Development of Sindhi Lexical Functional Grammar

Mutee U Rahman and Hameedullah Kazi
Department of Computer Science, Isra University, Hyderabad, Sindh 71000, Pakistan
muteeurahman@gmail.com , hkazi@isra.edu.pk

Abstract

This paper presents an on-going work on Sindhi morphology and grammar development. An LFG (Lexical Functional Grammar) model for Sindhi is developed where morphological constructions are modeled in Xerox Lexicon Compiler (LEXC), and syntactic constructions are modeled in LFG by using Xerox Linguistic Environment (XLE). Development of various grammatical constructions of Sindhi is discussed. Morphological constructions considered for development include: Nouns, Pronouns, Adjectives, Verbs, Adverbs, Postpositions and pronominal suffixation of verbs. Syntactic constructions include noun phrase, verbal complex, verbal subcategorization, adjuncts, coordination, and subordination. While developing morphology and syntax of Sindhi, Tense, Aspect, Mood and Agreement are also considered wherever applicable.

1. Introduction

The paper presents Sindhi computational grammar development project in which finite state morphology and LFG (Lexical Functional Grammar) frameworks are used to implement Sindhi morphology and syntax respectively. Finite state morphology is implemented in XFST (Xerox Finite State Technology) tools [1] and Syntax is implemented in XLE (Xerox Linguistic Environment [2]. Sindhi is a resource poor language in Computational Linguistics and Natural Language Processing domains. Neither Sindhi Morphology nor the Syntax is studied by researchers with computational linguistics perspective. Sindhi has rich inflectional and derivational morphology. Nouns adjectives and pronouns have number gender and case inflections [3]. Verb morphology includes number, gender, tense, aspect and mood inflections. Sindhi syntax features include free constituent ordering, agreement, complex noun phrase constructions, coordination, subordination, syntactic case formations, and pro-drop. LFG rules defined for Sindhi syntax quite reasonably handle these syntactic constructions. Following sections give an overview of finite state morphology and lexical functional grammar frameworks.

1.1. Finite State Morphology

Finite state transducers (FSTs) play an important role in language processing applications [4] and computational studies of morphologically complex languages. Efficient morphological parsers can be implemented by combining these FSTs and computational lexicon (repository of words). FSTs convert/translate lexical level constructs to surface level words by applying morphotactics (morpheme ordering rules. Their reversible nature makes reverse conversion/translation possible. This two level (lexical and surface) morphology plays important role in implementation of morphological analyzers for natural languages [17]. Figure-1 shows the process of two level (lexicon and surface) morphology modeling using FSTs. A sample orthography FST rule is given in Table-1. This rule is used by FSTs to convert intermediate level word into surface word. Finite state morphological models based on these FSTs are well known models and successfully been used for morphological modeling of many languages. They handle concatenative and non-concatenative morphology very well [5].

1.2. Lexical Functional Grammar

Lexical Functional Grammar (LFG) is a natural language syntax representation formalism based on generative grammars [6] [18]. LFG defines the structure of language and relationship among different aspects of linguistic structure. Various relations are de-fined at lexicon level as LFG has a rich lexical structure. LFG represents linguistic structure at different levels which include lexicon, constituency structure (c-structure) and functional structure (f-structure) levels.

Table 1. A sample orthograpghy rule

Singular	Intermediate	Plural	Rule
Mango	Mangos	Mangoes	ε→e /^____s#

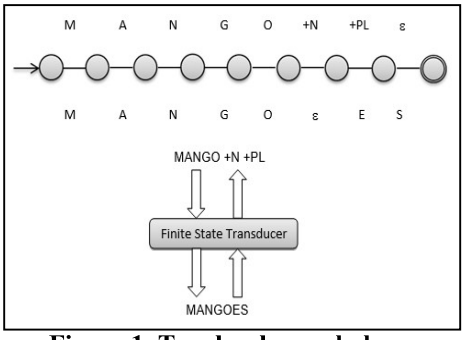


Figure 1. Two level morphology

The lexicon contains list of words or parts of words (smaller meaning bearing units) along-with information about these words including their distribution or syntax and morphology. Thus a lexical entry in LFG may include part of speech, number, gender, case, and argument structure in case of verbs and some postpositions and adjectives. Syntactic structure information in LFG is represented at two different levels. C-structure representation handles word or phrase grouping and their precedence in a phrase structure tree along-with some grouping and order constraints. F-structure represents more abstract relations between different functional constructs like subject, object, secondary object, oblique, complement, and open complement.

