

Development of Sindhi Lexical Functional Grammar

Muttee U Rahman & Hameedullah Kazi

Isra University, Hyderabad

Outline

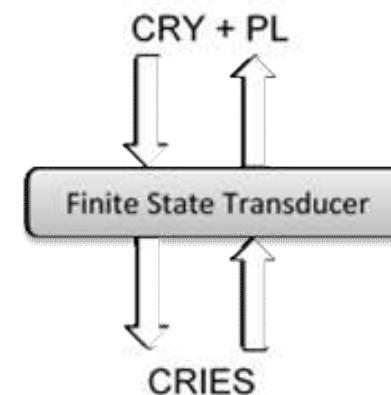
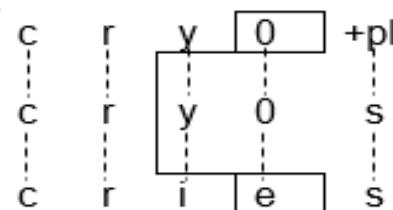
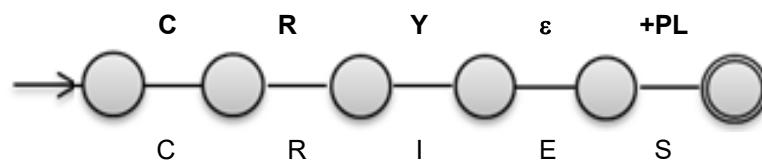
- Introduction
 - Background
 - Finite State Morphology
 - Lexical Functional Grammar
- Overall Development Model
- Implementing Sindhi Morphology
- Implementing Sindhi Syntax
- Coverage
- Conclusion

Background

- Presented work is about development of Sindhi Grammar
- Frameworks used include: Finite State Morphology and Lexical Functional Grammar
- Xerox Finite State Morphology Tools (XFST) and Xerox Linguistic Environment (XLE) are used for Implementation

Finite State Morphology

Singular	Intermediate	Plural	Rule
CRY	CRYS	CRIES	$y \rightarrow ie / ^_ s\#$



Lexical Functional Grammar

- Grammar based on generative grammars (Steedman, 1989), (Dalrymple, 2001)
- Defines linguistic structure at three different levels
 - Lexicon
 - C-structure (Constituent Structure)
 - F-structure (Functional Structure)

Lexical Functional Grammar

Lexicon

mAryO V (\uparrow PRED) = 'mAru<(\uparrow SUBJ), (\uparrow OBJ)>'

(\uparrow TENNSE) = Past

(\uparrow SUBJ NUM) = SG

(\uparrow SUBJ PERS) = 3

Ali N (\uparrow PRED) = 'Ali'

(\uparrow NUM) = SG

(\uparrow PERS) = 3

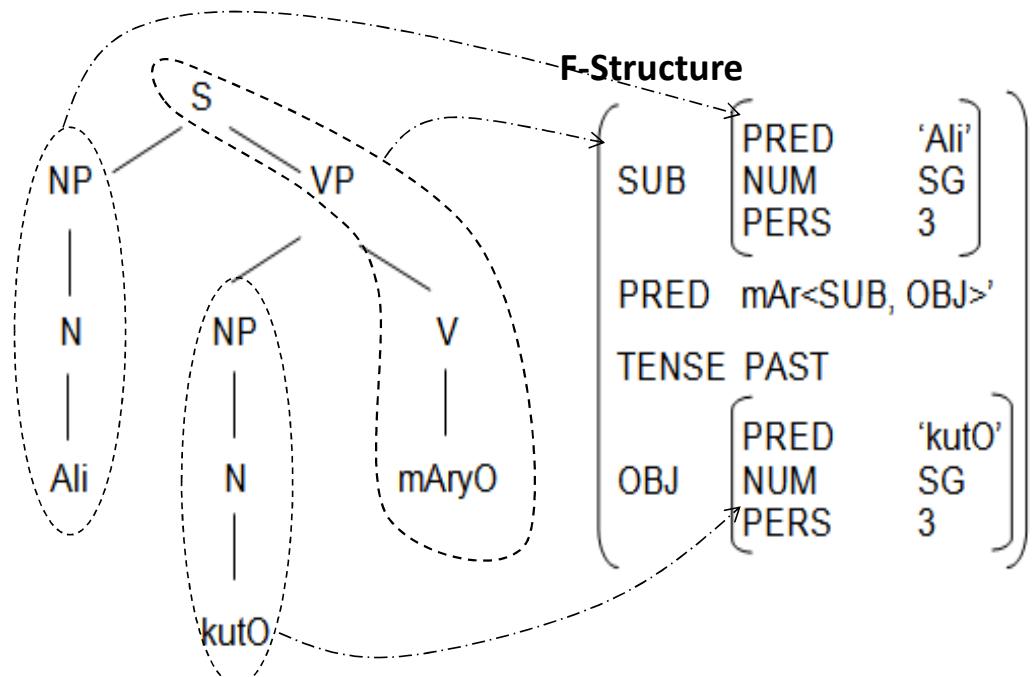
C-Structure Rules

1. $S \rightarrow NP \quad VP$
 $(\uparrow \text{SUBJ}=\downarrow) \quad \uparrow=\downarrow$

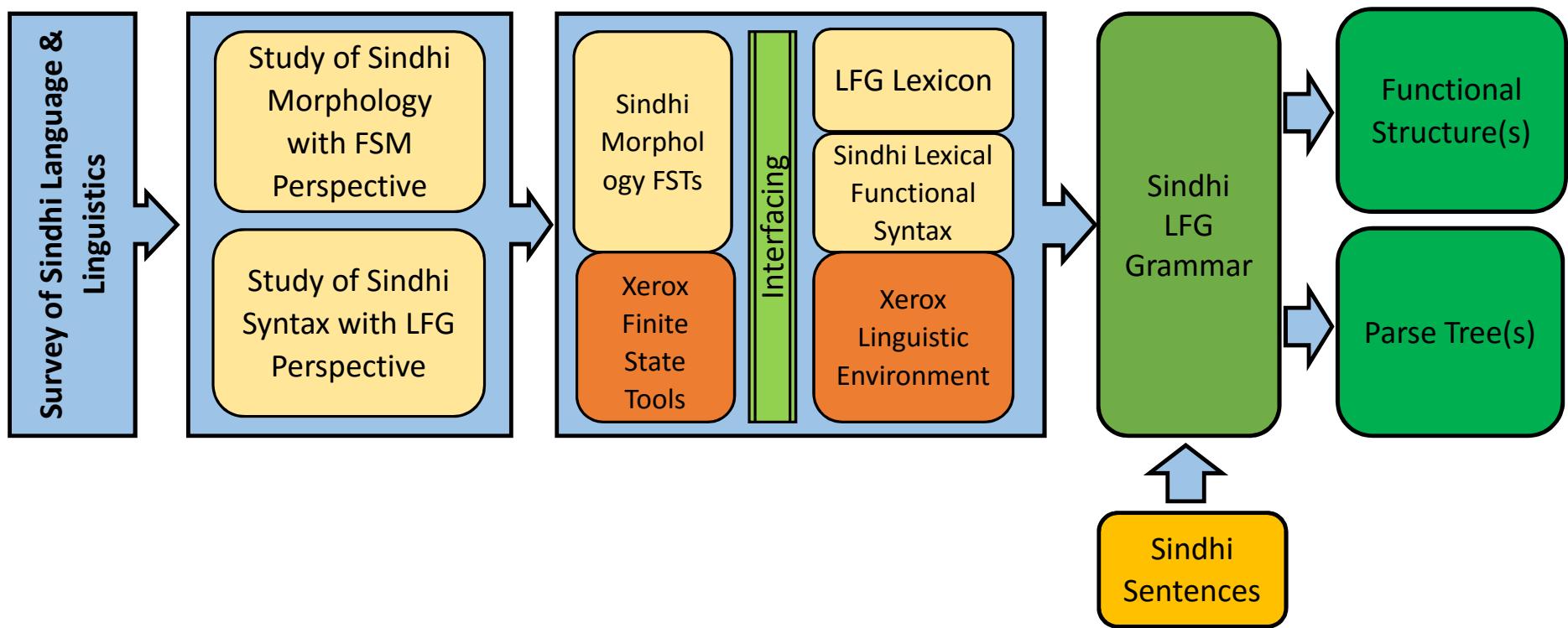
2. $NP \rightarrow N$
 $\uparrow=\downarrow$

