainty is added to the parser. Using that functionality, agreement between relative clause and head noun will be made easier.
*Example: *The girl which I like is good*
2. Unable to parse 'She is the girl who is good'. But the sentence 'She is the girl who is good and sleeping' is parsed correctly. Have to look for the reason.
3. Sentence relatives as discussed in the above analysis will be analyzed.
4. Sometimes a relative clause has optional relative pronoun as mentioned above. This will be taken care of in the future.
5. Topicalization will be covered when parser will be enhanced as given below:
Example: The girl whom I gave the book to is sleeping



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	28 th December 2004	0.1	Created
Zunaira Malik	6 th July, 2005	0.1	Created
Zunaira Malik	25 th July, 2005	0.2	Made comma optional in Pro-1 and made ADJP into a SET in Pro-3

Rule ID: EGR234

Rule Syntax:

ADJPmain -> (neg) ADJP [(comma) ADJP]* coord_conj (neg) ADJP
ADJPmain -> correlative ADJP coord_conj ADJP
ADJPmain -> ADJP

Rule Functional Description:

ADJPmain -> (neg: ^ !=!;)ADJP: ! \$ ^; [(comma:;) ADJP: ! \$ ^;]* coord_conj: ^ CONJ_FORM = ! CONJ_FORM;
(neg: ^ !=!;)ADJP: ! \$ ^, ^ADJ_FORM = !ADJ_FORM;

- 1) ADJPmain -> correlative: ^ !=!; ADJP: ! \$ ^; coord_conj: ^ CONJ_FORM = ! CONJ_FORM, ^CONJ_FORM = ^_ALLOWED_CONJ; ADJP: ! \$ ^;
- 2) ADJPmain -> ADJP: ! \$ ^;

Frequency: -

Description: This rule combines adjective phrases through coordinate conjunctions and correlatives.

c-structure: An adjective phrase can combine with another adjective phrase using a coordinate conjunction and correlatives. It can also take the negative “not” before the adjective in a sentence with coordinate conjunctions.

f-structure: The adjective phrases occurring before the coordinate conjunction can be more than one, separated by a comma in between which is optional (silky brown hair). However, in the case of correlatives no comma is currently allowed.

Example:

1. She has a blue or pink frock. (Production 1)
2. She has neither blue nor pink frock (Production 2)
3. She is good and not happy. (Production 1)
4. She is not good and not happy. (Production 1)



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

Rule Status: Active

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules:

Related POS: EPOS117, EPOS122

Replaces: EGR147

Reason:

Replaced by: -



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



Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT ([COMMA CAT]+ (COMMA)) CONJ CAT}$$

Here, the CAT (category) is ADJP (Adjective phrase). This rule combines adjective phrases with coordinate conjunctions.

Example: i) She wants to hide or to leave unnoticed.
ii) She wants to sit and to eat slowly.
iii) She doesn't want to run but to drive home.

Coordination of ADJP might be possible using correlative conjunctions. In that case there is an endorsing item at the beginning of the phrase which takes only a particular coordinator with it. For more information refer to the POS document for correlative.

Example: i) She is either good or bad.
ii) She is neither good nor bad.
iii) She is both good and bad.

Result: We decided on the above analysis.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	6 th July, 2005	0.1	Created

Rule ID: EGR235

Rule Syntax:

ADVPmain_premod -> ADVP (coord_conj)(ADVP)

ADVPmain_postmod -> ADVP (coord_conj)(ADVP)

Rule Functional Description:

- 1) ADVPmain -> ADVP:! \$ ^ADJUNCT ADV V_PREMOD, !ADV_TYPE =c V_MOD; (coord_conj: ^ ADJUNCT ADV V_PREMOD CONJ_FORM = !CONJ_FORM, ! CONJ_FORM =c 'AND'); (ADVP:! \$^ ADJUNCT ADV,!ADV_TYPE =c V_MOD ;).
- 2) ADVPmain -> ADVP:! \$ ^ADJUNCT ADV V_POSTMOD, !ADV_TYPE =c V_MOD; (coord_conj: ^ ADJUNCT ADV V_POSTMOD CONJ_FORM = !CONJ_FORM, ! CONJ_FORM =c 'AND'); (ADVP:! \$^ ADJUNCT ADV,!ADV_TYPE =c V_MOD ;)

Frequency: -

Description: this rule is used to denote the combination of adverbs with the conjunction AND before and after a verb occurring in a sentence.

c-structure: a main ADVP phrase can contain a single adverb or a combination of adverbs separated by the conjunction AND. This may occur before or after a verb in a sentence.

f-structure: The first production represents a adverbial phrase which is made into an ADJUNCT, then an ADVERB and lastly defined as a pre-modifying adverbial phrase(V_PREMOD). A check is placed for the adverb to be of V_MOD type, that is that is should modify a verb. Than comes the coordinate conjunction AND, which is optional and if occurs than is placed within the ADJUNCT ADV V_PREMOD set. The occurrence of a conjunction is followed by another adverbial phrase which is similar to the initial ADVP and contains a check for V_MOD type adverb.

The second production is similar to the first one with the exception of the V_POSTMOD label. These labels differentiate between the two ADVPmain productions because we need to identify them for correct translation into Urdu which is different due to the POST or PRE occurrence of an adverbial phrase.

Example:

- 1) She is happily and quickly eating. (V_PREMOD with conjunction)
- 2) She is eating quickly and skillfully. (V_POSTMOD with conjunction)
- 3) She is finally eating. (V_PREMOD)



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



4) She is eating quickly. (V_POSTMOD)

Rule Status: Active

Reference:

Related Rules: EGR215,

Related POS: EPOS117

Replaces: -

Reason: - addition in release 5 due to VPact_main requirement for ADVPs with conjunctions before and after verbs.

Replaced by: -



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

Analysis:

Future Work:

1) Need to incorporate this in the grammar as right now there are problems with the parser and it does not run the VPact_main production when it goes through the ADVPmain production to get to the ADVP. Right now, this production has been put flatly within the VPact_main rule. The present and the future form of the rule is given in **advanced grammar rules.txt**

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	16 th September, 04	0.1	Created
Kiran Khurshid	9 th December, 04	0.2	Added negative determiner to the predicative adjective
Zunaira Malik	10 th July, 05	0.3	Removal of end PPnmain and VPinfmain
Zunaira Malik	19 th July, 05	0.4	Addition of NPTYPE constraint in NPmain
Shanza Nayyer	23 rd July, 05	0.5	Addition of _VERB_FORM to NPmain

Rule ID: EGR236

Rule Syntax: Following is the constituent description of the rule.

VPpredlink -> [[(neg_det) ADJPmain] | NPmain | PPnmain]

Rule Functional Description: Following are the functional specifications of the rules.

VPpredlink ->[[(neg_det: ^PREDLINK !=;) ADJPmain: ^PREDLINK !=, ^PREDLINK ADJ_TYPE = PREDICATIVE;] | NPmain: ^PREDLINK !=, ^SUBJ NUM = ! NUM, ~[!NPTYPE =c GERUND] , ^_VERB_FORM =c 'be'; | PPnmain: ^PREDLINK = !;].

Frequency: -

Description: This production gives the detail of the predicative argument of the VP copular construction.

c-structure: When main verb of the form 'be' is used in a VP, it has a different analysis as compared to other main verbs because this verb shows linking of subject with another argument, which is giving information about the subject itself. VPpredlink consists of main verb which is followed by ADJPmain (adjectival phrase), NPmain or PPnmain. This type of construction is known 'copular construction' and the verb of such form is known as 'copular verb'. Note that there are other verbs which act as copular verbs e.g. 'she feels sick'.

f-structure: The linking verb *be* can have APJP, NP or PP as its argument. All these arguments become PREDLINK in f-structure. In NPmain the NUM is to be the same as that of SUBJ to avoid entries like **They are a good boy*. The NPTYPE constraint has been added in NPmain to not allow sentences like "He is working" to be parsed as predlinks as they do not fall under this category. Also the _VERB_FORM constraint is added in NP main to make sure that when ever the linking of copular verb is with NP, the copular verb is of form 'be'. This is to ensure that 'she played football' is not parsed as a copular construction.

Examples:

- 1) I am good. (ADJP as PREDLINK)
- 2) She is no different. (negative determiner with ADJP as PREDLINK)
- 3) I am a good girl. (NP as PREDLINK)
- 4) Cat is on the table. (PP as PREDLINK)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR111, EGR117, EGR144, EGR145

Related POS: EPOS121

Replaces: EGR026

Reason: Made a generic rule

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule:

Analysis: Predicative constructions involve a linking or *copular* verb which has a subject and another argument. The post-verbal category can be of a number of categories e.g. NP, PP, ADJP [1, pg. 69]. These constructions are handled through this rule.

Result: We decided on the above analysis.

Future Work: -



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	2 nd Nov, 04	0.1	Created
Zunaira Malik	19 th July, 05	0.2	Addition of EXPL in PRONTYPE constraint

Rule ID: EGR007

Rule Syntax: Following is the constituent description of the rule.

NPpro → pro

Rule Functional Description: Following are the functional specifications of the rule.

NPpro → pro: ^=!, !PRONTYPE =c {PERS,RELATIVE, EXPL};.

Frequency: -

Description: This rule shows the NP production having pronoun.

c-structure: The NPpro (NPpronoun) goes to simple pronoun. Pronoun is not preceded by any pre-nominal. Pronoun is given a separate POS because of its different features and behavior from noun. For details of pronoun see the relevant POS document.

f-structure: The pronoun is copied to its mother node. Pronoun can be personal or relative.

Examples:

- 1) He is a good boy.
- 2) Ahmad gave her a book.
- 3) The girl to whom I wrote is good.
- 4) It is raining.