Subsequent sections discuss related work, implementation of finite state morphology and lexical functional grammar for Sindhi with nominal and verbal elements, pronominal suffixes followed by conclusion and future work.

2. Related Work

Apart from (Rahman and Bhatti) [7] one cannot find any work in finite state morphology and LFG frameworks for Sindhi morphology and grammar developments. In this work Sindhi noun morphology is discussed and few basic FSTs are presented. However, Sindhi syntax representation efforts in Context Free Grammars and Linear Specification Language can be found in (Rahman and Shah) [8] and (Rahman, Shah and Memon) [9]. First study tries to represent selected Sindhi sentence structures in CFG rules which have over generation problems. Second study tries to cope with over generation by using LSL (Linear Specification Language) but again lacks the agreement problem solution and feature representations. The only comparatively comprehensive research study available but not yet published is "Implementing GF Resource Grammar for Sindhi" [10]. The study tries to investigate the Sindhi morphology and syntax from computational perspective in grammatical framework [11]. However, the study does not cover the most of the parts of Sindhi morphology and syntax. Neither the morphological analyzer nor the syntax analyzer is proposed or designed; only the resource grammar library is made available as a shared resource.

Among south Asian languages Urdu is extensively studied with LFG perspective [12]. Urdu became part of parallel grammar project (Butt Helge and King) [13] and was analyzed with large scale grammar development perspective. It was found that basic analysis decisions made for European languages are applicable to typologically different language Urdu. In Pargram project parallel computational grammar of different languages is being developed within LFG framework. Various research articles discussing different syntactic issues in Urdu LFG including complex predicates, clitics, argument structure, argument srambling in noun phrases and verb phrases can be found on official website of Urdu Pargram [14]. Jafar Rizvi in his PhD thesis [15] also presented Urdu syntax analysis in LFG.

3. Implementation of Finite State Morphology and Lexical Functional Grammar for Sindhi

Overall implementation model is shown in Figure-2. Survey of Sindhi language and linguistics provides foundations of the work. Based on these foundations Sindhi grammar is analyzed and studied with LFG perspective and Sindhi morphological constructions are studied with finite state morphology perspective.

Later Xerox Finite State Tools Lexicon Compiler and Xerox Linguistic Environment are used to develop Sindhi morphology and Syntax respectively. Different components are interfaced with each other in XLE to parse and analyze Sindhi sentences. Apart from finite state morphology full form lexicon for postpositions is also developed in LFG. As a result, parse tree and functional structure analysis are generated.

Following sections discuss morphology and syntax implementation details.