3. $VP \rightarrow NP \quad V$

4. -----



Grammar Engineering Process



Implementing Morphology

- Morphological paradigms of different POS classes are modeled by incorporating the inflection rules in FSTs using XFST scripts

Implementing Morphology

! SINDHI NOUN MORPHOLOGY

Multichar_Symbols

+Noun +Adjective +Adverb +Verb
+Common +Proper +Abstract !Noun Types
+Animate +Inanimate !Noun Consent

+A₁ +A₂ +A₃ +A₄ +A₅ +A₆ +A₇ +A₈ +A₉ +A₁₀ +A₁₁ +A₁₂ +A₁₃ +A₁₄ +A₁₅ +A₁₆ +A₁₇ +A₁₈ +A₁₉ +A₂₀ +A₂₁ +A₂₂ +A₂₃ +A₂₄ +A₂₅ +A₂₆ +A₂₇ +A₂₈ +A₂₉ +A₃₀ +A₃₁ +A₃₂ +A₃₃ +A₃₄ +A₃₅ +A₃₆ +A₃₇ +A₃₈ +A₃₉ +A₄₀ +A₄₁ +A₄₂ +A₄₃ +A₄₄ +A₄₅ +A₄₆ +A₄₇ +A₄₈ +A₄₉ +A₅₀ +A₅₁ +A₅₂ +A₅₃ +A₅₄ +A₅₅ +A₅₆ +A₅₇ +A₅₈ +A₅₉ +A₆₀ +A₆₁ +A₆₂ +A₆₃ +A₆₄ +A₆₅ +A₆₆ +A₆₇ +A₆₈ +A₆₉ +A₇₀ +A₇₁ +A₇₂ +A₇₃ +A₇₄ +A₇₅ +A₇₆ +A₇₇ +A₇₈ +A₇₉ +A₈₀ +A₈₁ +A₈₂ +A₈₃ +A₈₄ +A₈₅ +A₈₆ +A₈₇ +A₈₈ +A₈₉ +A₉₀ +A₉₁ +A₉₂ +A₉₃ +A₉₄ +A₉₅ +A₉₆ +A₉₇ +A₉₈ +A₉₉ +A₁₀₀ +A₁₀₁ +A₁₀₂ +A₁₀₃ +A₁₀₄ +A₁₀₅ +A₁₀₆ +A₁₀₇ +A₁₀₈ +A₁₀₉ +A₁₁₀ +A₁₁₁ +A₁₁₂ +A₁₁₃ +A₁₁₄ +A₁₁₅ +A₁₁₆ +A₁₁₇ +A₁₁₈ +A₁₁₉ +A₁₂₀ +A₁₂₁ +A₁₂₂ +A₁₂₃ +A₁₂₄ +A₁₂₅ +A₁₂₆ +A₁₂₇ +A₁₂₈ +A₁₂₉ +A₁₃₀ +A₁₃₁ +A₁₃₂ +A₁₃₃ +A₁₃₄ +A₁₃₅ +A₁₃₆ +A₁₃₇ +A₁₃₈ +A₁₃₉ +A₁₄₀ +A₁₄₁ +A₁₄₂ +A₁₄₃ +A₁₄₄ +A₁₄₅ +A₁₄₆ +A₁₄₇ +A₁₄₈ +A₁₄₉ +A₁₅₀ +A₁₅₁ +A₁₅₂ +A₁₅₃ +A₁₅₄ +A₁₅₅ +A₁₅₆ +A₁₅₇ +A₁₅₈ +A₁₅₉ +A₁₆₀ +A₁₆₁ +A₁₆₂ +A₁₆₃ +A₁₆₄ +A₁₆₅ +A₁₆₆ +A₁₆₇ +A₁₆₈ +A₁₆₉ +A₁₇₀ +A₁₇₁ +A₁₇₂ +A₁₇₃ +A₁₇₄ +A₁₇₅ +A₁₇₆ +A₁₇₇ +A₁₇₈ +A₁₇₉ +A₁₈₀ +A₁₈₁ +A₁₈₂ +A₁₈₃ +A₁₈₄ +A₁₈₅ +A₁₈₆ +A₁₈₇ +A₁₈₈ +A₁₈₉ +A₁₉₀ +A₁₉₁ +A₁₉₂ +A₁₉₃ +A₁₉₄ +A₁₉₅ +A₁₉₆ +A₁₉₇ +A₁₉₈ +A₁₉₉ +A₂₀₀ +A₂₀₁ +A₂₀₂ +A₂₀₃ +A₂₀₄ +A₂₀₅ +A₂₀₆ +A₂₀₇ +A₂₀₈ +A₂₀₉ +A₂₁₀ +A₂₁₁ +A₂₁₂ +A₂₁₃ +A₂₁₄ +A₂₁₅ +A₂₁₆ +A₂₁₇ +A₂₁₈ +A₂₁₉ +A₂₂₀ +A₂₂₁ +A₂₂₂ +A₂₂₃ +A₂₂₄ +A₂₂₅ +A₂₂₆ +A₂₂₇ +A₂₂₈ +A₂₂₉ +A₂₃₀ +A₂₃₁ +A₂₃₂ +A₂₃₃ +A₂₃₄ +A₂₃₅ +A₂₃₆ +A₂₃₇ +A₂₃₈ +A₂₃₉ +A₂₄₀ +A₂₄₁ +A₂₄₂ +A₂₄₃ +A₂₄₄ +A₂₄₅ +A₂₄₆ +A₂₄₇ +A₂₄₈ +A₂₄₉ +A₂₅₀ +A₂₅₁ +A₂₅₂ +A₂₅₃ +A₂₅₄ +A₂₅₅ +A₂₅₆ +A₂₅₇ +A₂₅₈ +A₂₅₉ +A₂₆₀ +A₂₆₁ +A₂₆₂ +A₂₆₃ +A₂₆₄ +A₂₆₅ +A₂₆₆ +A₂₆₇ +A₂₆₈ +A₂₆₉ +A₂₇₀ +A₂₇₁ +A₂₇₂ +A₂₇₃ +A₂₇₄ +A₂₇₅ +A₂₇₆ +A₂₇₇ +A₂₇₈ +A₂₇₉ +A₂₈₀ +A₂₈₁ +A₂₈₂ +A₂₈₃ +A₂₈₄ +A₂₈₅ +A₂₈₆ +A₂₈₇ +A₂₈₈ +A₂₈₉ +A₂₉₀ +A₂₉₁ +A₂₉₂ +A₂₉₃ +A₂₉₄ +A₂₉₅ +A₂₉₆ +A₂₉₇ +A₂₉₈ +A₂₉₉ +A₃₀₀ +A₃₀₁ +A₃₀₂ +A₃₀₃ +A₃₀₄ +A₃₀₅ +A₃₀₆ +A₃₀₇ +A₃₀₈ +A₃₀₉ +A₃₁₀ +A₃₁₁ +A₃₁₂ +A₃₁₃ +A₃₁₄ +A₃₁₅ +A₃₁₆ +A₃₁₇ +A₃₁₈ +A₃₁₉ +A₃₂₀ +A₃₂₁ +A₃₂₂ +A₃₂₃ +A₃₂₄ +A₃₂₅ +A₃₂₆ +A₃₂₇ +A₃₂₈ +A₃₂₉ +A₃₃₀ +A₃₃₁ +A₃₃₂ +A₃₃₃ +A₃₃₄ +A₃₃₅ +A₃₃₆ +A₃₃₇ +A₃₃₈ +A₃₃₉ +A₃₄₀ +A₃₄₁ +A₃₄₂ +A₃₄₃ +A₃₄₄ +A₃₄₅ +A₃₄₆ +A₃₄₇ +A₃₄₈ +A₃₄₉ +A₃₅₀ +A₃₅₁ +A₃₅₂ +A₃₅₃ +A₃₅₄ +A₃₅₅ +A₃₅₆ +A₃₅₇ +A₃₅₈ +A₃₅₉ +A₃₆₀ +A₃₆₁ +A₃₆₂ +A₃₆₃ +A₃₆₄ +A₃₆₅ +A₃₆₆ +A₃₆₇ +A₃₆₈ +A₃₆₉ +A₃₇₀ +A₃₇₁ +A₃₇₂ +A₃₇₃ +A₃₇₄ +A₃₇₅ +A₃₇₆ +A₃₇₇ +A₃₇₈ +A₃₇₉ +A₃₈₀ +A₃₈₁ +A₃₈₂ +A₃₈₃ +A₃₈₄ +A₃₈₅ +A₃₈₆ +A₃₈₇ +A₃₈₈ +A₃₈₉ +A₃₉₀ +A₃₉₁ +A₃₉₂ +A₃₉₃ +A₃₉₄ +A₃₉₅ +A₃₉₆ +A₃₉₇ +A₃₉₈ +A₃₉₉ +A₄₀₀ +A₄₀₁ +A₄₀₂ +A₄₀₃ +A₄₀₄ +A₄₀₅ +A₄₀₆ +A₄₀₇ +A₄₀₈ +A₄₀₉ +A₄₁₀ +A₄₁₁ +A₄₁₂ +A₄₁₃ +A₄₁₄ +A₄₁₅ +A₄₁₆ +A₄₁₇ +A₄₁₈ +A₄₁₉ +A₄₂₀ +A₄₂₁ +A₄₂₂ +A₄₂₃ +A₄₂₄ +A₄₂₅ +A₄₂₆ +A₄₂₇ +A₄₂₈ +A₄₂₉ +A₄₃₀ +A₄₃₁ +A₄₃₂ +A₄₃₃ +A₄₃₄ +A₄₃₅ +A₄₃₆ +A₄₃₇ +A₄₃₈ +A₄₃₉ +A₄₄₀ +A₄₄₁ +A₄₄₂ +A₄₄₃ +A₄₄₄ +A₄₄₅ +A₄₄₆ +A₄₄₇ +A₄₄₈ +A₄₄₉ +A₄₅₀ +A₄₅₁ +A₄₅₂ +A₄₅₃ +A₄₅₄ +A₄₅₅ +A₄₅₆ +A₄₅₇ +A₄₅₈ +A₄₅₉ +A₄₆₀ +A₄₆₁ +A₄₆₂ +A₄₆₃ +A₄₆₄ +A₄₆₅ +A₄₆₆ +A₄₆₇ +A₄₆₈ +A₄₆₉ +A₄₇₀ +A₄₇₁ +A₄₇₂ +A₄₇₃ +A₄₇₄ +A₄₇₅ +A₄₇₆ +A₄₇₇ +A₄₇₈ +A₄₇₉ +A₄₈₀ +A₄₈₁ +A₄₈₂ +A₄₈₃ +A₄₈₄ +A₄₈₅ +A₄₈₆ +A₄₈₇ +A₄₈₈ +A₄₈₉ +A₄₉₀ +A₄₉₁ +A₄₉₂ +A₄₉₃ +A₄₉₄ +A₄₉₅ +A₄₉₆ +A₄₉₇ +A₄₉₈ +A₄₉₉ +A₅₀₀ +A₅₀₁ +A₅₀₂ +A₅₀₃ +A₅₀₄ +A₅₀₅ +A₅₀₆ +A₅₀₇ +A₅₀₈ +A₅₀₉ +A₅₁₀ +A₅₁₁ +A₅₁₂ +A₅₁₃ +A₅₁₄ +A₅₁₅ +A₅₁₆ +A₅₁₇ +A₅₁₈ +A₅₁₉ +A₅₂₀ +A₅₂₁ +A₅₂₂ +A₅₂₃ +A₅₂₄ +A₅₂₅ +A₅₂₆ +A₅₂₇ +A₅₂₈ +A₅₂₉ +A₅₃₀ +A₅₃₁ +A₅₃₂ +A₅₃₃ +A₅₃₄ +A₅₃₅ +A₅₃₆ +A₅₃₇ +A₅₃₈ +A₅₃₉ +A₅₄₀ +A₅₄₁ +A₅₄₂ +A₅₄₃ +A₅₄₄ +A₅₄₅ +A₅₄₆ +A₅₄₇ +A₅₄₈ +A₅₄₉ +A₅₅₀ +A₅₅₁ +A₅₅₂ +A₅₅₃ +A₅₅₄ +A₅₅₅ +A₅₅₆ +A₅₅₇ +A₅₅₈ +A₅₅₉ +A₅₆₀ +A₅₆₁ +A₅₆₂ +A₅₆₃ +A₅₆₄ +A₅₆₅ +A₅₆₆ +A₅₆₇ +A₅₆₈ +A₅₆₉ +A₅₇₀ +A₅₇₁ +A₅₇₂ +A₅₇₃ +A₅₇₄ +A₅₇₅ +A₅₇₆ +A₅₇₇ +A₅₇₈ +A₅₇₉ +A₅₈₀ +A₅₈₁ +A₅₈₂ +A₅₈₃ +A₅₈₄ +A₅₈₅ +A₅₈₆ +A₅₈₇ +A₅₈₈ +A₅₈₉ +A₅₉₀ +A₅₉₁ +A₅₉₂ +A₅₉₃ +A₅₉₄ +A₅₉₅ +A₅₉₆ +A₅₉₇ +A₅₉₈ +A₅₉₉ +A₆₀₀ +A₆₀₁ +A₆₀₂ +A₆₀₃ +A₆₀₄ +A₆₀₅ +A₆₀₆ +A₆₀₇ +A₆₀₈ +A₆₀₉ +A₆₁₀ +A₆₁₁ +A₆₁₂ +A₆₁₃ +A₆₁₄ +A₆₁₅ +A₆₁₆ +A₆₁₇ +A₆₁₈ +A₆₁₉ +A₆₂₀ +A₆₂₁ +A₆₂₂ +A₆₂₃ +A₆₂₄ +A₆₂₅ +A₆₂₆ +A₆₂₇ +A₆₂₈ +A₆₂₉ +A₆₃₀ +A₆₃₁ +A₆₃₂ +A₆₃₃ +A₆₃₄ +A₆₃₅ +A₆₃₆ +A₆₃₇ +A₆₃₈ +A₆₃₉ +A₆₄₀ +A₆₄₁ +A₆₄₂ +A₆₄₃ +A₆₄₄ +A₆₄₅ +A₆₄₆ +A₆₄₇ +A₆₄₈ +A₆₄₉ +A₆₅₀ +A₆₅₁ +A₆₅₂ +A₆₅₃ +A₆₅₄ +A₆₅₅ +A₆₅₆ +A₆₅₇ +A₆₅₈ +A₆₅₉ +A₆₆₀ +A₆₆₁ +A₆₆₂ +A₆₆₃ +A₆₆₄ +A₆₆₅ +A₆₆₆ +A₆₆₇ +A₆₆₈ +A₆₆₉ +A₆₇₀ +A₆₇₁ +A₆₇₂ +A₆₇₃ +A₆₇₄ +A₆₇₅ +A₆₇₆ +A₆₇₇ +A₆₇₈ +A₆₇₉ +A₆₈₀ +A₆₈₁ +A₆₈₂ +A₆₈₃ +A₆₈₄ +A₆₈₅ +A₆₈₆ +A₆₈₇ +A₆₈₈ +A₆₈₉ +A₆₉₀ +A₆₉₁ +A₆₉₂ +A₆₉₃ +A₆₉₄ +A₆₉₅ +A₆₉₆ +A₆₉₇ +A₆₉₈ +A₆₉₉ +A₇₀₀ +A₇₀₁ +A₇₀₂ +A₇₀₃ +A₇₀₄ +A₇₀₅ +A₇₀₆ +A₇₀₇ +A₇₀₈ +A₇₀₉ +A₇₁₀ +A₇₁₁ +A₇₁₂ +A₇₁₃ +A₇₁₄ +A₇₁₅ +A₇₁₆ +A₇₁₇ +A₇₁₈ +A₇₁₉ +A₇₂₀ +A₇₂₁ +A₇₂₂ +A₇₂₃ +A₇₂₄ +A₇₂₅ +A₇₂₆ +A₇₂₇ +A₇₂₈ +A₇₂₉ +A₇₃₀ +A₇₃₁ +A₇₃₂ +A₇₃₃ +A₇₃₄ +A₇₃₅ +A₇₃₆ +A₇₃₇ +A₇₃₈ +A₇₃₉ +A₇₄₀ +A₇₄₁ +A₇₄₂ +A₇₄₃ +A₇₄₄ +A₇₄₅ +A₇₄₆ +A₇₄₇ +A₇₄₈ +A₇₄₉ +A₇₅₀ +A₇₅₁ +A₇₅₂ +A₇₅₃ +A₇₅₄ +A₇₅₅ +A₇₅₆ +A₇₅₇ +A₇₅₈ +A₇₅₉ +A₇₆₀ +A₇₆₁ +A₇₆₂ +A₇₆₃ +A₇₆₄ +A₇₆₅ +A₇₆₆ +A₇₆₇ +A₇₆₈ +A₇₆₉ +A₇₇₀ +A₇₇₁ +A₇₇₂ +A₇₇₃ +A₇₇₄ +A₇₇₅ +A₇₇₆ +A₇₇₇ +A₇₇₈ +A₇₇₉ +A₇₈₀ +A₇₈₁ +A₇₈₂ +A₇₈₃ +A₇₈₄ +A₇₈₅ +A₇₈₆ +A₇₈₇ +A₇₈₈ +A₇₈₉ +A₇₉₀ +A₇₉₁ +A₇₉₂ +A₇₉₃ +A₇₉₄ +A₇₉₅ +A₇₉₆ +A₇₉₇ +A₇₉₈ +A₇₉₉ +A₈₀₀ +A₈₀₁ +A₈₀₂ +A₈₀₃ +A₈₀₄ +A₈₀₅ +A₈₀₆ +A₈₀₇ +A₈₀₈ +A₈₀₉ +A₈₁₀ +A₈₁₁ +A₈₁₂ +A₈₁₃ +A₈₁₄ +A₈₁₅ +A₈₁₆ +A₈₁₇ +A₈₁₈ +A₈₁₉ +A₈₂₀ +A₈₂₁ +A₈₂₂ +A₈₂₃ +A₈₂₄ +A₈₂₅ +A₈₂₆ +A₈₂₇ +A₈₂₈ +A₈₂₉ +A₈₃₀ +A₈₃₁ +A₈₃₂ +A₈₃₃ +A₈₃₄ +A₈₃₅ +A₈₃₆ +A₈₃₇ +A₈₃₈ +A₈₃₉ +A₈₄₀ +A₈₄₁ +A₈₄₂ +A₈₄₃ +A₈₄₄ +A₈₄₅ +A₈₄₆ +A₈₄₇ +A₈₄₈ +A₈₄₉ +A₈₅₀ +A₈₅₁ +A₈₅₂ +A₈₅₃ +A₈₅₄ +A₈₅₅ +A₈₅₆ +A₈₅₇ +A₈₅₈ +A₈₅₉ +A₈₆₀ +A₈₆₁ +A₈₆₂ +A₈₆₃ +A₈₆₄ +A₈₆₅ +A₈₆₆ +A₈₆₇ +A₈₆₈ +A₈₆₉ +A₈₇₀ +A₈₇₁ +A₈₇₂ +A₈₇₃ +A₈₇₄ +A₈₇₅ +A₈₇₆ +A₈₇₇ +A₈₇₈ +A₈₇₉ +A₈₈₀ +A₈₈₁ +A₈₈₂ +A₈₈₃ +A₈₈₄ +A₈₈₅ +A₈₈₆ +A₈₈₇ +A₈₈₈ +A₈₈₉ +A₈₉₀ +A₈₉₁ +A₈₉₂ +A₈₉₃ +A₈₉₄ +A₈₉₅ +A₈₉₆ +A₈₉₇ +A₈₉₈ +A₈₉₉ +A₉₀₀ +A₉₀₁ +A₉₀₂ +A₉₀₃ +A₉₀₄ +A₉₀₅ +A₉₀₆ +A₉₀₇ +A₉₀₈ +A₉₀₉ +A₉₁₀ +A₉₁₁ +A₉₁₂ +A₉₁₃ +A₉₁₄ +A₉₁₅ +A₉₁₆ +A₉₁₇ +A₉₁₈ +A₉₁₉ +A₉₂₀ +A₉₂₁ +A₉₂₂ +A₉₂₃ +A₉₂₄ +A₉₂₅ +A₉₂₆ +A₉₂₇ +A₉₂₈ +A₉₂₉ +A₉₃₀ +A₉₃₁ +A₉₃₂ +A₉₃₃ +A₉₃₄ +A₉₃₅ +A₉₃₆ +A₉₃₇ +A₉₃₈ +A₉₃₉ +A₉₄₀ +A₉₄₁ +A₉₄₂ +A₉₄₃ +A₉₄₄ +A₉₄₅ +A₉₄₆ +A₉₄₇ +A₉₄₈ +A₉₄₉ +A₉₅₀ +A₉₅₁ +A₉₅₂ +A₉₅₃ +A₉₅₄ +A₉₅₅ +A₉₅₆ +A₉₅₇ +A₉₅₈ +A₉₅₉ +A₉₆₀ +A₉₆₁ +A₉₆₂ +A₉₆₃ +A₉₆₄ +A₉₆₅ +A₉₆₆ +A₉₆₇ +A₉₆₈ +A₉₆₉ +A₉₇₀ +A₉₇₁ +A₉₇₂ +A₉₇₃ +A₉₇₄ +A₉₇₅ +A₉₇₆ +A₉₇₇ +A₉₇₈ +A₉₇₉ +A₉₈₀ +A₉₈₁ +A₉₈₂ +A₉₈₃ +A₉₈₄ +A₉₈₅ +A₉₈₆ +A₉₈₇ +A₉₈₈ +A₉₈₉ +A₉₉₀ +A₉₉₁ +A₉₉₂ +A₉₉₃ +A₉₉₄ +A₉₉₅ +A₉₉₆ +A₉₉₇ +A₉₉₈ +A₉₉₉ +A₉₉₉