Rule Status: Active

Reference: -

Related Rules: -

Related POS: EPOS104

Replaces: EGR121

Reason: addition of EXPL in PRONTYPE constraint

Replaced by: -

Reason: -



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



Analysis: Refer to the related POS document for details of pronoun.

Result: -

Future Work: Following work has been left for future:

When a pronoun and a noun are combined (emphatic pronoun):

Example:

- 1) We students are demanding that the administration give us two hours for lunch.
- 2) The administration has managed to put us students in a bad situation.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	3 rd Nov, 04	0.1	Created
Zunaira Malik	25 th July, 05	0.2	Addition of NNUM constraints

Rule ID: EGR238

Rule Syntax: Following is the constituent description of the rule.
POSTDET -> [NUMBERP|QUANTP]

Rule Functional Description: Following are the functional specifications of the rule:

POSTDET -> [NUMBERP: ^NUMBER =!, ^NNUM = !NUM; | QUANTP: ^QUANT = !, ^NNUM = !NNUM;].

Frequency: -

Description: This rule shows the functional and constituent structure of post- determiner phrase.

c-structure: This production is for the phrases which can occur after the central determiner (DETP). POSTDET can have NUMBERP(Number phrase) or QUANTP(quantifier phrase).

f-structure: Everything in POSTDET goes as a NUMBER to the mother node i.e. PRENOMP.

Examples:

In the following examples the underlined part is the determiner phrase:

- 1) The few books I read are missing. (QUANTP)
- 2) Two books are torn. (NUMBERP)
- 3) Hundreds of people. (NUMBERP)

Rule Status: Active

Reference:

- [1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"
- [2] "A Lexicalized Tree Adjoining Grammar for English", Institute for Research in Cognitive Science, University of Pennsylvania
- [3] <http://webster.commnet.edu/grammar/determiners/determiners.htm>
- [4] B. A. Hockey and Heather Mateyak, "Determining Determiner Sequencing: A Syntactic Analysis for English" University of Pennsylvania.

Related Rules: EGR133, EGR137

Related POS:

Replaces: EGR132

Reason: addition of NUM constraint to NUMBERP and QUANTP

Replaced by: -

Reason: -



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



Analysis: Following is the in-depth analysis of the rule.

Analysis 1:-

Result: -

Future Work: Following issues need to be addressed, and constraints to be included respectively, after further analysis:

The post determiners will be further analyzed so that all possible post determiners along with their agreement properties can be covered.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	16 ⁿ September, 04	0.1	Created
Kiran Khurshid	22 nd December, 04	0.2	Added coordinate conjunctions and subordinate correlatives
Shanza Nayyer	3 rd August, 05	0.3	Sub_ConjP_opt added and correlative in subordinate conjunction (COMP) deleted.

Rule ID: EGR239

Rule Syntax: Following is the constituent description of the rule.

S -> S_coord_conjP
S -> Smain
S -> Smain Sub_ConjP_opt
S -> Smain Sub_ConjP
S -> correlative Smain (comma) Sub_ConjP_opt
S -> Sub_ConjP (comma) Smain
Sub_ConjP -> sub_conj Smain
Sub_ConjP_opt -> (sub_conj) Smain

Rule Functional Description: Following are the functional specifications of the rule.

- 1) S -> S_coord_conjP: ^ = !;.
- 2) S -> Smain: ^ = !, ^CLAUSE_TYPE = !CLAUSE_TYPE;.
- 3) S -> Smain: ^ = !; Sub_ConjP: ^ ADJUNCT = !, !_POSITION = c MIDDLE;.
- 4) S -> Smain: ^ = !; Sub_ConjP_opt: ^ COMP = !;.
- 5) S -> correlative: ^ = !; Smain: ^ = !; (comma:;) Sub_ConjP_opt: ^ ADJUNCT = !, !_POSITION = c MIDDLE, ^_ALLOWED_CONJ = !CONJ_FORM;.
- 6) S -> Sub_ConjP: ^ ADJUNCT = !, !_POSITION = c START; (comma:;) Smain: ^ = !;.
- 7) Sub_ConjP -> sub_conj: ^ = !; Smain: ^ = !;.
- 8) Sub_ConjP_opt -> (sub_conj: ^ = !, !_POSITION = c MIDDLE;) Smain: ^ = !;.

Frequency: -

Description: This rule shows the subordinate coordination at sentence level.

c-structure: These productions show subordinate and coordinate conjunctions at sentence level. The first production uses the coordinate conjunction phrase rule. The second production allows a sentence without a conjunction to parse. The third and fourth phrase caters subordinate conjunction at the middle of two sentences. The fifth and sixth production allows correlatives with the subordinate conjunctions. The seventh production caters subordinate conjunction at the start of the sentence. In such case, the two sentences are separated by comma. The eighth production is the subordinate clause which is being used by the above five productions.



f-structure: In (Sub_ConjP) subordinate conjunction phrase, the sub-ordinate clause becomes the COMP in *f-structure* when it is sub-categorized by the verb (as in production 4) and otherwise becomes ADJUNCT (as in production 3). Right now the sentence initial subordinate conjunction is made an ADJUNCT (as in production 6).

Examples:

- 1) She came home and he gave her the book. (Production 1)
- 2) I would be going to the party. (Production 2)
- 3) Girl came home after/as I went to school (Production 3)
- 4) She took my money that/because I left on the table (Production 3)
- 5) She slept while I went to school (Production 3)
- 6) She said that I should go to school (Production 4)
- 7) As the strength of the defenders failed, so the courage of the attackers grew. (Production 5)
- 8) While I went to school, she slept (Production 6)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

[2] www.grammar.englishclub.com

[3] Radolph Quirk, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik, "A Comprehensive Grammar of the English Language"

Related Rules: EGR140

Related POS: EPOS117, EPOS118, EPOS122

Replaces: EGR028

Reason: Coordinate and subordinate conjunction analyzed in greater depth and separate rules made for them

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis1: Conjunctions are words that "join". Conjunctions join two parts of a sentence. They are discussed in detail below:

1. Form

Conjunctions have three basic forms:

Single Word

Example: and, but, because, although

Compound (often ending with *as* or *that*)

Example: provided that, as long as, in order that

Correlative (which surround an adverb or adjective)

Example: so...that

2. Function

Conjunctions are divided into two basic types.

Coordinating Conjunctions are used to join two parts of a sentence that are grammatically equal. The two parts may be single words or clauses,

Example: i) Jack and Jill went up the hill.

ii) The water was warm but I didn't go swimming.

Subordinating Conjunctions are used to join a subordinate dependent clause to a main clause, for example:

- *I went swimming, **although** it was cold.*

3. Position

Coordinating Conjunctions always come between the words or clauses that they join.

Subordinating Conjunctions usually come at the beginning of the subordinate clause. [2]

Analysis2: The coordination rule must cater for more than two conjuncts as in *they close the door, locked it, and walked to the car*. The rules must be formulated differently, according to whether the non-final conjuncts are separated from one another only by commas, or whether a conjunction is also required. Moreover, the rule must allow for two-part conjunctions, such as *either...or, neither...nor*.

NP coordination often involves number, person, and gender mismatches between the individual conjuncts and the entire coordinated NP.

Two requirements must be met by an analysis of NP coordination. First, it must be possible to assert constraints about and assign values to attributes in both the individual conjuncts and in the f-structure on the coordination as a whole. Second, there must be an algorithm to construct the number, person, and gender values of the coordinated NP from the values of the individual conjuncts.

The type of conjunction involved in the coordinated phrases plays a role in determining agreement. The general rule in English for determining the number, gender, and person of a coordinated NP is as follows:

- 1) If any conjunct is plural, the entire NP is plural;
- 2) If any conjunct is masculine, the entire NP is masculine;
- 3) If any conjunct is first person, the entire NP is first person;
- 4) If there is no first person conjunct and any conjunct is second person, the entire NP is second person; otherwise the NP is third person. [1, p.145]



Analysis 2:

Another kind of embedding occurs when one clause is made a constituent of another clause: the case normally described as SUBORDINATION. Consider the following

Example: i) The weather has been remarkably warm [since we returned from Italy last week].

Clauses which are embedded in other clauses are subordinate clauses (e.g. since we returned from Italy last week). The relation between the two clauses is one of 'part to whole'.

Subordination of clauses is not confined to clauses which are immediate constituents of other clauses. There are also clauses which are constituents of phrases, and which therefore are only indirectly embedded within a larger clause (as the relative clause used as the postnominal modifier)

Embedding gives rise to the theoretical possibility of grammatical units having indefinite length.

Example: i) This is the house [that Jack built].

ii) This is the malt [that lay in the house [that Jack built]].

iii) This is the rat [that ate the malt [that lay in the house [that Jack built]]] [3, pg. 44]

Result: The above analyses are used for the rule.

Future Work:

1) Right now subordinate conjunction is catered partially. Every subordinate conjunction and its semantic behavior will be covered in the later stage.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	28 th December 2004	0.1	Created
Zunaira Malik	9 th August, 2005	0.2	Addition of neg and optional comma, also New Release

Rule ID: EGR240

Rule Syntax:

ADJPmain -> (neg) ADJP [(comma) ADJP]* (coord_conj) (neg) ADJP

ADJPmain -> correlative ADJP coord_conj ADJP

ADJPmain -> ADJP

Rule Functional Description:

- 1) ADJPmain -> (neg: ^=!) ADJP: ! \$ ^; [(comma:;) ADJP: ! \$ ^;]* (coord_conj: ^ CONJ_FORM = ! CONJ_FORM;) (neg: ^ = !;) ADJP: ! \$ ^, ^ ADJ_FORM = ! ADJ_FORM;.
- 2) ADJPmain -> correlative: ^=!; ADJP: ! \$ ^; coord_conj: ^ CONJ_FORM = ! CONJ_FORM, ^CONJ_FORM = ^_ALLOWED_CONJ; ADJP: ! \$ ^;.
- 3) ADJPmain -> ADJP: ^=!;

Frequency: -

Description: This rule combines adjective phrases through coordinate conjunctions and correlatives.

c-structure: An adjective phrase can combine with another adjective phrase using a coordinate conjunction and correlatives.

f-structure: The adjective phrases occurring before the coordinate conjunction can be more than one, separated by a comma in between. However, in the case of correlatives no comma is currently allowed.