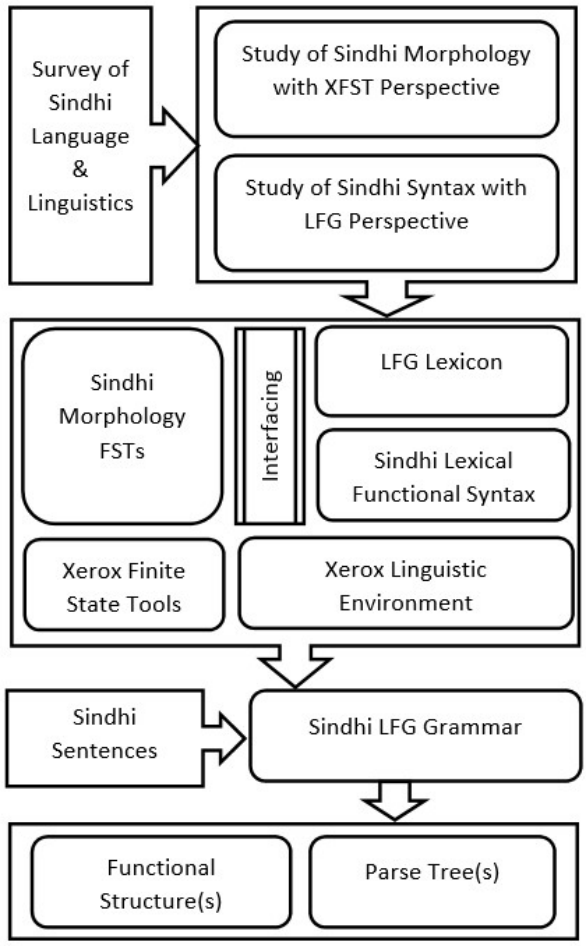


Figure 2. Sindhi grammar development model

3.1. Implementing Morphology

Inflectional morphology of various word classes of Sindhi is implemented by incorporating the inflection rules in finite state models using Xerox LEXC (Lexicon Compiler) [16]. Different morphological paradigms of nouns, pronouns, adjectives, adverbs and verbs are represented in finite state transducers in LEXC syntax. These scripts are compiled to generate finite state machines which represent Sindhi lexicon.

Sindhi nouns are inflected by number gender and case. Different paradigms are modeled in FST rules and resulting transducers act as function machines in which either upper side represents the input and lower side represents the output or vice versa. The reversible property of these FSTs makes them very useful. When lower side becomes input these FSTs function as morphological analyzers and when upper side is input these will function as surface form generators.

While defining finite state lexicon for Sindhi nouns in LEXC, a Root lexicon named Nouns is defined which is further extended to various sub-lexicons. These sub-lexicons actually model various inflectional paradigms of nouns. LEXC script fragment defining Sindhi noun morphology is shown in figure 3.

```
!SINDHI NOUN MORPHOLOGY
Multichar_Symbols
+Noun +Adjective +Adverb +Verb
+Common +Proper +Abstract !Noun Types
+Animate +Inanimate !Noun Concept
+Accusative +Dative +Ergative
+Genitive +Instrumental + Locative
+Nominative +Oblique +Vocative !Noun
Cases
+Count +Mass +Gerund +Measure +City
+Country +FirstName +LastName
+FullName +Name
+Fem +Masc !Gender
+Sg +Pl !Number
+1st +2nd +3rd !Person

LEXICON Root
Nouns;

LEXICON Nouns
       !Boy (Animate Common Noun)
CHOkirO+Noun+Common+Count+Animate:CHO
kir N_Cat1;
       ...
LEXICON N_Cat1
       +Sg+Masc+Nominative:O      #;
       +Sg+Masc+Oblique:E         #;
       +Sg+Masc+Vocative:A        #;
       +Sg+Fem+Nominative:Ia      #;
       +Sg+Fem+Vocative:I         #;
       +Sg+Fem+Oblique:I          #;
       +Pl+Masc+Nominative:A      #;
       +Pl+Masc+Oblique:ani       #;
       +Pl+Masc+Vocative:aO       #;
       +Pl+Fem+Nominative:yUN     #;
       +Pl+Fem+Oblique:yani       #;
       +Pl+Fem+Vocative:yUN       #;
```

Figure 3. LEXC fragment of Sindhi noun morphology

It can be seen that script starts with multi character symbol declarations which are used to define morphological tags. Stem forms of noun are placed in root lexicon (Nouns in this case) followed by sequence of tags representing different features of noun. Stem along-with these features will produce intermediate word form shown after colon (:) following the tag

sequence. This intermediate form is further inflected based on various feature sequences defined in sub-lexicon N_Cat1. For example, consider the stem form and tag sequence given below:

```
CHOkir+Noun+Common+Count+Animate
```

This will produce intermediate animate common count noun form "CHOkir", this transducer is followed by another transducer in series (via N_Cat1 sub-lexicon link) which takes further input tags as shown below:

```
+Sg+Masc+Nominative
```

This tag sequence produces the singular masculine nominative morpheme "O". The overall concatenated tag sequence preceded by stem (upper side) and concatenated output (lower side) are given below:

```
Upper:        CHOkir+Noun+Common+Count+
              Animate+Sg+Masc+Nominative
Intermediate: CHOkir          O
Lower:        CHOkirO
```

While going from upper to lower side surface form "CHOkirO" of stem "CHOkir" with features specified in tag sequence is generated; going from lower to upper will give following morphological analysis of noun "CHOkirO".

```
CHOkir    {"+Noun"    "+Common"    "+Count"
"+Animate" "+Sg" "+Masc" "+Nominative"}
```