! Boy (Animate Common Noun)

CHOkir+Noun+Common+Count+Animate:N Cat1;

...

LEXICON N Cat1;

+Sg+Masc+Nominative:O #;

+Sg+Masc+Oblique:E #;

+Sg+Masc+Voc

Implementing Morphology

- Following inflections are handled (wherever applicable)
 - Number (CHOkirO, CHOkirA)
 - Gender (CHOkiro, CHOkirIa)
 - Case (CHOkirO, CHOkirE)
 - Tense (likHu, likHAN, likHiyO)
(AhE, huO, hUNdO)
 - Aspect (likHu, likHando)
 - Mood (likHu, likHijANI)
-
- The diagram illustrates the morphological features listed in the slide. It uses curly braces to group the features into two main categories: 'Noun, Pronoun, Adj, Adv, Postposition, Verb' and 'Verb'. The first category groups Number, Gender, Case, and Tense (in blue). The second category groups Aspect and Mood (both in blue). The Tense feature is also explicitly noted as 'not yet analyzed by Sindhi Grammarians' in red text at the bottom of the list.
- Noun, Pronoun, Adj, Adv, Postposition, Verb**
- Verb**
- Tense Aspect and Mood not yet analyzed by Sindhi Grammarians**

Noun Cases

- Case

Case	Case Marker	Example	
Nominative	Ø	CHOkirO	CHOkirO
Accusative / Dative	-E	CHOkir-E	
Postpositional	-E	CHOkir-E	
Locative	-E	CHOkir-E	
Instrumental	-E	sONT-E sAN	
Possessive / Genitive	-E	CHOkir-E JO	
Ablative	-AN	gHaru	
Vocative	-A	CHOkirO	

Oblique
Form

- Noun case morphology is further complicated by number and gender inflections in combination with cases

Pronouns

- Pronouns are declined for number and gender
- Marked by **Nominative, Oblique and Genitive** Cases

Case	Masculine	Feminine
Nom.Sg.	kehRO: CHOKirO	kehRI CHOKirI
Nom.pl.	kehRA CHOKirA	kehRyUN CHOKirUN
Obl.sg	kehRE CHOKirE	kehRIa CHOKirIa
Obl.pl	kehRani CHOKirani	kehRiyuni CHOKiruni
Gen.sg.	muhinjo CHOKirO	muhinJI: CHOKirI
Gen.pl.	muhinja CHOKirA	muhinjUN CHOKiriUN

Pronominal Suffixes

- Sindhi is one of few Indo-Aryan languages with pronominal suffixes
- Three types of pronominal suffixes are

S.No.	Pronominal Suffix Type		Syntactic Role	Example	
1	Nominal Suffix	اسمیہ ضمیر متصل	Noun	پُٹم، پُٹس، چاچھین	puTa-mi
2	Verbal Suffix	فعلیہ ضمیر متصل	Verb	ماریانس، اثتوں، لکندم	mAri-yAN-si
3	Postpositional Suffix	جري ضمیر متصل	Pronoun	کین، سائنس، وئنون	kHE-na

Verbs

- Verbs are further classified into
 - Main Verbs (Transitive & Intransitive)
 - Compound / Complex Verbs
 - Participles (Present Participle, Past Participle, Future Participle, Verbal Noun, Conjunctive Participle)
 - Infinitives
 - Auxiliary
 - Copula
 - Modal