Example:

1. She has a blue or pink frock. (Production 1)
2. She has neither blue nor pink frock (Production 2)
3. The girl is not good and not happy (Production 1)
4. Silky Brown hair (optional comma , Production 1)



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

Rule Status: Active

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules:

Related POS: EPOS117, EPOS122

Replaces: EGR147

Reason: Addition of neg and optional comma, also New Release

Replaced by: -



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



Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT} ([\text{COMMA CAT}]^+ (\text{COMMA})) \text{ CONJ CAT}$$

Here, the CAT (category) is ADJP (Adjective phrase). This rule combines adjective phrases with coordinate conjunctions.

Example: i) She wants to hide or to leave unnoticed.
ii) She wants to sit and to eat slowly.
iii) She doesn't want to run but to drive home.

Coordination of ADJP might be possible using correlative conjunctions. In that case there is an endorsing item at the beginning of the phrase which takes only a particular coordinator with it. For more information refer to the POS document for correlative.

Example: i) She is either good or bad.
ii) She is neither good nor bad.
iii) She is both good and bad.

Result: We decided on the above analysis.

Future Work:



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	15 th Feb, 05	0.1	Created
Zunaira Malik	9 th Aug, 05	0.2	New Release

Rule ID: EGR241

Rule Syntax: Following is the constituent description of the rule.

PARENPN -> open_paren [ADVP | NPmain | NUMBERP] close_paren

Rule Functional Description: Following are the functional specifications of the rule.

PARENPN -> open_paren: ^=! ; [ADVP: ^=! ; | NPmain: ^=! ; | NUMBERP: ^=! ;] close_paren: ^=! ;.

Frequency: -

Description: This rule shows the parenthetical phrase.

c-structure: A parenthetical phrase starts with an open_paren (symbol) and ends with a close_paren. Within these symbols an adverbial phrase, or a noun phrase or a number phrase must occur.

f-structure: The parenthetical phrase becomes a part of the noun's ADJUNCT-PAREN.

Examples:

- 1) The energy-recovery experiment (ERX) successfully completed.
- 2) There are three (3) projects.
- 3) 6-county area (about 60 by 60mi)

Rule Status: Active

Reference:

- (1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR105, EGR111, EGR133

Related POS: -

Replaces: -EGR149

Reason: - New Release

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis1: Parentheticals are introduced by a special c-structure rule which includes the required punctuation and allows a limited number of constituents within it, e.g., NPs, imperatives, APs. The PARENPN constituent appears in selected c-structure positions depending on the language. Ideally, almost any constituent can be followed by a parenthetical; however, in practice this allows for extensive ambiguity, and so the parentheticals appear only in select positions, as dictated by the corpus at hand. [1, p.147]

The PARENPN constituent corresponds to an ADJUNCT-PAREN feature in the f-structure, thus rendering it distinct from a plain ADJUNCT. ADJUNCT-PAREN is defined not to be a set which restricts parentheticals to one per constituent. However, some corpora may allow more than one parenthetical per constituent, which would then require ADJUNCT-PAREN to be defined as a set. Note that constituents which normally do not appear with adjuncts, such as proper names and pronouns, can freely occur with ADJUNCT-PARENS. [1, pp. 147-148]

Result: We decided on the above analysis.

Future Work:

- Currently the parenthesis allow only adverbial phrase, noun phrase or a number phrase, as this is the most common possible combinations in the corpus. However, later it may allow more things.



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	4 th Nov, 04	0.1	Created
Zunaira Malik	9 th Aug, 05	0.2	New Release

Rule ID: EGR242

Rule Syntax: Following is the constituent description of the rule.
PREDETQUANTP -> quant p

Rule Functional Description: Following are the functional specifications of the rule.
PREDETQUANTP -> quant: ^!=!, ^QTYPE = GEN; p: ! PFORM =c 'of', ^PFORM = 'of';.

Frequency: -

Description: Following is the constituent structure and functional annotation description of the quantifier phrase occurring before a determiner (PREDETQUANTP).

c-structure: This rule takes quantifier with preposition 'of' before determiner

f-structure: Quantifier with preposition goes as the SPEC of the PRENOMP. Quantifier has QTYPE as genitive if it can occur with preposition. There is a constraint that after quantifier only preposition 'of' can come, and no other preposition can come.

Examples:

- 1) All of the trees are dying.
- 2) Some of the girls are good.

Rule Status: Active

Reference: -

Related Rules:

Related POS: EPOS110, EPOS104

Replaces: EGR136

Reason: New Release

Replaced by: -

Reason: -



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



Analysis: -

Result: -

Future Work: -



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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	11 th June, 04	0.1	Created
Zunaira Malik	9 th Aug, 05	0.3	New Release

Rule ID: EGR243

Rule Syntax: Following is the constituent description of the rule.

QUANTP -> (ADVP) quant

Rule Functional Description: Following are the functional specifications of the rules.

QUANTP -> (ADVP: ^ADJUNCT ADV= !;) quant: ^=!, ^QTYPE = NONGEN;.

Frequency: -

Description: This production gives the detail of the Quantifier Phrase.

c-structure: It shows the variation of quantifier phrase. A quantifier phrase which occurs after the DET as a POSTDET can have adverb as a modifier.

f-structure: If adverb occurs before a quantifier, it becomes a modifier and is represented by ADJUNCT of quantifier in the f-structure.

Examples:

- 1) Many girls went to the concert.
- 2) Too many boys are attending this training.
- 3) Some few books were torn.
- 4) A few books are missing.

Rule Status: Active

Reference: -

Related Rules: EGR105

Related POS: EPOS110

Replaces: EGR137

Reason: New Release

Replaced by: -

Reason: -



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

Analysis: Please refer to EPOS007 and relevant grammar rule documents for details.

Result: -

Future Work:

1) Some checks need to be added so that the quantifier which acts as a modifier of a quantifier as in production 2 is controlled for erroneous entries like **many several books are missing*.

2) Also not all adverbial intensifiers can modify a quantifier as in example below. It needs to be checked.

Example: i) Very few girls are going.

**Very many girls are going.*



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	15 th April, 04	0.1	Created
Zunaira Malik	9 th August, 05	0.3	New Release

Rule ID: EGR244

Rule Syntax: Following is the constituent description of the rule.

NPgen -> (PRENOMP) Ngen

Ngen -> n apostr

Rule Functional Description: Following are the functional specifications of the rule.

NPgen -> (PRENOMP: ^ = !;) Ngen: ^ = !;.

Ngen -> n: ^ = !; apostr: ^ CASE = GEN;.

Frequency: -

Description: This production gives the detail of the genitive (possessive) construction of NP.

c-structure: Everything that precedes genitive NP is made a specifier/modifier of NPgen which later becomes the specifier of head noun.

NPgen can have the DETP (determiner phrase) followed by NUMBERP (numberP), ADJP (adjective phrase) and Ngen. The Ngen can have the n (noun) and the apostr (apostrophe)

f-structure: The structure of DETP, NUMBERP goes as SPEC of NPgen and the structure of ADJP becomes an ADJUNCT of NPgen. Noun genitive is distinguished from simple noun by the feature CASE which is genitive as given in the rule above.

Examples:

- 1) The three good books' first page
- 2) Three times the two big tables' width

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR130, EGR133, EGR114

Related POS: EPOS105, EPOS109

Replaces: EGR138

Reason: New Release

Replaced by: -

Reason: -



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

Analysis: Following is the in-depth analysis of the above rule:

Analysis: Although prenominal genitives pattern like articles and quantifiers as far as the syntax is concerned, they do differ in terms of complexity. Genitives give rise to more complex specifiers:

Example: The driver's dog

The genitive NP *the driver's* is embedded under the SPEC feature of *dog*, indication that the entire NP (*the driver's*) serves as a specifier of *dog*. Since the *driver's* is itself an NP with a specifier (the article *the*), a SPEC feature is embedded under the higher one. Note that any further material in the genitive NP would also be embedded under the SPEC feature of *dog*.

Example: the nice driver's dog

The adjective *nice* would be analyzed as an adjunct modifying *driver* and be placed inside the SPEC feature modifying *dog*. [1, p. 103]

Future Work:

1) Catering multiple genitives

Example: John's this uncle's dog

2) Incase of multiple genitives, it should be kept in mind that second genitive cannot be preceded by an article.

*Example: *John's the uncle's dog*



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	4 th Nov, 04	0.1	Created
Zunaira Malik	9 th Aug, 05	0.2	New Release

Rule ID: EGR245

Rule Syntax: Following is the constituent description of the rule.

Sdecl -> (Start_AdjunctCl) NPmain VPmain

Rule Functional Description: Following are the functional specifications of the rule.

Sdecl -> (Start_AdjunctCl: ^ ADJUNCT = !;) NPmain: ^SUBJ = !, ^SUBJ CASE = NOM; VPmain: ^!= !;.