Above morphological analysis says that "CHOkirO" is a morphological form of stem "CHOkir" which is a common animate count noun in singular masculine form with nominative case. In the same way oblique morphological form (used as base for various syntactic cases of nouns) "CHOkirE" is generated by producing and concatenating the oblique morpheme "E" by input tag sequence given below and output sequence "CHOkir" and "E".

```
CHOkir+Noun+Common+Count+Animate+Sg+M
asc+Oblique
```

Total twelve (12) different inflections of stem "CHOkir" are taken care of. A total of 21 different common noun categories are identified according to their inflectional properties. For every category a different sub-lexicon is defined. Usually proper nouns are not inflected therefore their entries only contain the feature tags. For example, the proper noun Pakistan has following entry in the lexicon.

```
Pakistan+Noun+Proper+Inanimate+Countr
y+Masc+Sg+3rd:pAkistAn#;
```

It says that Pakistan is an inanimate masculine singular proper noun which is a country and has surface form "pAkistAn". Most of the proper nouns have this type of entry. However, in Sindhi there are exceptional cases of proper noun inflections. For example, a person name "dOdO" can have number, and case inflections "dOdA" (plural or singular vocative) and "dOdE" (oblique form). A sub-lexicon is defined to handle these inflections.

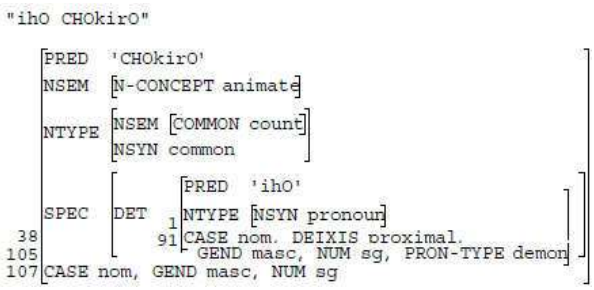
Verb in Sindhi is a morphologically complex word class. Verbs are marked by number, gender, case, tense, aspect and mood. Various categories of auxiliary verbs are also inflected by number, gender, and case; auxiliaries may also be used as tense and aspect markers with inflections. Copula verbs also undergo morphological changes. Due to many different categories of verbs reasonably good number of tags is used while implementing verb morphology. Verb lexicon covers auxiliary verb, copula verb and main verb morphology. Morphological analyses show that a verb in Sindhi can have up to 75 different morphological forms. Implementation strategy of verb morphology is identical to noun morphology discussed above. Pronoun, Adjective, and adverb morphology is also modeled on same lines.

3.2. Implementing Syntax

Different syntactic constructions of Sindhi are implemented in XLE by defining Sindhi LFG rules. Morphology defined in LEXC scripts is compiled to finite state transducers (discussed in previous section) and integrated to LFG grammar via morphology syntax interface in XLE environment.

3.2.1. Nominal Elements. Nominal elements include nouns, pronouns, adjectives, adverbs and phrases constituted by these elements. Different NP constructions implemented include: pronoun-noun, adjective-noun, and pronoun-adjective-noun combinations. These noun phrase combinations are further complicated by coordination, postpositional phrases and relative clauses. F-structure analysis of a noun phrase with demonstrative pronoun-noun combination is shown in Figure 4. Demonstrative pronoun "ihO" (this) is treated as a determiner in noun SPEC (specification). Different cases of nominal elements including nominative, accusative, dative, ablative, locative, instrumental, participant, genitive/possessive, agentive and vocative are taken care of. Different complications of syntactic case

marking are handled by defining a special case phrase



KP [14] which represents case marked noun phrase constructions.

Figure 4. Noun phrase with demonstrative pronoun.

Figure 5 shows an example of a case phrase with dative and accusative case marking. F-structure chart shows two possibilities of case "dat" and "acc"; the proper noun "Ali" therefore can either be in dative or accusative case as "khE" is case marker for both these cases. However, ambiguity of case of "Ali" will be resolved when other syntactic elements in the sentence are present and depending on whether "Ali" is direct object or indirect object or sometimes a dative subject.