Implementing Syntax

- Nominal Elements
 - Nouns, Pronouns, Adjectives, Adverbs
 - Phrases constituted by above elements
 - Complicated by coordination, postpositional phrases and relative clauses and Cases Marking
- Verbal Elements
 - Verb Subcategorization
 - SUBJ, OBJ, OBJ2, OBL, PREDLINK, COMP, XCOMP
- Adjuncts
 - ADJUNCT, XADJUNCT (Open Adjuncts)

NP Constructions

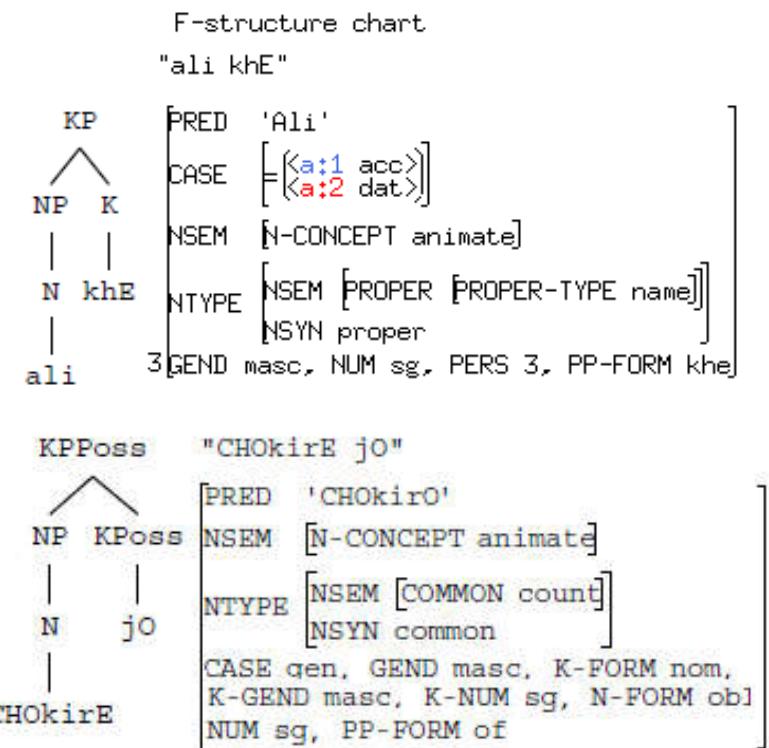
- Noun (CHOkirO)
- Pronoun-Noun (ihO CHOkirO)
- Adj-Noun (suTHO CHOkirO)
- Pronoun-Adj-Noun
(ihO suTHO CHOkirO)

"ihO CHOkirO"

```
PRED  'CHOkirO'  
NSEM  [N-CONCEPT animate]  
NTYPE [NSEM [COMMON count]]  
      [NSYN common]  
  
SPEC  [DET  1[PRED  'ihO'  
          NTYPE [NSYN pronour]]]  
      38  91[CASE nom. DEIXIS proximal.  
      105  GEND masc, NUM sg, PRON-TYPE demon]  
      107[CASE nom, GEND masc, NUM sg]
```

Case Marking

- Syntactic Case Marking is handled by using special Case Phrase KP (Bogel., et al, 2009)
- Accusative & Dative Case with “khE” marker
- Genitive case is special as it holds agreement
 - KPPoss is used

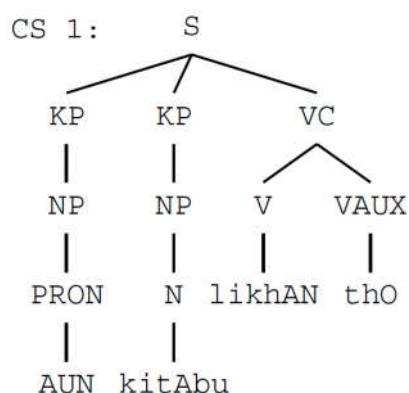


Verb Subcategorization

SUBJECT & OBJECT

ali kitAbu likHE tHO
Ali.NN book.NC write.Aoirst be.Aux.Pres
Ali Writes a book.
(↑ PRED)=’LIKHU<(↑ SUB) (↑ OBJ)>’

"AUN kitAbu likhAN tho"



PRED	'likhu<[1:AUN], [34:kitAbu]>'
SUBJ	[PRED 'AUN' NTYPE [NSYN pronoun] 1 CASE nom, GEND masc, NUM sg, PERS 1, PRON-TYPE personal]
OBJ	[PRED 'kitAbu' NSEM [N-CONCEPT inanimate] NTYPE [NSEM [COMMON count] NSYN common] 34 CASE acc, GEND masc, N-FORM obl, NUM sg]
TNS-ASP	[MOOD indicative, PERF -, PROG -, TENSE pres, TENSE-FORM aorist]
67	AUXTYPE tho, GEND masc, NUM sg, VTYPE main

Verb Subcategorization

SUBJECT Only

ali	dORE	tHO
Ali.NN	run.Aoirst	be.Aux.Pres
Ali runs.		
(↑PRED)=’dORi<(↑SUB)’		

Verb Subcategorization

Passives: SUBJ → NULL, OBJ → SUBJ

kitAbu likHijE thO

book.NC write.Pass.Aorist be.Aux.Pres

Book is being written/Book writing takes place.

(↑PRED)=’LIKHU<(NULL ↑SUB)>’

"kitAbu likhijE"

kitAbu likHibO AhE

book.NC write.Pass.Fut is.Aux.Pres

Book writing takes place.

(↑PRED)=’LIKHU<(NULL ↑SUB)>’

PRED	'likhu<NULL, [1:kitAbu]>'
	[PRED 'kitAbu'
	NSEM [N-CONCEPT inanimate]
SUBJ	[NSEM [COMMON count]]
	[NSYN common
	1[GEND masc, N-FORM obl, NUM sg]
35	[GEND masc, NUM sg, PASSIVE +, VTYP main]

Verb Subcategorization

Passives: NULL Arguments

likHibO

AhE

write.Pass.Fut.Sg.Masc is.Aux.Pres.Sg

Writing takes place.

(↑PRED)=’LIKHU<(NULL) >’

likHi;jE

tHO

write.Pass.Aorist.Sg be.Aux.Pres.Sg.Masc

(It’s) being written.