Frequency: -

Description: This rule shows the sentence level production of declarative sentence of English Grammar.

c-structure: This rule (declarative sentence) consists of a Noun Phrase (NP) and a Verb Phrase (VP). Adverbial Phrase (ADVP) can optionally come as sentence modifier.

f-structure: The structure contained in NP is equal to the SUBJ structure of mother node S. A constraint is applied which makes sure that the subject case of NP is nominative. All that is contained in VP is equivalent to the mother node S. ADVP becomes ADJUNCT.

Examples:

- 1) Aslam is sleeping well.
- 2) He went to school everyday.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR111, EGR146, EGR105

Related POS:

Replaces: EGR139

Reason: New Release

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis: As English is SOV ordered language and is not free ordered (unlike Urdu), hence NP is considered as a SUBJ preceding VP in declarative sentences. Some languages allow practically any order of constituents in declarative clauses. However many languages have relatively strict c-structure requirements on root declarative. For example, English and French require the subject to precede the VP in simple declaratives, and have further requirements on the internal structure of VP. [1, p. 18]

Example:

Aslam slept.

[[Aslam]_{NP} [slept]_{VP}]_S

Result: Above analysis was finalized for the rule.

Future Work: Further functional annotations might be added in the later stage.



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	4 th Nov, 04	0.1	Created
Zunaira Malik	9 th Aug, 05	0.2	New Release

Rule ID: EGR246

Rule Syntax: Following is the constituent description of the rule.
Smain -> ([PPnmain | VPparticiple]) [Sint | Sdecl | Simp].

Rule Functional Description: Following are the functional specifications of the rule.
Smain -> ([PPnmain: ^ ADJUNCT PREP = !; | VPparticiple: ^ ADJUNCT PARTICIPLE = !;]) [Sint: ^=!, ^CLAUSE_TYPE = INTERROGATIVE, ^_FLAG = INTERROGATIVE; | Sdecl: ^=!, ^CLAUSE_TYPE = DECLARATIVE, ^_FLAG = DECL; | Simp: ^=!, ^CLAUSE_TYPE = IMPERATIVE;].

Frequency: -

Description: This rule shows the root level production of a sentence of English Grammar.

c-structure: A sentence may be an interrogative sentence (Sint), a declarative sentence (Sdecl) or an imperative sentence (Simp). It can have a PP or a Participle phrase in the beginning as an adjunct.

f-structure: All the features of the Sint, Sdecl or Simp are passed on to the mother node of Smain, where a feature CLAUSE_TYPE is introduced to contain the type of sentence being formed. FLAG is used to control over generation.

Examples:

- 1) Aslam is sleeping.
- 2) Did he go to sleep?
- 3) Give him the book.
- 4) Crying he went to school.
- 5) In the morning I had breakfast.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

[2] <http://englishplus.com/grammar/glossary.htm>

Related Rules: EGR115, EGR116, EGR108, EGR139

Related POS:

Replaces: EGR140

Reason: New Release

Replaced by: -



Analysis: Following is the in-depth analysis of the rule.

Analysis: A sentence may be of the following three types:

1. **Declarative Sentence:** A declarative sentence states an idea. It does not give a command or request, nor does it ask a question. A declarative sentence usually ends in a period, though it may end in an exclamation point. [2]
 - a. *Example:* She goes to school.
2. **Imperative Sentence:** An imperative sentence asks, requests, or commands someone to do something. An imperative sentence drops the subject. Sometimes when simply the verb of an imperative sentence is referred to, it is said to be in the imperative mood [2].
 - a. *Example:* Go away!
3. **Interrogative Sentence:** An interrogative sentence is a sentence that asks a direct question. It is punctuated with a question mark at the end. In English an interrogative sentence normally changes the word order so that the verb or part of the verb comes before the subject.
 - a. *Example:* Do you like this?

Result: The above analysis was finalized for the rule.

Future Work: Exclamatory sentences will be done later. They may be covered in any one of the above mentioned types of sentences.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	16 th September 2004	0.1	Created
Zunaira Malik	9 th August, 2005	0.2	New Release

Rule ID: EGR247

Rule Syntax:

S_coord_conjP -> (correlative) [Smain [(comma) Smain]* coord_conj]] Smain

Rule Functional Description:

S_coord_conjP -> (correlative: ^ = !;) [Smain: !\$ ^, ^CLAUSE_TYPE=!CLAUSE_TYPE; [(comma:;) Smain: !\$ ^, ^CLAUSE_TYPE=!CLAUSE_TYPE;]* coord_conj: ^ CONJ_FORM = !CONJ_FORM, [!CONJ_FORM =c 'AND' || !CONJ_FORM =c 'OR'];] Smain: !\$ ^, ^CLAUSE_TYPE=!CLAUSE_TYPE;.

Frequency: -

Description: “and” and “or” are coordinate conjunctions allowed between any of the three types of sentences: declaratives, interrogatives and imperatives. This rule is used to parse the sentences coordinated by “and” and “or” only.

c-structure: Smain represents declarative, imperative and interrogative sentences. It can be coordinated using coordinate conjunctions. If more than one sentence is to be coordinated, it is separated by a comma until the second last sentence comes. Then the coordinate conjunction “and” or “or” occurs before the last sentence occurs.

f-structure: The only coordinate conjunction allowed are “and” and “or”, hence the constraints are placed on CONJ_FORM. These can be paired with correlatives like “either”, “neither” and “both”.

Example:

Interrogative sentences: What is your name AND what do you do to earn a living?
Would you like me to do this for you OR are you not comfortable this way?

Declarative sentences: He was looking at the bird AND it flew away.
She may do the work OR I will complete it.

Imperative sentences: Run AND catch the bus.
Get dressed OR wait.



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

Rule Status: Active

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR140

Related POS: EPOS117

Replaces: EGR141

Reason: New Release

Replaced by: -



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

Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT} ([\text{COMMA CAT}]^+ (\text{COMMA})) \text{ CONJ CAT}$$

Here, the CAT (category) is a clause.

All three sentence types, namely declarative, interrogative and imperative, can combine using a coordinate conjunction. This rule allows only the same type of sentences to combine. This is ensured by the unification of the feature `CLAUSE_TYPE`.

Example: i) What is your name AND what do you do to earn a living? (Interrogative sentences)

ii) Would you like me to do this for you OR are you not comfortable this way? (Interrogative sentences)

iii) He was looking at the bird AND it flew away. (Declarative sentences)

iv) She may do the work OR I will complete it. (Declarative sentences)

v) Run AND catch the bus. (Imperative sentences)

vi) Get dressed OR wait. (Imperative sentences)

Result: We decided on the above analysis.

Future Work:

- At the moment any number of sentences can join using comma and coordinate conjunction. It will be seen later what the limit of sentence coordination is.
- If the semantics of coordination is ignored (or generalized) then all the three separate rules for sentence coordination can and might be merged.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	16 th September 2004	0.1	Created
Zunaira Malik	9 th August, 2005	0.2	New Release

Rule ID: EGR248

Rule Syntax:

S_coord_conjP -> Sdecl coord_conj Smain

Rule Functional Description:

S_coord_conjP -> Sdecl: !\$^, ; coord_conj: ^ CONJ_FORM = !CONJ_FORM, [!CONJ_FORM =c 'BUT' || !CONJ_FORM =c 'SO']; Smain: ! \$ ^,.

Frequency: -

Description:

c-structure: The coordinate conjunctions “but” and “so” can coordinate a declarative sentence with any other type of sentence, i.e. declarative sentence, imperative sentence or an interrogative sentence. Note that in such sentences multiple declarative sentences are not allowed before the coordinate conjunction.

f-structure: The only coordinate conjunction allowed are "but" and "so", hence the constraints are placed on CONJ_FORM.

Example:

Declaratives with declaratives:

She is going to school but I would not like her to go.

Declaratives with interrogatives:

She was upset so we stopped her from going.

She was looking good but why was she upset?

Declaratives with imperatives:

She is not interested so why are you disturbing her?

Anything can happen but be ready to face it.

The bus is leaving so run.

Rule Status: Active

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR139, EGR140

Related POS: EPOS117

Replaces: EGR142

Reason: New Release

Replaced by: -



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

Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT} ([\text{COMMA CAT}]^+ (\text{COMMA})) \text{ CONJ CAT}$$

Here, the CAT (category) is a clause.

It is observed that declarative sentences can combine with other declarative, imperative and interrogative sentences using the coordinate conjunction "and", "or" and "so". Hence, this rule combines declarative sentences with Smain (which represents all the three types of sentences just mentioned).

Example: (As given above)

Result: We decided on the above analysis.

Future Work:

- If the semantics of coordination is ignored (or generalized) then all the three separate rules for sentence coordination will be merged.
- To allow the following examples, the rule may later be altered by allowing multiple declaratives before the coordinate conjunction "so":
 - I am happy, you are happy so we all are happy.
 - I am listening; you are listening so we all are listening.
 - She is interested, she's not even listening so why are you disturbing her.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	16 th September 2004	0.1	Created
Zunaira Malik	9 th August, 2005	0.2	New Release

Rule ID: EGR249

Rule Syntax:

S_coord_conjP -> Simp coord_conj Sdecl

Rule Functional Description:

S_coord_conjP -> Simp: !\$^; coord_conj: ^ CONJ_FORM = !CONJ_FORM, [!CONJ_FORM =c 'AND' || !CONJ_FORM =c 'OR' || !CONJ_FORM =c 'BUT']; Sdecl: ! \$ ^;.

Frequency: -

Description:

c-structure: The imperative sentences can coordinate with declarative sentences using the coordinate conjunction "and", "or" and "but". Note that in such sentences multiple imperative sentences are not allowed before the coordinate conjunction.

f-structure: The only coordinate conjunction allowed are "and", "or" and "but", hence the constraints are placed on CONJ_FORM.