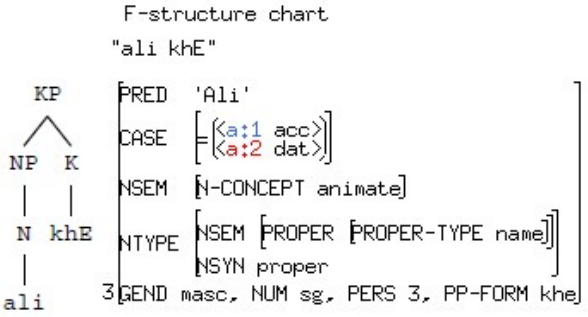


Figure 5. Dative and accusative case marking of noun.

For genitive case separate phrase KPPoss (possessive case phrase) is defined which reflects special agreement features required for agreement by different constituents of a sentence. LFG definition of KPPoss in XLE format is given below:

```
KPPoss -->    NP: {(! N-FORM)=c obl |
                   (! NTYPE NSYN)= proper}
                   ^=!;
              KPoss: ^=!.
```

LFG lexicon entry of KPoss (possessive case marker) "jO" showing extra attributes is given below.

```
jO        KPoss * (^ PP-FORM)=of
                  (^ K-NUM)=sg
                  (^ K-GEND)=masc
                  (^ K-FORM)=nom
                  (^ CASE)=gen.
```

7. It may be noted that extra attributes K-NUM, K-GEND, and K-FORM (K represents case) are introduced here to reflect the possessive case marker attributes to be agreed with possessed noun attributes. An example of KPPoss with genitive noun marking is shown in figure 6.

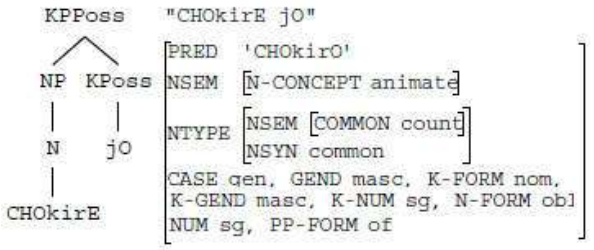


Figure 6. Genitive case marking with possessive case phrase.

3.2.2. Verbal Elements. Verbal elements include verbs which subcategorize (require arguments) for different grammatical functions. These grammatical functions include subject (SUBJ), object (OBJ), secondary object (OBJ2), oblique (OBL), PREDLINK, complement (COMP) and open complement XCOMP. Noun phrases (including all nominal elements) either define these functions or play essential role in their definition within a sentence. Sentence constituents therefore include verbs, their arguments and adjunct (ADJUNCT) elements which do not subcategorize for verbs. Different Verb categories include predicative verbs (main verbs and copula verbs), modal verbs and auxiliary verbs. In Sindhi main, auxiliary and modal verbs are combined to make verbal complex. Auxiliaries are also used to mark tense, aspect and mood. LFG implementation of verbal syntax includes verbal subcategorization for different grammatical functions listed above, verbal complex, and tense-aspect-mood analysis. Tense coverage include aorist formations, present, past and future tenses. Aspectual formations including perfective, imperfective-habitual and imperfective-continuous are analyzed by implemented LFG rules. Verb mood is also analyzed, coverage of different mood constructions includes: subjunctive, presumptive, imperative, declarative or indicative, permissive, prohibitive, capacitive, suggestive, and compulsive moods. A short version of sentence definition in LFG format is given below:

```
S--> NP:(^   SUBJ)=!   (!   GEND)=(^
GEND);)
       (KP: (^ OBJ2)=! (! CASE)=c dat)
```

```
(KP: (^ OBL)=! {(! CASE)=c inst
| (! CASE)=c agent})
(KP: (^ OBJ)=! {(! CASE)=c acc
| (! CASE)=c nom})
VC: (! NUM)=(^NUM) (! GEND)=(^
GEND) ^=!.
```

Above rules define sentence S as a sequence of noun phrase (NP) which is a subject, followed by optional case phrases (KPs) which include indirect object (OBJ2), oblique (OBL) and direct object (OBJ) followed by verb complex which may include combinations of different verb types. Above given rule defines the general structure of Sindhi sentence. Different constraints like (! GEND) = (^ GEND) and (! CASE=c dat) are placed to ensure gender case and number agreement. Consider following present tense sentence where subject and object are in nominative case.

