



## Reference No:

## Revision History:

Name	Change Date	Version	Description of Changes
Tafseer Ahmed	6 <sup>th</sup> Nov, 04	0.1	Created
Sara Hussain	7 <sup>th</sup> Jan, 2005	4.1.0.0	Added SEM_TYPE feature with TEMPORAL and PLACE options

POS ID: UPOS104

Part of Speech Name: Noun

Part of Speech String: n

## Lexical Entry Template:

<word>: n, ^ PRED = '<logical form of word>', ^ NUM = {SG, PL}, ^ GEND = {M, F}, ^ FORM = {NOM, OBL}, ^ CTYPE = {COUNT, MASS}, ^ NTYPE = {PROPER, COMMON}, ^ SEM\_TYPE = {TEMPORAL, PLACE}, ^ ANIM = {+, -}, ^ RESPECT = {NORESPECT, FAMILIAR, USUAL, EXTRA}.

## Sample Lexical Entry:

لڑکی : n, ^ Pred = 'لڑکی', ^ GEND = F, ^ NUM = SG, ^ FORM = {NOM, OBL}, ^ CTYPE = COUNT, ^ NTYPE = COMMON, ^ ANIM = +, ^ RESPECT = {NORESPECT, FAMILIAR, USUAL, EXTRA}.

لڑکیاں : n, ^ Pred = 'لڑکی', ^ GEND = F, ^ NUM = PL, ^ FORM = NOM, ^ GEND = F, ^ CTYPE = COUNT, ^ NTYPE = COMMON, ^ ANIM = +, ^ RESPECT = {NORESPECT, FAMILIAR, USUAL, EXTRA}.

لڑکیوں : n, ^ Pred = 'لڑکی', ^ GEND = F, ^ NUM = PL, ^ FORM = OBL, ^ CTYPE = COUNT, ^ NTYPE = COMMON, ^ ANIM = +, ^ RESPECT = {NORESPECT, FAMILIAR, USUAL, EXTRA}.

لڑکی : n, ^ Pred = 'لڑکی', ^ GEND = F, ^ NUM = SG, ^ FORM = {NOM, OBL}, ^ CTYPE = COUNT, ^ NTYPE = COMMON, ^ ANIM = +, ^ RESPECT = {NORESPECT, FAMILIAR, USUAL, EXTRA}.

آج : n, ^ Pred = 'آج', ^ GEND = M, ^ NUM = SG, ^ FORM = {NOM, OBL}, ^ CTYPE = COUNT, ^ NTYPE = COMMON, ^ SEM\_TYPE = TEMPORAL, ^ ANIM = -, ^ RESPECT = NORESPECT.

گھر : n, ^ Pred = 'گھر', ^ GEND = M, ^ NUM = SG, ^ FORM = {NOM, OBL}, ^ CTYPE = COUNT, ^ NTYPE = COMMON, ^ ANIM = NEG, ^ RESPECT = NORESPECT.

**Description:** A noun is the name of a person, place, thing, or idea. Whatever exists, we assume, can be named, and that name is a noun.

The description of the features is following:

NUM (Number) can be: SG (singular): Example: لڑکی  
PL (plural): Example: لڑکیاں  
It cannot be null or both.

GEND (Gender) can be: M (masculine): Example: لڑکا  
F (feminine): Example: لڑکی  
It cannot be null or both. For mass nouns, it is always singular.

FORM can be: NOM (Nominative): Example: لڑکیاں



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

OBL (Oblique): It cannot be null or both.	Example: لڑکیوں
CTYPE (Count Type) can be: COUNT MASS It cannot be null or both.	Example: کتاب Example: پانی
NTYPE (Noun Type) can be: COMMON PROPER It cannot be null or both.	Example: لڑکا Example: احمد
SEM_TYPE (Semantic Type) can be: TEMPORAL PLACE It cannot be both but can be NULL.	Example: آج Example: گھر
ANIM (Animated) can be: + (Positive): - (Negative): It cannot be null or both.	Example: لڑکا Example: کتاب
RESPECT can be: NORESPECT, FAMILIAR, USUAL, EXTRA It cannot be null. It is used for agreement especially in Predicate Link agreement.	

The detail of the feature values is discussed in analysis section.

**Examples:**

شاعری، کمپیوٹر، ڈائری، بینسل، بستہ، بوتل، کتاب، آدمی، گھوڑا، حامد، احمد  
(Temporal) آج، بفتہ، صبح  
(Place) گھر، کلکتہ، لاڑکانہ

**POS Status:** Active

**Reference:**

- [1] <http://webster.commnet.edu/grammar/nouns.htm>
- [2] "Urdu Sarf au Nahav", Maulvi Abdul Haq

**Related POS:**

**Replaces:** - UPOS04

**Reason:** -

**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:**

Number of a noun can be singular (SG) or plural (PL) or both. It is a mandatory feature. If any noun has both singular and plural values they will be written in the form of set, e.g. {SG, PL}. Same is for Gender feature.

In Urdu, noun can be of two forms. One is Nominative (NOM) i.e. *lerka* لڑکا and other is Oblique (OBL) i.e. *lerkey* لڑکے as singular. Nominative form is used when noun phrase is in Nominative case. When noun phrase is in some other case Oblique form is used along with case marker. For further details see POS Document of case markers and Grammar rule document for Case Phrase (KP).

Nouns can be of two types with respect to their physical properties: countable and mass nouns. This property helps in choosing right specifier for them. It is reflected as feature CTYPE. 'کتاب' is an example of Count Noun while 'پانی' is an example of Mass Noun.

With respect to their use, two major categories of nouns are common noun and proper noun. Proper Noun is a name of particular person, place or thing whereas common noun is a noun which can be used for generic type like man, horse, tree, book. These types are reflected in feature NTYPE. These are useful because on their basis some classes of specifiers and modifiers can be restricted to some typical type.

A noun can be animated or unanimated. It can help in deciding correct case marker for noun in some cases. For example, when a noun occurs at object position, accusative case is used for animated nouns and nominative case is used for unanimated nouns. e.g. اس نے کرسی اٹھائی, اس نے بچے کو اٹھایا.

In Urdu there is a rule to use plural agreements for singular nouns if one wants to give respect. Therefore a feature of respect is added which will have RESPECT value for words like 'چچا', 'والدہ' etc. Such words will have USUAL respect. For words like محمد ﷺ respect value will be EXTRA. This rule is not applicable on unanimated nouns; e.g. قرآن plural will not be used.

Semantic features shows that a noun can be temporal (آج، بوقتہ، صبح) or spatial (لاڑکانہ، کلکتہ، گھر). These features are significant when temporal or spatial nouns occur as adjuncts in a sentence (either alone or in a postposition phrase). The order of adjuncts in a sentence is such that temporal adjuncts are followed by place (spatial) adjuncts. To encode such ordering scheme semantic features in nouns have been introduced.

**Result:** We decided on Analysis above.

**Future Work:** 1) More noun types as abstract noun may be added if found significant for syntax rules.

2) The different case forms (e.g. nominative, oblique) in temporal / place nouns such as بوقتہ، کارخانہ، لاڑکانہ، کلکتہ etc. needs to be analyzed.