Example:

1. Get dressed and you'll feel much better.
2. Run or you'll miss the bus.
3. Stand but you'll get tired after some time.

Rule Status: Active

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR108, EGR139

Related POS: EPOS117

Replaces: EGR143

Reason: New Release

Replaced by: -



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



Analysis:

Analysis 1: The basic approach to constituent coordination in LFG is as follows [1]:

$$\text{SCCOORD (CAT)} = \text{CAT (} [\text{COMMA CAT}]^+ (\text{COMMA}) \text{) CONJ CAT}$$

Here, the CAT (category) is a clause.

It is observed that imperative and declarative sentences can combine using any three of the coordinate conjunctions, namely, "and", "or" and "but". This rule is used to parse such sentences.

Example: i) Sit but you'll have to stand after some time.
ii) Run but you won't be able to catch the bus.
iii) Get dressed and you'll feel much better.
iv) Run or you'll miss the bus.

Result: We decided on the above analysis.

Future Work:

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	16 th September, 04	0.1	Created
Zunaira Malik	9 th August, 05	0.2	New Release

Rule ID: EGR250

Rule Syntax: Following is the constituent description of the rule.

VP-> {VPpred_main1 | VPpred_main2} (NP) VPpredlink (VPadjunct)

Rule Functional Description: Following are the functional specifications of the rules.

VP -> [VPpred_main1: ^ = !;|VPpred_main2: ^ = !;] (NP: ^ OBJ = !;) VPpredlink: ^ = !; (VPadjunct: ^ = !;).

Frequency: -

Description: This production gives the detail of the VP production having a copular construction.

c-structure: It shows a copular construction of verb phrase. This contains a auxiliaries and main verb (VPpred_main1 and VPpred_main2), followed by the predicate argument (VPpredlink).

f-structure: The copular verb become the head of the sentence and the predicate argument becomes the PREDLINK in f-structure.

Examples:

- 1) I am good.
- 2) I am being a very good girl.
- 3) She painted the door green.

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR103, EGR104

Related POS: -

Replaces: EGR102

Reason: New Release

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule:

Analysis: Predicative construction involves a linking or *copular* verb which has a subject and another argument, as in the example below:

Example: i) The beacon is on the roof.
ii) The tractor is red.

The post-verbal argument can be of a number of categories, e.g., NP, PP, AP etc. Due to the semantic relationship between the subject and the phrase after linking verb, these verbs are given special sub-categorization frames. Traditionally, this has been done by having the post-verbal phrase 'be' an XCOMP whose subject is controlled by the linking verb's subject. However, a new analysis, termed the PREDLINK analysis, is used. Under both approaches, linking verbs may have their own c-structure category and their own VP rule which allows the post-verbal NP, AP, and PP to be assigned the appropriate grammatical function. [1, p. 69]

Result: The above analysis was finalized.

Future Work:

1) The sentence like *the chicken is cooked* will create ambiguity. It will be parsed having *is* as the main verb and *cooked* as an ADJP and will also be parsed through the production having *cooked* as the main verb. This kind of ambiguity will be catered in the later stage using some statistical method.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	6 th September 2004	0.1	Created
Kiran Khurshid	22 nd September 2004	1.0	Approved by SA
Zunaira Malik	9 th august, 2005	1.2	New Release

Rule ID: EGR251

Rule Syntax:

AuxPinf -> aux

AuxPinf -> aux aux

AuxPinf -> aux aux aux

Rule Functional Description:

AuxPinf -> aux: ^_AUX1 = !, ^_MAUX = !, ^_AUX1 _MORPH_FORM = TO_INF, ^_MAUX TNS_ASP _JUNK = 'check';.

AuxPinf -> aux: ^_AUX1 = !, ^_AUX1 _MORPH_FORM = TO_INF; aux: ^_MAUX = !, !_MORPH_FORM = c ^_AUX1 _ALLOWED_FORM, ^_MAUX TNS_ASP _JUNK = 'two_aux_in_auxpinf';.

AuxPinf -> aux: ^_AUX1 = !, ^_AUX1 _MORPH_FORM = TO_INF; aux: ^_AUX2 = !, !_MORPH_FORM = c ^_AUX1 _ALLOWED_FORM; aux: ^_MAUX = !, !_MORPH_FORM = c ^_AUX2 _ALLOWED_FORM;.

Frequency: -

Description: This rule is for all the combinations of auxiliaries that can occur in a to-infinitive clause.

c-structure: An infinitive auxiliary phrase can allow zero to three auxiliaries.

f-structure: The values starting with “_” are junk values only to check wellformedness of the phrase.

Example:

- (1) You need to run to catch the bus
- (2) We are glad to have invited you.



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



Rule Status: Active

Reference:

- (1) Pam Peters, "The Cambridge Guide to English Usage"
- (2) Quirk et al, "A Comprehensive Grammar of the English Language"

Related Rules: -

Related POS: EPOS101

Replaces: - EGR107

Reason: - New Release

Replaced by: -



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

Analysis:

Analysis 1: The basic nonfinite forms of verbs such as (to) ask, (to) go, (to) decide are called **infinitives** [1]. There are 5 basic forms of full verbs given on [pp. 96-97, 2]. The base form of the verb has two nonfinite forms: 1) the bare infinitive and 2) the *to*-infinitive. This rule is used to parse the auxiliaries that can occur in *to*-infinitive phrases.

If we relate the structure of the nonfinite verb phrase to that of the finite verb phrase, we can tabulate the eight possible combinations as follows [pg. 153, 2]:

Type of verb phrase	Infinitive	Type
Simple	to examine	
Complex	to have examined	B
	to be examining	C
	to have been examining	D
	to have been examined	BC
	to be being examined	CD
	to have been being examined (rare)	BCD

Where Type B, C and D mean [pg. 152, 2]:

Type B: perfective aux (HAVE) + -ed participle

Type C: progressive aux (BE) + -ing participle

Type D: passive aux (BE) + -ed participle

To cover all these combinations of auxiliaries the above three productions are used. Note that “have”, “been”, “being” and “be” are all auxiliaries named as “aux”. The `_AUX1` and `_AUX2` are temporary structures to check the well-formedness of the sentence. Similarly, all other values starting with “_” are used to check well-formedness.

Result: We decided on the above analysis.

Future Work:



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	7 th September 2004	0.1	Created
Kiran Khurshid	22 nd September 2004	1.0	Approved by SA
Kiran Khurshid	20 th December 2004	1.1	Added negation and the optional auxiliary "do" for emphasis
Zunaira Malik	9 th August, 2005	1.2	New Release

Rule ID: EGR252

Rule Syntax:

Simp -> [(aux) [VPimp_act | VPimp_pass]] (VPimp_oblig) (VPimp_adjunct)
Simp -> [aux neg [VPimp_act | VPimp_pass]] (VPimp_oblig) (VPimp_adjunct)
Simp -> (aux) VPimp_pred
Simp -> aux neg VPimp_pred

Rule Functional Description:

1) Simp -> [(aux: !_AUX_FORM =c 'do', !_MORPH_FORM =c BARE, ^EMPHASIS = POS;) [VPimp_act: ^=!, ^SUBJ PRED = 'PRO', ^SUBJ PRONTYPE = NULL; | VPimp_pass: ^=!, ^SUBJ PRED = 'PRO', ^SUBJ PRONTYPE = NULL;]] (VPimp_oblig: ^ = !;) (VPimp_adjunct: ^ = !;).

2) Simp -> [aux: ^_AUX2 = !, !_AUX_FORM =c 'do', !_MORPH_FORM =c BARE; neg: ^=!; [VPimp_act: ^=!, ^SUBJ PRED = 'PRO', ^SUBJ PRONTYPE = NULL; | VPimp_pass: ^=!, ^SUBJ PRED = 'PRO', ^SUBJ PRONTYPE = NULL;]] (VPimp_oblig: ^ = !;) (VPimp_adjunct: ^ = !;).

3) Simp -> (aux: ^ = !, !_AUX_FORM =c 'do', !_MORPH_FORM =c BARE, ^EMPHASIS = POS;) VPimp_pred: ^=!, ^SUBJ PRED = 'PRO', ^SUBJ PRONTYPE = NULL;.

4) Simp -> aux: ^ = !, !_AUX_FORM =c 'do', !_MORPH_FORM =c BARE; neg: ^=!; VPimp_pred: ^=!, ^SUBJ PRED = 'PRO', ^SUBJ PRONTYPE = NULL;.

Frequency: -

Description: This rule is used to parse the imperatives.

c-structure: An imperative sentence can contain an active verb phrase or a passive verb phrase followed by a prepositional phrase or an infinitive verb phrase allowed only in imperative sentences.

f-structure: The prepositional phrase can either become an oblique of the verb phrase or an adjunct, depending on the sub-categorization of the verb. The infinitive verb phrase is optional so it becomes a part of the adjunct of the verb phrase. The subject of an imperative sentence is implicitly understood to be a second person pronoun. The auxiliary "do" introduces an EMPHASIS feature to denote emphasis in the sentence. The negative particle "not" introduces an _NEG_MARKER feature to denote negation. See the analysis below for further details.