Ali	KHatu	likhE	thO
Ali.Nom.M	Nom.M.Sg	Write.Aorist	Aux.Pres
Ali	Letter	Write	Be

Ali writes a letter.

Parse tree and functional structure analysis of above sentence are shown in figure 7 and figure 8 respectively. Subject and object case is identified morphologically, tense form in combination with present tense auxiliary "thO" identifies the tense, and aspect is also identified by morphological form of main verb "likhE" which is neither progressive nor perfective. Aspect is undefined therefore.

Consider following sentence:

Ali	CHOkirE-khE	KHatu
Ali.Nom.M	boy.Obl.Sg.M-Dat	letter.Nom.M.Sg

likhE	payO
write.Aorist.Sg	Aux.Cont

Ali is writing a letter to the boy.

It can be seen in above sentence that there are three verbal arguments, a subject "Ali", an indirect object "CHOkirO" in oblique form with dative case marker, and a direct object "KHatu" with nominative case. Verb aspect is continuous / progressive identified by "payO" auxiliary used as a continuity marker. Main verb "likhE" is in aorist form. Figure 9 and figure 10 show parse tree and f-structure analysis of above sentence respectively. Indirect object is subcategorized as OBJ2 with dative case. Auxiliaries "thO" and "payO" also define the indicative mood.

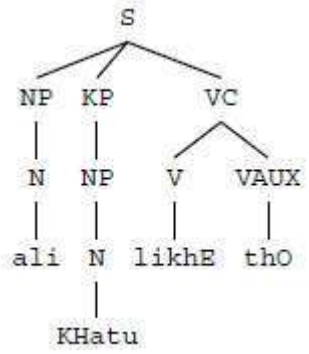


Figure 7. Sample present tense sentence with unspecified aspect.

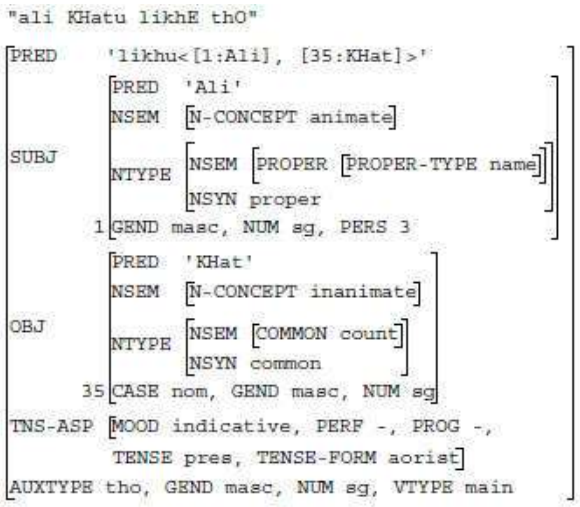


Figure 8. LFG analysis of sample sentence with SUBJ, OBJ subcategorization in present tense with undefined aspect.

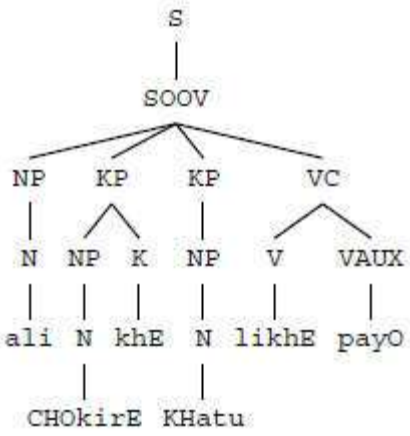


Figure 9. Sample sentence with imperfective continuous aspect.

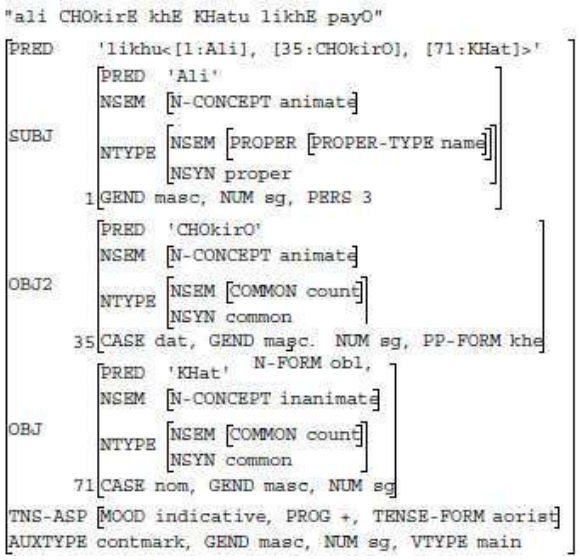


Figure 10: LFG analysis of sample sentence with SUBJ, OBJ, and OBJ2 subcategorization in aorist tense form with imperfective continuous aspect.