(↑PRED)=’LIKHU<(NULL) >’

Verb Subcategorization

Object-2 (OBJ-θ, Secondary OBJ)

ali CHOkirE=khe KHatu likhE
Ali boy.Obl=dat letter.Nom write

(↑PRED)=’likhu<(↑SUB) (↑ OBJ2) (↑ OBJ)>’

SUB: ali

OBJ2: CHOkirO

OBJ: KHatu

F-structure #1

"ali CHOkirE khE KHatu likhE"

```
PRED      'likhu<[1:Ali], [35:CHOkirO], [71:KHatu]>'  
          | PRED  'Ali'  
          | NSEM  [N-CONCEPT animate]  
          | NTYPES [NSEM [PROPER [PROPER-TYPE name]]]  
          | NSYN  proper  
          | 1[GEND masc, NUM sg, PERS 3]  
          |  
          | PRED  'CHOkirO'  
          | NSEM  [N-CONCEPT animate]  
          | NTYPES [NSEM [COMMON count]]  
          | NSYN  common  
          | 35[CASE dat, GEND masc, N-FORM obl,  
          |      NUM sg, PP-FORM khe]  
          |  
          | PRED  'KHatu'  
          | NSEM  [N-CONCEPT inanimate]  
          | NTYPES [NSEM [COMMON count]]  
          | NSYN  common  
          | 71[CASE nom, GEND masc, NUM sg]  
          |  
          | TNS-ASP [TENSE-FORM aorist]  
          | GEND masc, NUM sg, VTYPE main
```

Verb Subcategorization

Oblique

tUN CHOkirE=khE ali=khAN KHatu likhArAi
you boy=dat ali=abl letter write.caus2
(↑PRED)=’khAu<(↑ SUB) (↑ OBL) (↑ OBJ2) (↑ OBJ)>’

SUB: tUN

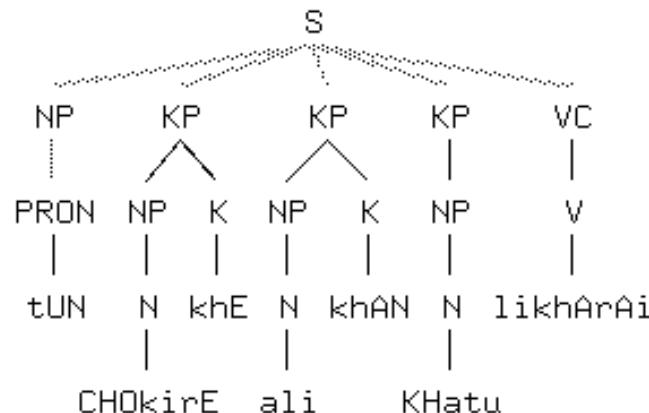
OBL: Ali

OBJ2: CHOkirO

OBJ: KHatu

Verb Subcategorization

"tUN CHOkirE khE ali khAN KHatu likhArAi"



F-structure #1

'tUN CHOkirE khE ali khAN KHatu likhArAi'

```

PRED 'likhu<[1:tUN],[35:CHOkir0],[71:Ali],[108:KHat]>'
  PRED 'tUN'
  SUBJ [NSYN pronoun]
  [CASE nom, GEND masc, NUM sg, PERS 2, PRON-TYPE personal]
  PRED 'CHOkir0'
  OBJ2 [NSEM [N-CONCEPT animate]]
  [NTYPE [NSEM [COMMON count]]
  [NSYN common]
  [CASE dat, GEND masc, N-FORM obl, NUM sg, PP-FORM khe]
  PRED 'Ali'
  OBJ [NSEM [N-CONCEPT animate]]
  [NTYPE [NSEM [PROPER [PROPER-TYPE name]]]
  [NSYN proper]
  [CASE agent, GEND masc, NUM sg, PERS 3, PP-FORM khan]
  PRED 'KHat'
  OBL [NSEM [N-CONCEPT inanimate]]
  [NTYPE [NSEM [COMMON count]]
  [NSYN common]
  [CASE nom, GEND masc, NUM sg]
  GEND masc, V-Formimperative, V-Form2 causative, VTYP main]
  
```

Verb Subcategorization

Complement (COMP)

Ali sOchyO [ta Ahmed kela khAE thO]

ali.Nom thought [that Ahmed bananas eat be.PresAux

(↑PRED)=’soch<(↑SUB) ↑COMP>’

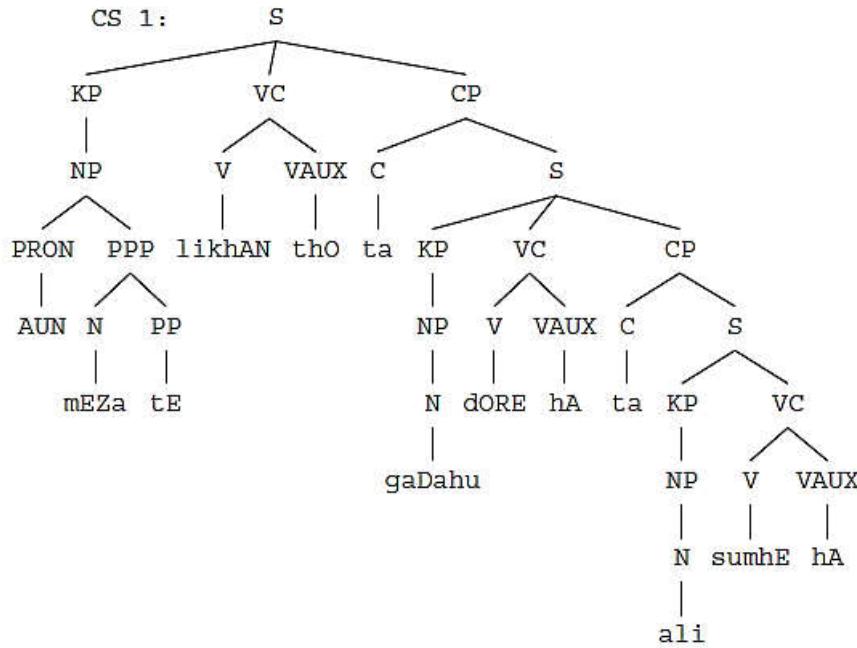
SUB: Ali

COMP: ‘khau<(↑SUB) ↑OBJ>’

SUB: Ahmed

OBJ: kela

Complement (COMP)



"AUN mEZa tE likhAN thO ta gaDahu dORE hA ta ali sumhE hA"

```

PRED      'likhu<[1:AUN], [119:dORi]>'
PRED      'AUN'
SUBJ      {
  ADJUNCT {
    {
      PRED      'mEZa'
      NSEM      [N-CONCEPT inanimate]
      NTTYPE     [NSEM [COMMON count]]
      67        [NSYN common]
      34        [CASE loc, PP-FORM on]
      1         [GEND fem, N-FORM obl, NUM sg]
      NTTYPE     [NSYN pronoun]
      1         [CASE nom, GEND fem, NUM sg, PERS 1, PRON-TYPE personal]
    }
  }
}
  121      [CASE acc, GEND masc, N-FORM obl, NUM sg, PERS 3]
PRED      'dORi<[121:gaDahu], [187:sumhu]>'
PRED      'gaDahu'
SUBJ      {
  NSEM      [N-CONCEPT animate]
  NTTYPE     [NSEM [COMMON count]]
  121      [NSYN common]
}
  189      [CASE acc, GEND masc, NUM sg, PERS 3]
PRED      'sumhu<[189:Ali]>'
PRED      'Ali'
SUBJ      {
  NSEM      [N-CONCEPT animate]
  NTTYPE     [NSEM [PROPER [PROPER-TYPE name]]]
  189      [NSYN proper]
}
  187      [CASE acc, GEND masc, NUM sg, PERS 3]
TNS-ASP   [MOOD contrafactual, TENSE-FORM aorist]
  187      [C-TYPE that, NUM sg, VTTYPE main]
TNS-ASP   [MOOD contrafactual, TENSE-FORM aorist]
  119      [C-TYPE that, NUM sg, VTTYPE main]
TNS-ASP   [MOOD indicative, PERF -, PROG -, TENSE pres, TENSE-FORM aorist]
  69       [AUXTYPE tho, GEND masc, NUM sg, VTTYPE main]
  