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

Example:

- (1) Go home (Production 1)
- (2) Go home to study in the morning (Production 1)
- (3) Do go home (Production 1)
- (4) Do not go home (Production 2)
- (5) Get dressed (Production 3)
- (6) Be guided by me (Production 3)
- (7) Be quiet (Production 3)
- (8) Do Get dressed (Production 3)
- (9) Do be guided by me (Production 3)
- (10) Do be quiet (Production 3)
- (11) Do not get dressed (Production 4)
- (12) Do not be guided by me (Production 4)
- (13) Do not be quiet (Production 4)

Rule Status: Active

Reference:

- (1) Quirk et al, "A Comprehensive Grammar of the English Language"
- (2) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR109, EGR144

Related POS: EPOS101, EPOS120

Replaces: EGR108

Reason: New Release

Replaced by: -



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

Analysis:

Analysis: Directives typically take the form of an imperative sentence, which differs from a declarative sentence in that [pg. 827, 1]:

- It generally has no subject
- It has either a main verb in the base, or (less commonly) an auxiliary in the base form followed by the appropriate form of the main verb

Example: (i) Give me the book.

(ii) Be quiet

The following table summarizes the structural types of imperatives [pg. 830, 1]:

		1 st Person	2 nd Person	3 rd Person
Without subject		--	(1) Open the door	--
With Subject	Without "let"	--	(2) You open the door	(3) Someone open the door
	With "let"	(4) Let me open the door. Let's open the door	--	(5) Let someone open the door

A distinctive feature of imperatives cross linguistically is that they lack an overt subject. The subject is understood to be a second person PRO and must be provided either as a part of the imperative rule or by the imperative morphology [pg. 28, 2].

Example: (1) Push the button.

The imperatives can be both in active and in passive form. The passive form is rare but nevertheless it is found in imperatives.

Example: ACTIVE: (1) Bring the book.

PASSIVE [pg.827, 1]:

(1) Get dressed.

(2) Be guided by me.

The auxiliary "do" in the bare form can also occur before both active and passive imperatives. This adds a feature of emphasis to the sentence.

Example: ACTIVE: (1) Do bring the book

PASSIVE: (2) Do get dressed

To negate the imperatives the auxiliary "do" and the negative particle "not" occurs before the active and passive imperatives. The negative particle introduces a feature _NEG_MARKER to denote that the imperative is negative. The auxiliary "do" is in bare form.

Example: ACTIVE: (1) Do not bring the book

PASSIVE: (2) Do not get dressed

Result: We decided on the above analysis

Future Work:

- (1) The imperatives with "let" and having a subject in them will be done later.

Example: (i) Let me give you the book.

- (2) The imperatives may contain adverbs in them. They will be done later.

Example: (i) Quickly open the door.

(ii) Do your work quietly.



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

- (3) Imperatives without “let”, but having a subject will be catered later because they may be confused with vocatives [pg. 829, 1].
Example: (i) Mary, play on my side
(ii) Mary play on my side
- (4) The imperatives with reflexive pronouns will be done later.
Example: (i) Help yourself / yourselves.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	7 th September 2004	0.1	Created
Kiran Khurshid	22 nd September 2004	1.0	Approved by SA
Kiran Khurshid	10 th December 2004	1.1	Removed the auxiliary "do" from VPimp_act. Moved it to Simp.
Zunaira Malik	9 th August, 2005	1.2	New Release

Rule ID: EGR253

Rule Syntax:

VPimp_pred -> v VPpredlink (VPimp_adjunct)
 VPimp_act -> [VPimp_itrans_trans | VPimp_dtrans]
 VPimp_pass -> aux v
 VPimp_inf -> inf [VPimp_itrans_trans | VPimp_dtrans | VPimp_pass | VPimp_pred]
 VPimp_oblig -> [(PPnmain) | (adv) | (VPimp_inf)]
 VPimp_adjunct -> (PPnmain) (VPimp_inf)

Rule Functional Description:

- 1) VPimp_pred -> v: ^!=!, !_MORPH_FORM =c {BARE, TO_INF}, !_VERB_FORM =c 'be'; VPpredlink: ^!=!;
 (VPimp_adjunct: ^ = !;).
- 2) VPimp_act -> [VPimp_itrans_trans: ^!=!; | VPimp_dtrans: ^!=!;].
- 3) VPimp_pass -> aux: ^ _AUX1 = !, !_MORPH_FORM =c {BARE, TO_INF}, [_AUX_FORM =c 'get' || !_AUX_FORM =c 'be']; v: ^ = !, !VOICE =c PASSIVE;.
- 4) VPimp_inf -> inf: [VPimp_itrans_trans: ^!=!, ^_MORPH_FORM = TO_INF; | VPimp_dtrans: ^!=!, ^_MORPH_FORM = TO_INF; | VPimp_pass: ^!=!, ^_AUX1 _MORPH_FORM = TO_INF; | VPimp_pred: ^!=!, ^_MORPH_FORM =c TO_INF;].
- 5) VPimp_oblig -> [(PPnmain: ! \$ ^OBL;)|(adv: ^OBL =!, !ADV_TYPE =c V_MOD;)] (VPimp_inf: ^XCOMP = !, ^XCOMP SUBJ PRED = 'pro', ^XCOMP SUBJ PRONTYPE = NULL;)].
- 6) VPimp_adjunct -> (PPnmain: ! \$ ^ADJUNCT PREP;) (VPimp_inf: ^ADJUNCT INFINITIVE = !, ^ADJUNCT INFINITIVE SUBJ PRED = NULL;).

Frequency: -

Description: This rule represents all possible VPs in an imperative sentence.



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

c-structure: An adjective phrase or a noun phrase can follow a verb in an imperative verb phrase or it can be a phrase consisting of active or passive verb phrases. Moreover, an imperative sentence can be followed by an infinitive phrase. This infinitive phrase behaves different from the one which occurs in declarative sentences because it cannot contain all combinations of auxiliaries. Hence VPimp_inf is made to cater the infinitive phrases possible in imperatives only.

f-structure: All feature values beginning with “_” are used to check the well-formedness of the sentence.

Example:

- (1) Be quiet. (Production 1)
- (2) Be a good girl (Production 1)
- (3) Attend the meeting. (Production 2)
- (4) Get introduced to him to be famous. (Production 3)
- (5) Be guided by me. (Production 3)
- (6) Run to catch the bus. (Production 4)
- (7) Run to school (Production 5)
- (8) Stop to eat in the morning. (Production 6)

Rule Status: Active

Reference:

- (1) Quirk et al, “A Comprehensive Grammar of the English Language”

Related Rules: -

Related POS: EPOS101, EPOS119, EPOS106

Replaces: - EGR109

Reason: - New Release

Replaced by: -



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

Analysis:

Analysis: Imperative sentences can contain a verb phrase in passive or active form, which may be followed by a to-infinitive clause.

Example: (1) Give me the book to read. (Active VP followed by to-infinitive phrase)

(2) Get introduced to him to be famous. (Passive VP followed by to-infinitive phrase)

Also, an imperative sentence can contain a predicate link which points to the second person overt subject. It can contain a noun phrase or an adjective phrase followed by “be”. The existing rules for the predicate link analysis in declarative sentences could not be used in imperatives, because the combination

Example: (1) Be good

(2) Be quiet

In the passive form, the imperatives can start with the two auxiliaries: “be” and “get”. These auxiliaries are followed by the appropriate form of the verb.

Example: (1) Get dressed

(2) Be guided by me.

Result: We decided on the above analysis

Future Work:

- In the active form, the imperatives usually contain the verb in the bare form only, but in rare cases an auxiliary can precede the verb. In such cases the appropriate form of the verb follows this auxiliary (usually “be” or “have”):
Example: (1) Be listening to this station the same time tomorrow night. [pg. 827, 1]
(2) Start the book and have finished it before you go to bed. [pg. 827, 1]
These constructions will be catered later.
- Verb phrases with adverbs will be done later.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Kiran Khurshid	7 th September 2004	0.1	Created
Kiran Khurshid	22 nd September 2004	1.0	Approved by SA
Zunaira Malik	9 th August, 2005	1.1	New Release

Rule ID: EGR254

Rule Syntax:

VPimp_itrans_trans -> v (NPmain)

VPimp_dtrans -> v NPmain NPmain

Rule Functional Description:

1) VPimp_itrans_trans -> v: ^!=!, !VOICE =c ACTIVE, !_MORPH_FORM =c {BARE, TO_INF};
(NPmain: ^OBJ=!, ^CASE = ACC;).

2) VPimp_dtrans -> v: ^!=!, !VOICE =c ACTIVE, !_MORPH_FORM =c {BARE, TO_INF}; NPmain : ^OBJ2 = !, ^ OBJ2 CASE = DAT; NPmain : ^OBJ = !, ^OBJ CASE = ACC, ~[^OBJ PRED =c 'pro'];.

Frequency: -

Description: These rules represent the intransitive, transitive and ditransitive verb phrases allowed in imperative sentences.

c-structure: An active imperative verb phrase can either be intransitive, transitive or ditransitive.

f-structure: An intransitive verb would not need a noun phrase as an object. A transitive verb phrase would contain a verb followed by a noun phrase. This noun phrase becomes a part of the structure OBJ of the verb phrase, having ACCUSATIVE CASE. In ditransitive verb phrase the first noun phrase becomes OBJ2 having DATIVE CASE and the second object becomes OBJ having ACCUSATIVE CASE and which cannot be a pronoun.

Example:

- 1) Go. (Production 1)
- 2) Cut the apple. (Production 1)
- 3) Give me the book. (Production 2)



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

Rule Status:

Reference:

(1) Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR111

Related POS: EPOS003

Replaces: - EGR110

Reason: - New release

Replaced by: -



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

Analysis:

Analysis: The imperatives can contain an intransitive, transitive or ditransitive verb, all in bare form. The verb phrases in declaratives can occur in any form and can contain auxiliaries but in imperatives only the bare form verb are allowed and no auxiliaries occur. Hence, the existing rules for declaratives could not be reused in imperatives.