4. Pronominal Suffixes

Sindhi pronominal suffixes may appear with nouns, verbs, postpositions, and adverbs of place. Pronominal suffixes are treated as special lexical entries in lexicon. For example, consider transitive verb "likhu" (write); when appears with 1st person pronominal suffix "-iyami" becomes "likh-iyami" (I wrote). Morphological analysis of "likhiyami" is given below:

```
{likhiyami "+Token" | likhu "+Verb"
"+Psx" "+SSg" "+S1P" "+SMF"
"+SObl" "+Sg" "+PastPart"}
```

Above morphological analysis says that "likhiyami" is a morphological form of root "likhu". +Psx attribute says that this is a pronominal suffixed form. The tag pattern "+Sxxx" represent different attributes of subject reflected by pronominal suffix. +PastPart tag says that verb form is a past participle. F-structure analysis of "likhiyami" is shown in figure 11. It can be seen that different attributes of verb "likhu" in f-structure are extracted from morphological tags given in above morphological analysis. Pronominal suffixation may cause a complete sentence replaced with single word form with all its verbal and nominal elements. Syntax analysis therefore needs to extract / reconstruct this information from morphology. In this case the sentence "mUN likhiyO" (I wrote) is replaced by "likhiyami". This reconstruction can be seen in verbal subcategorization of "likhu" where the SUBJ argument contains the value 'pro' which represents a pronoun with gender, noun form, number and person attributes (feminine, oblique, singular, 1st person). Oblique singular 1st person pronoun in Sindhi is "mUN" which can either be feminine or masculine. As the verb is in past participle form therefore its aspect is perfective.

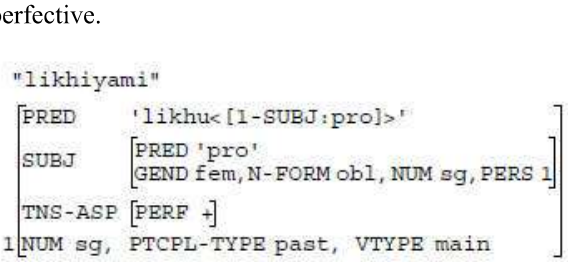


Figure 11: F-structure of pronominal suffixed verb "likhiyami".

5. Coverage

Morphological coverage includes: finite state models of nouns, pronouns, adjectives, adverbs and verbs. Full form LFG lexicon of postpositions, conjunctions and few adverbs. Case, mood, tense and aspect morphology of nominal and verbal elements is also implemented. Table 2 shows some figures about morphology coverage. Interestingly adjectives have more average inflections per stem as compared to nouns. This is due to degree change inflections of native Sindhi adjectives where inflections are doubled as compared to nouns. For example, a masculine noun with 'O' ending can have up to 12 inflections and an adjective with 'O' ending will also have almost 12 inflections. However, with internal morphological change when degree changing morphology is applied number of inflections becomes double. For example, adjective 'naNdHO' (small) becomes 'naNdHaRO' all inflections of naNdHO will also be applied to naNdHaRO as well and this will double the number of inflectional forms. Pronoun inflections per stem is also 3.58 due to number gender and case inflections (mostly in wh-pronons).

Syntax coverage include noun phrase constructions with all nominal elements, verbal subcategorization with SUB, OBJ, OBL, OBJ2, COM, XCOMP, ADJUNCT, XADJUNCT, and PREDLINK, coordination, subordination, mood, case, aspect, tense, and agreement.