```

Verb Subcategorization

Open Complement (XCOMP)

Ali KHatu likhaNra **gHurE** thO

Ali letter write.inf want be.AuxPres

(↑PRED)=’gHuru<(↑ SUBJ) (↑ XCOMP)>’

SUB: Ali

XCOMP: ‘kara<(↑SUBJ) ↑OBJ>’

SUB: Ali

OBJ: KHatu

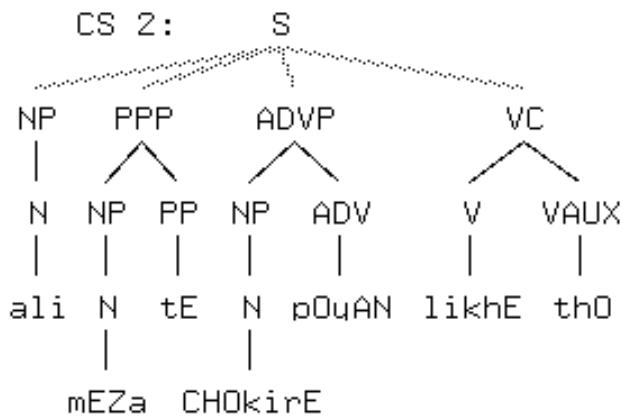
"ali KHatu likhaNra ghurE th0"	
PRED	'ghuru<[1:Ali], [69:likhu]>' PRED 'Ali' NSEM [N-CONCEPT animate]
SUBJ	NTYPE [NSEM [PROPER [PROPER-TYPE name]]] NSYN proper 1[GEND masc, N-FORM obl, NUM sg, PERS 3]
	[PRED 'likhu<[1:Ali], [35:KHAT]>' SUBJ [1:Ali] PRED 'KHAT' NSEM [N-CONCEPT inanimate]
XCOMP	OBJ [NSEM [COMMON count]] NSYN common 35[GEND masc, N-FORM obl, NUM sg] 69[GEND masc, NUM pl, VFORM inf, VTYP main]
	TNS-ASP [MOOD indicative, PERF -, PROG -, TENSE pres. AUXTYPE tho, GEND masc, NUM sg, VTYP main, TENSE-FORM aorist]

ADJUNCT

- Postpositional and adverbial phrases which do not fit in verb sub-categorization frames are called adjuncts
 - bHOLRO bAG mEN kHAE tHO
 - bHOLRO bAG mEN vaNra tE kHAE tHO
- Phrasal level Adjuncts
 - suTHO aiN suhiNrU CHOkirO

ADJUNCT

"ali mEZa tE CHOkirE pOyAN likhE th0"



	"ali mEZa tE CHOkirE pOyAN likhE th0"
PRED	'likhu<[1:Ali]>'
SUBJ	[PRED 'Ali' NSEM [N-CONCEPT animate] NTYPE [NSEM [PROPER [PROPER-TYPE name]]] NSYN proper 1[GEND masc, NUM sg, PERS 3]
OBJ	[PRED 'tE<[35:mEZa]>' PRED 'mEZa' NSEM [N-CONCEPT inanimate] NTYPE [NSEM [COMMON count]] NSYN common 35[GEND fem, N-FORM obl, NUM sg]
ADJUNCT	[CASE loc, PP-FORM on 69[PRED 'pOyAN<[72:CHOkir0]>' PRED 'CHOkir0' NSEM [N-CONCEPT animate] NTYPE [NSEM [COMMON count]] NSYN common 72[GEND masc, N-FORM obl, NUM sg] 105[ATYPE space]
	TNS-ASP [MOOD indicative, PERF -, PROG -, TENSE pres, TENSE-FORM aorist]
	112[AUXTYPE tho, GEND masc, NUM sg, VTYPE main]

XAJUNCT

- XADJUNCTs are embedded sentences where SUBJ is controlled from outside
- The only pattern found is marked by conjunctive participles
 - hU dORI gHaru vayO
 - Ali kitAbu likHI mAnI kHAdHI

hU dORI gHaru vayO

Ali kitAbu likHI mAnI kHAdHI

More Research is required on XADJUNCT Patterns in Sindhi

XAJUNCT

"CHOkir0 kitAbu likhI dORy0"

PRED	'likhu<[1:CHOkir0], [34:kitAbu]>'
SUBJ	[PRED 'CHOkir0' NSEM [N-CONCEPT animate] NTYPE [NSEM [COMMON count] NSYN common 1 [CASE nom, GEND masc, NUM sg, PERS 3]
OBJ	[PRED 'kitAbu' NSEM [N-CONCEPT inanimate] NTYPE [NSEM [COMMON count] NSYN common 34 [CASE nom, GEND masc, NUM sg]
XADJUNCT	[PRED 'dORi<[1:CHOkir0]>' SUBJ [1:CHOkir0] 91 [GEND masc, NUM sg, VTYPE main]
68	GEND masc, PTCPL-TYPE conjunctive, VTYPE main]

Pronominal Suffixes

Suffixes attached to verbs, construct different morphological forms, syntactically cause pro-drop



coverage

- Morphology

- FST Models (Nouns, Pronouns, Adjectives, Verbs)
- LFG Lexicon Postpositions, Conjunctions, Adverb
- Features
 - Gender, Number, Case, Mood, Aspect, Tense

- Syntax

- Partially Free Word Order
- SUB, OBJ, OBL, OBJ2, COM, XCOMP, ADJUNCT, XADJUNCT, PREDLINK
- Coordination, Subordination, Mood, Case, Aspect, Tense, Agreement

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

Conclusion & Future Work

- Development in current state covers the morphological and syntactic constructions discussed in above.
- Basic morphology and syntax constructs in Sindhi are identified and modeled.
- Morphological analysis shows interesting results like adjectives have more average inflections than nouns
- Pronouns have 3.58 average inflections per word.
- Also verb can have up to 75 different morphological forms (or even more)

Conclusion & Future Work

- 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,
 - Adverbial agreement
 - Prodrop phenomenon in Sindhi.

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 FiniteState Operations", in proc. *Workshop on Computational Approaches to Semitic 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 Algorithms 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.

Acknowledgements



مرکز حاسوبی ایجاد لسانی



CLT07, CLT09, CLT10, CLT12, CLT14, CLT16