Example: (1) Go (intransitive)
(2) Give the book (transitive)
(3) Give me the book (ditransitive)

Result: We decided on the above analysis

Future Work:

- (1) Only some verbs can be used in imperatives. It will be decided later which verbs will be allowed and which wont be.



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



Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	6 th September, 04	0.1	Created
Zunaira Malik	9 th August, 05	0.2	New Release

Rule ID: EGR255

Rule Syntax: Following is the constituent description of the rule.

Sint -> pro VP

Sint -> pro {modal|aux} Sdecl

Sint -> pro {VPpred_main1| VPpred_main2} VPpredlink

Rule Functional Description: Following are the functional specifications of the rule.

1) Sint -> pro: ^SUBJ=!, !PRONTYPE =c INTERROG; VP: ^=!, ^_FLAG = INTERROGATIVE;.

2) Sint -> pro: ^OBJ=!, !PRONTYPE =c INTERROG; [modal: ^HelpVP=!, ^TNS_ASP= !TNS_ASP;|aux: ^HelpVP=!, ~[!_MORPH_FORM =c {BARE,TO_INF}], ^TNS_ASP= !TNS_ASP, ^VOICE = !VOICE;] Sdecl: ^=!, !_FLAG =c INTERROGATIVE;.

3) Sint -> pro: ^SUBJ=!, !PRONTYPE =c INTERROG; [VPpred_main1: ^=!;|VPpred_main2: ^ = !;] VPpredlink: ^=!;.

Frequency: -

Description: This rule shows the interrogative production of English Grammar.

c-structure: The above productions show some of the interrogative constructions. The above productions uses wh-words which act as pronouns in the interrogative sentence. These kind of interrogatives can take an intransitive verb(as in production1), transitive verb(as in production2) and copular verb(as in production3).

f-structure: The first production makes pronoun or noun phrase as a SUBJ if it is of type interrogative.

The second production is for transitive or ditransitive verbs. It uses Sdecl rule. The constraints on which auxiliary should be followed in the series of auxiliaries, is done through the m-structure (represented by variable starting with underscore).

The third production caters copular constructions.

Examples:

- 1) What works? Whose books will be published? Who is being examined? (for production 1)
- 2) What have you been teaching? Whose book are you reading? What have you been teaching me? What are you doing? (for production 2)
- 3) What is good? What is your name? Who are you (for production 3)



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

Rule Status: Active

Reference:

- [1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"
- [2] Pam Peters, "The Cambridge Guide to English Usage"

Related Rules: EGR139, EGR127, EGR103, EGR104

Related POS: EPOS103, EPOS114, EPOS101

Replaces: EGR115

Reason: - New release

Replaced by: -



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

Analysis: Following is the in-depth analysis of the rule.

Analysis 1: Interrogatives often have substantially different c-structures from their declarative counterparts. English places interrogative words (*wh*-words) in a certain position (clause initial). Yes-no questions may be formed by subject-auxiliary inversion as in example below:

Example:

You have driven this tractor.

Have you driven this tractor?

For questions formed with so-called *wh*-words (who, what, etc.), these differences include the appearance of interrogative phrases. These phrases often appear in restricted c-structure positions. Due to these substantial differences in c-structure, interrogatives are treated with a separate set of c-structure rules, including one set for yes-no questions and one for *wh*-questions. This allows a simple way of introducing STMT-TYPE INTERROGATIVE, which all interrogatives have, and the special punctuation which usually accompanies root interrogatives. However, a number of problems arise with regard to the appropriate distribution of interrogative phrases, especially in multiple questions, and with ensuring the correct form of the auxiliaries in subject-auxiliary inversion constructions. [1, pg. 24]

Analysis 2: interrogative words include pronouns:

who, what, which, whom, whose

and adverbs:

when, where, why, how

Both can be used in either direct or indirect questions.

Who's there?

He asked who's there?

What do you want?

They inquired what I wanted?

In English *interrogative constructions*, the normal subject-verb order is inverted, and the subject *they/he/you* follows the first (auxiliary) part of the verb. Modern English always brings in *do* to for the interrogative when the verb is not itself an auxiliary as in *I like red wine* formed interrogative as *Do you like red wine?*

Modern grammars apply the term interrogative to the particular 'sentence function' or 'clause type' that expresses a question, rather than the distinctive verb form. This recognizes the fact that an interrogative construction can express other speech functions, such as imperative. In the US and elsewhere, the sentence *Why don't you open the door?* is a polite way of instructing someone to do something. [2, pg. 288]

Result: We decided on the above analysis.

Future Work:

- 1) Auxiliaries agreement in copular verb will be catered as in **Are Ayesha being a good girl?*
- 2) Passive constructions in interrogatives will be examined.
- 3) As in analysis 2, interrogative construction having a passive mood will be modeled.
- 4) Tag questions with declaratives, imperatives and exclamatives will be covered.
Example: Open the door, won't you? The boat hasn't left, has it?
- 5) Negative *wh*-questions will be covered.
Example: Who hasn't had any money? Why didn't you tell me?



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





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

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	6 th September, 04	0.1	Created
Zunaira Malik	9 th August, 05	0.2	New Release

Rule ID: EGR256

Rule Syntax: Following is the constituent description of the rule.

Sint -> (adv) {modal|aux} Sdecl

Sint -> (adv) VPpred_main3 NP VPpredlink

Sint -> (adv) {modal|aux} NP { VPpred_main2 | VPpred_main3} VPpredlink

Rule Functional Description: Following are the functional specifications of the rule.

1) Sint -> (adv: ^ADJUNCT ADV=!, !SEM_TYPE =c INTERROG;) [modal: ^HelpVP=!, ^TNS_ASP= !TNS_ASP;|aux: ^HelpVP=!, ~[! _MORPH_FORM =c {BARE, TO_INF}], ^TNS_ASP= !TNS_ASP, ^VOICE = !VOICE;] Sdecl: ^=!, ^_FLAG = INTERROGATIVE;.

2) Sint -> (adv: ^ADJUNCT ADV=!, !SEM_TYPE =c INTERROG;) VPpred_main3: ^=!, NP: ^SUBJ = !; VPpredlink: ^=!, ^_FLAG = INTERROGATIVE;.

3) Sint -> (adv: ^ADJUNCT ADV=!, !SEM_TYPE =c INTERROG;) [modal: ^HelpVP=!, ^TNS_ASP= !TNS_ASP;|aux: ^HelpVP=!, ~[! _MORPH_FORM =c {BARE, TO_INF}], ^TNS_ASP= !TNS_ASP, ^VOICE = !VOICE;] NP: ^SUBJ = !; [VPpred_main2: ^=!, |VPpred_main3: ^=!,] VPpredlink: ^=!, ^_FLAG = INTERROGATIVE;.

Frequency: -

Description: These rules show the interrogative production of English Grammar.

c-structure: The above productions show some of the interrogative constructions having interrogatives acting as adverbials. For details check the analysis below.

Production 1 models an interrogative having split VP. An auxiliary or modal occurs initially followed by an NP and then by a sequence of auxiliaries (optional) and then the main verb. Verb can be transitive or intransitive.

Part of this rule uses the declarative sentence production rule.

Production 2 and 3 are copular constructions with or without Helping Verb Phrase.

f-structure: The features starting with underscore are for checking well-formedness. The adverbial interrogative is made ADJUNCT in f-structure.

When the first helping verb is encountered in the production 1, it is added to f-structure as a junk structure to be used for agreement checking. Later this helping verb's structure is used to check the right auxiliaries are followed after NP.

Production 2 and 3 also uses VPpred_main 2 and 3 (the production made for interrogatives having helping verb and mainverb). VPpredlink is the argument part that is a property of the subject.

Examples:



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

- 1) Why have I been writing you? Why may you give me a pen? (for production 1)
- 2) Why are you good? Are you good? Is your name Imran? Are you on the table? (for production 2)
- 3) Why are you being good? Are you being good? (for production 3)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

[2] Pam Peters, "The Cambridge Guide to English Usage"

Related Rules: EGR139, EGR103, EGR104, EGR117

Related POS: EPOS102, EPOS114, EPOS101

Replaces: EGR116

Reason: - New Release

Replaced by: -



Analysis: Following is the in-depth analysis of the rule.

Analysis 1: Interrogatives often have substantially different c-structures from their declarative counterparts. English places interrogative words (*wh*-words) in a certain position (clause initial). Yes-no questions may be formed by subject-auxiliary inversion as in example below:

Example:

You have driven this tractor.

Have you driven this tractor?

For questions formed with so-called *wh*-words (who, what, etc.), these differences include the appearance of interrogative phrases. These phrases often appear in restricted c-structure positions. Due to these substantial differences in c-structure, interrogatives are treated with a separate set of c-structure rules, including one set for yes-no questions and one for *wh*-questions. This allows a simple way of introducing STMT-TYPE INTERROGATIVE, which all interrogatives have, and the special punctuation which usually accompanies root interrogatives. However, a number of problems arise with regard to the appropriate distribution of interrogative phrases, especially in multiple questions, and with ensuring the correct form of the auxiliaries in subject-auxiliary inversion constructions. [1, pg. 24]

Analysis 2: interrogative words include pronouns:

who, what, which, whom, whose

and adverbs:

when, where, why, how

Both can be used in either direct or indirect questions.

Who's there?

He asked who's there?

What do you want?

They inquired what I wanted?

In English *interrogative constructions*, the normal subject-verb order is inverted, and the subject *they/he/you* follows the first (auxiliary) part of the verb. Modern English always brings in *do* to for the interrogative when the verb is not itself an auxiliary as in *I like red wine* formed interrogative as *Do you like red wine?*

Modern grammars apply the term interrogative to the particular 'sentence function' or 'clause type' that expresses a question, rather than the distinctive verb form. This recognizes the fact that and interrogative construction can express other speech functions, such as imperative. In the US and elsewhere, the sentence *Why don't you open the door?* is a polite way of instructing someone to do something. [2, pg. 288]

Result: We decided on the above analysis.