5. Conclusion and Future Work

Development in current state covers the morphological and syntactic constructions discussed in above sections. Basic morphology and syntax

constructs in Sindhi are identified and modeled. Morphological analysis shows interesting results like adjectives have more average inflections than nouns, and pronouns have 3.58 average inflections per word. Also verb can have up to 75 different morphological forms. Though the basic constructs of Sindhi morphology and Syntax are implemented yet many complexities are subject to further research and development including: pronominal suffixation with nominal elements, pronominal suffixation with postpositions, NP coordination model, verbal complex constructions which form complex predicates, and pro-drop. Also the morphological lexicon size and coverage requires more enhancements. Developed model will be tested against synthesized test-suit covering all morphology and syntax patterns implemented and real time corpus test suit being developed.

Table 2: Morphology coverage

Word Class	Stems	Morphological Forms / Inflections	Average Inflections / Stem
Verbs	100	5013	50.13
Nouns	323	1729	5.35
Pronouns	79	283	3.58
Adjectives	71	394	5.55
Adverbs	38	38	1.00
Total	**611**	**7457**	**12.20**

6. References

[1] K. R. Beesley, and K. Lauri. "Finite-state morphology: Xerox tools and techniques." CSLI, Stanford (2003).

[2] C. Dick, M. Dalrymple, R. Kaplan, T. H. King, John Maxwell, and Paula Newman. "XLE documentation." Palo Alto Research Center (2008).

[3] M. U. Rahman. "Sindhi Morphology and Noun Inflections", in proc. *Conference on Language & Technology (CLT-09)*, pp. 74-81. 2009.

[4] R. Emmanuel, and Y. Schabes. Finite-state language processing. MIT press, 1997.

[5] K. R. Beesly. "Arabic Morphology Using Only FinateState Operations", in proc. *Workshop on Computational Approaches to Semetic languages*, Montreal, Quebec, pp. 50-57. (1998).

[6] M. J. Steedman. A Generative Grammar for Jazz Chord Sequences. Music Perception 2 (1): 52–77. JSTOR 40285282. (1989).

[7] M. U. Rahman., and M. I. Bhatti. "Finite State Morphology and Sindhi Noun Inflections.", in proc. *Pacific Asia Conference on Language, Information and Computation (PACLIC 24)*. Sendai, Japan. pp. 669 – 676 (2010).

[8] M. U. Rahman, A. Shah "Grammar Checking Model for Local Languages.", in proc. *SCONEST (Student Conference on Engineering Sciences and Technology)* Karachi. (2003).

[9] M U. Rahman, A. Shah, R.A. Memon. Partial Word Order Syntax of Urdu/Sindhi and Linear Specification Language. *Journal of Independent Studies and Research (JISR) Volume 5*, Number2, July 2007. pp. 13 – 18.

[10] J. D. Oad. *Implementing GF Resource Grammar for Sindhi*. Unpublished Master's Thesis. Department of Applied Information Technology Chalmers University of Technology Gothenburg, Sweden. (2012).

[11] A. Ranta. "Grammatical Framework: A Type-Theoretical Grammar Formalism", *Journal of Functional Programming* 14 (2): 145–189. (2004).

[12] M. Butt, *The Structure of Complex Predicates in Urdu,* CSLI Publications, Stanford. (1995).

[13] M. Butt, D. Helge, T. H. King, H. Masuichi, and C. Rohrer. "The parallel grammar project." in proc. "*Workshop on Grammar engineering and evaluation*" Volume 15, pp. 1-7. Association for Computational Linguistics, 2002.

[14] T. Bögel,, M. Butt, A. Hautli, and S. Sulger. "Urdu and the modular architecture of ParGram", in proc. *Conference on Language and Technology*, vol. 70. Lahore. 2009.

[15] S. M. J. Rizvi. *Development of Alorithms and Computational Grammar for Urdu*, PhD thesis. ept. of Computer and Information Science PIAS. Islamabad. 2007

[16] L. Karttunen, *Finite-State Lexicon Compiler*. Technical Report, ISTL-NLTT2993-04-02, Xerox Palo Alto Research Center, Palo Alto, California (1993)

[17] L. Karttunen, and K. R. Beesley. "Twenty-five years of finite-state morphology." *Inquiries Into Words*, a Festschrift for Kimmo Koskenniemi on his 60th Birthday (2005): 71-83.

[18] M. Dalrymple. *Lexical-Functional Grammar*. John Wiley & Sons, Ltd, 2001.