Future Work:

- 1) Auxiliaries agreement in copular verb will be catered as in **Are Ayesha being a good girl?*
- 2) Passive constructions in interrogatives will be examined.
Example: What is being examined by the doctor?
- 3) As in analysis 2, interrogative construction having a passive mood will be modeled.
- 4) Tag questions with declaratives, imperatives and exclamatives will be covered.
Example: Open the door, won't you? The boat hasn't left, has it?
- 5) Negative *wh*-questions will be covered.
Example: Who hasn't had any money? Why didn't you tell me?
- 6) Interrogative Semantics as in Analysis 2 will be covered.



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





Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Shanza Nayyer	16 th September, 04	0.1	Created
Kiran Khurshid	22 nd December, 04	0.2	Added coordinate conjunctions and subordinate correlatives
Zunaira Malik	9 th August, 05	0.3	addition of Sub_ConjP_opt production and New release

Rule ID: EGR257

Rule Syntax: Following is the constituent description of the rule.

S -> S_coord_conjP
S -> Smain
S -> Smain Sub_ConjP
S -> Smain Sub_ConjP_opt
S -> correlative Smain (comma) Sub_ConjP_opt
S -> correlative Smain Sub_ConjP
S -> Sub_ConjP comma Smain
Sub_ConjP -> sub_conj Smain
Sub_ConjP_opt -> (sub_conj) Smain

Rule Functional Description: Following are the functional specifications of the rule.

- 1) S -> S_coord_conjP: ^ = !;.
- 2) S -> Smain: ^ = !, ^CLAUSE_TYPE = !CLAUSE_TYPE;.
- 3) S -> Smain: ^ = !; Sub_ConjP: ^ ADJUNCT SUBORD = !, !_POSITION = c MIDDLE;.
- 4) S -> Smain: ^ = !; Sub_ConjP_opt: ^ COMP = !;.
- 5) S -> correlative: ^ = !; Smain: ^ = !; (comma;.) Sub_ConjP_opt: ^ ADJUNCT SUBORD = !, ^_ALLOWED_CONJ = !CONJ_FORM;.
- 6) S -> correlative: ^ = !; Smain: ^ = !; Sub_ConjP: ^ COMP = !, !_POSITION = c MIDDLE, ^_ALLOWED_CONJ = !CONJ_FORM;.
- 7) S -> Sub_ConjP: ^ ADJUNCT SUBORD = !, !_POSITION = c START; comma; Smain: ^ = !;.
- 8) Sub_ConjP -> sub_conj: ^ = !; Smain: ^ = !;.
- 9) Sub_ConjP_opt -> (sub_conj: ^ = !, !_POSITION = c MIDDLE;) Smain: ^ = !;.

Frequency: -

Description: This rule shows the subordinate coordination at sentence level.

c-structure: These productions show subordinate and coordinate conjunctions at sentence level. The first production uses the coordinate conjunction phrase rule. The second production allows a sentence without a conjunction to parse. The third and fourth phrase caters subordinate conjunction at the middle of two sentences. The fifth and sixth production allows correlatives with the subordinate conjunctions. The seventh production caters subordinate conjunction at the start of the sentence. In such case, the two sentences are separated by comma. The eighth production is the subordinate clause which is being used by the above five productions.



f-structure: In (Sub_ConjP) subordinate conjunction phrase, the sub-ordinate clause becomes the COMP in *f-structure* when it is sub-categorized by the verb (as in production 4) and otherwise becomes ADJUNCT (as in production 3). Right now the sentence initial subordinate conjunction is made an ADJUNCT (as in production 7).

Examples:

- 1) She came home and he gave her the book. (Production 1)
- 2) I would be going to the party. (Production 2)
- 3) Girl came home after/as I went to school (Production 3)
- 4) She took my money that/because I left on the table (Production 3)
- 5) She slept while I went to school (Production 3)
- 6) She said that I should go to school (Production 4)
- 7) As the strength of the defenders failed, so the courage of the attackers grew. (Production 5)
- 8) So she said that I should go to school (Production 6)
- 9) While I went to school, she slept (Production 7)

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

[2] www.grammar.englishclub.com

[3] Radolph Quirk, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik, "A Comprehensive Grammar of the English Language"

Related Rules: EGR0140

Related POS: EPOS117, EPOS118, EPOS122

Replaces: EGR117

Reason: addition of Sub_ConjP_opt production and New release

Replaced by: -

Reason: -



Analysis: Following is the in-depth analysis of the rule.

Analysis1: Conjunctions are words that "join". Conjunctions join two parts of a sentence. They are discussed in detail below:

1. Form

Conjunctions have three basic forms:

Single Word

Example: and, but, because, although

Compound (often ending with *as* or *that*)

Example: provided that, as long as, in order that

Correlative (which surround an adverb or adjective)

Example: so...that

2. Function

Conjunctions are divided into two basic types.

Coordinating Conjunctions are used to join two parts of a sentence that are grammatically equal. The two parts may be single words or clauses,

Example: i) Jack and Jill went up the hill.

ii) The water was warm but I didn't go swimming.

Subordinating Conjunctions are used to join a subordinate dependent clause to a main clause, for example:

- *I went swimming, **although** it was cold.*

3. Position

Coordinating Conjunctions always come between the words or clauses that they join.

Subordinating Conjunctions usually come at the beginning of the subordinate clause. [2]

Analysis2: The coordination rule must cater for more than two conjuncts as in *they close the door, locked it, and walked to the car*. The rules must be formulated differently, according to whether the non-final conjuncts are separated from one another only by commas, or whether a conjunction is also required. Moreover, the rule must allow for two-part conjunctions, such as *either...or, neither...nor*.

NP coordination often involves number, person, and gender mismatches between the individual conjuncts and the entire coordinated NP.

Two requirements must be met by an analysis of NP coordination. First, it must be possible to assert constraints about and assign values to attributes in both the individual conjuncts and in the f-structure on the coordination as a whole. Second, there must be an algorithm to construct the number, person, and gender values of the coordinated NP from the values of the individual conjuncts.

The type of conjunction involved in the coordinated phrases plays a role in determining agreement. The general rule in English for determining the number, gender, and person of a coordinated NP is as follows:

- 1) If any conjunct is plural, the entire NP is plural;
- 2) If any conjunct is masculine, the entire NP is masculine;
- 3) If any conjunct is first person, the entire NP is first person;
- 4) If there is no first person conjunct and any conjunct is second person, the entire NP is second person; otherwise the NP is third person. [1, p.145]



Analysis 2:

Another kind of embedding occurs when one clause is made a constituent of another clause: the case normally described as SUBORDINATION. Consider the following

Example: i) The weather has been remarkably warm [since we returned from Italy last week].

Clauses which are embedded in other clauses are subordinate clauses (e.g. since we returned from Italy last week). The relation between the two clauses is one of 'part to whole'.

Subordination of clauses is not confined to clauses which are immediate constituents of other clauses. There are also clauses which are constituents of phrases, and which therefore are only indirectly embedded within a larger clause (as the relative clause used as the postnominal modifier)

Embedding gives rise to the theoretical possibility of grammatical units having indefinite length.

Example: i) This is the house [that Jack built].

ii) This is the malt [that lay in the house [that Jack built]].

iii) This is the rat [that ate the malt [that lay in the house [that Jack built]]] [3, pg. 44]

Result: The above analyses are used for the rule.

Future Work:

1) Right now subordinate conjunction is catered partially. Every subordinate conjunction and its semantic behavior will be covered in the later stage.

Reference No:

Revision History:

Name	Change Date	Version	Description of Changes
Zunaira Malik	20 th May, 05	0.1	Created

Rule ID: EGR200

Rule Syntax: Following is the constituent description of the rule
VP adjunct -> (VP inf main) (VP participle) (PP n main)

Rule Functional Description: Following are the functional specifications of the rules.
VP adjunct -> (VP inf main: ^ADJUNCT INFINITIVE = !;) (VP participle: ^ADJUNCT PARTICIPLE = !;) (PP n main: ! \$ ^ADJUNCT PREP = !;).

Frequency: -

Description: This production gives the detail of the VP adjunct production

c-structure:

f-structure:

Examples:

Rule Status: Active

Reference:

[1] Miriam Butt, Tracy Holloway King, "A Grammar Writer's Cookbook"

Related Rules: EGR208, EGR209, EGR210

Related POS: -

Replaces:

Reason:

Replaced by: -

Reason: -

Analysis: Following is the in-depth analysis of the rule:

Analysis: Predicative construction involves a linking or *copular* verb which has a subject and another argument, as in the example below:

Example: i) The beacon is on the roof.

ii) The tractor is red.

The post-verbal argument can be of a number of categories, e.g., NP, PP, AP etc. Due to the semantic relationship between the subject and the phrase after linking verb, these verbs are given special sub-categorization frames. Traditionally, this has been done by having the post-verbal phrase 'be' an XCOMP whose subject is controlled by the linking verb's subject. However, a new analysis, termed the PREDLINK analysis, is used. Under both approaches, linking verbs may have their own c-structure category and their own VP rule which allows the post-verbal NP, AP, and PP to be assigned the appropriate grammatical function. [1, p. 69]

Result: The above analysis was finalized.

Future Work:

1) The sentence like *the chicken is cooked* will create ambiguity. It will be parsed having *is* as the main verb and *cooked* as an ADJP and will also be parsed through the production having *cooked* as the main verb. This kind of ambiguity will be catered in the later stage using some statistical method.