Computational Linguistics
Issues and Challenges
Sindhi Pronominal Suffixes, Prodrop, and Particles

Mutee U Rahman
Isra University, Hyderabad
Outline

Part - I

• Pronominal Suffixes
• Pro-Drop
• Pro-drop patterns
• Syntactic Issues & Analysis

Part - II

• Particles

• Conclusion
Case

Ali=nE  Tafseer=kO  dEkhA
Ali =ERG Tafseer=Acc  see.Perf
Ali saw Tafseer

Ali  Tafseer dEkhA??
Pronouns

mEN dORUN
(may) I run

dORUN
Pronominal Suffixes

Pronominal suffix is a morphological variation in a word which reflects the pronoun along with existing word class
Pronominal Suffixes

• Sindhi is one of few Indo-Aryan languages (along-with Kashmiri, Punjabi and Saraiki) with pronominal suffixes.
Pronominal Suffixes

• In Sindhi Pronominal suffixes may appear on
  • Nouns, Postpositions
  • Verbs
    • Main
    • Auxiliaries
    • Modals
    • Copula verbs
  • Adverbs of place

Pro Drop

- Pro drop (Pronoun dropping) is a phenomenon in languages where pronouns are omitted and can be inferred grammatically or pragmatically.
  - [have you] ever been there? (English)
  - [tavhĩː] achoː thaːʔ (you are coming)
- Pronominal suffixes (may) cause pro drop in Sindhi

  **ach-ĩ**  **thoː**
  come-1PSg  be.1P.Sg
  (You.sg) are coming
Pronominal Suffixes with Nouns

• Pronominal suffixes which may appear on nouns, to mark genitive possessor pronoun

puṭu-\text{mi} \hspace{1cm} ac^\text{he:} \hspace{1cm} t^\text{ho:}

son.M.SG-1P.SG \hspace{1cm} come.IMPF.3P.SG \hspace{1cm} be.PRES.M.SG

My son comes.

• Subject Drop

• Any thing else dropped?
Postpositions

Follow Nouns, Pronouns etc..

yAsIn nE gadhE kO mArA

mEN darvAzE tak gayA

khE, mEN, tE, sAN, tAN, jO
Pronominal Suffixes with Postpositions

• With postpositions pronominal suffixes mark pronouns with different postpositional cases

  kʰeː-si  likʰ-u
  CM-SG3P.OBL  write.IMP-2P.SG
  (You) Write him.

• Subject & Object are dropped
• Something more is dropped??

کیس لکَ.
هن کی لکَ.
تون هن کی لکَ.
Pronominal Suffixes with Auxiliary Verbs

• On Auxiliary verbs these suffixes mark subjects (ultimately cause Subject Prodrop)

kita:bu likʰiyo: a:he:-mi
book.M.SG write.PERF.SG is.PRES.AUX-1P.SG
(I) have written the book.

ٚتَاب لكيو
آهي.
مون ٖتَاب لكيو
آهي.
Pronominal Suffixes with Modals

• With modals again subject prodrop is observed.

kitaːbu    likhiː:    sagʰ-ʔː    tʰoː:
book.M.SG  write.CONJPART  can-2P.SG  be.PRES.AUX.1P.SG
(you) can write a book.

تُونَ بِعَتَابٍ لَكِ
سُكرَتْهَنَ يَوِّ.
Pronominal Suffixes with Copula Verb

• a:he: is used as copula verb and is suffixed with first person noun suffix causing subject prodrop

kita:bu  a:he:-mi
book.M.SG  be.COP.SG-1P.SG
(I) have a book.

書いて、余部に書いた儿.
Mun も わ 再 挙書
Ahi.
Pronominal Suffixes with Main Verbs

• With main verbs these suffixes mark subject

kitaːbu likʰiyoː:-mi
book.M.SG write.PERF.SG-1P.SG
(I) wrote (a) book.

عتاب لكيوم.
عتاب لکیم.
مون عتاب لکیو.
mErE pAs tum hO
tum mErE pAs hO
tum pAs hO mErE
Pronominal Suffixes with Adverbs of Place

• Object pro-drop can be observed here

\[ \text{hu: } \text{vaṭa-si } \text{a:yo:} \]
He.M.Sg.2P  near-2P.SG
\[ \text{cameM.SG.PAST.2P } \text{(He) came to (him).} \]

\[ \text{هوُ وُتَس آيوُ.} \]
\[ \text{هوُ هُن وت آيوُ.} \]
Dual Pronominal Suffixes

• Dual pronominal suffixes may also appear on verbs and can possibly cause subject and object pro-drop at once

\[ \text{díṭ}^{\text{h}} \text{o}: -\text{mā}: -\text{va} \]
\[ \text{saw.SG}-1\text{P}.\text{SG.OBL}-2\text{P}.\text{PL.OBL} \]
\[ \text{(I) saw/looked (you / at you)}. \]

• The verb \( \text{díṭ}^{\text{h}} \text{o}: \) (saw) has \( 1\text{P}.\text{SG.OBL} \) suffix (-mā:) and \( 2\text{P}.\text{PL.OBL} \) suffix (-va) resulting a verb \( \text{díṭ}^{\text{h}} \text{o}: -\text{mā}: -\text{va} \) in which subject pronoun (I) and object pronoun (you.PL) are dropped but are inferable from suffixes
Pronominal Suffixes and Syntax

• Pronominal Suffixes cause Pro-Drop
• Subject is inferable
• Object is inferable
• However, along-with Subject the case marking information is also dropped

puṭu-si a:yọ:  huna-jo: puṭu a:yọ:
Son-3P.Sg   come.Perf.2P.Sg He.Obl-Gen son.Sg.M come.Perf.2P.Sg
His son came. His son came.
Pronominal Suffixes and Syntax

- One morphological form represents a syntactic phrase.
- Postpositional case markers are dropped and not reconstructable without going into the lexical semantics of verb.

<table>
<thead>
<tr>
<th>Case</th>
<th>Form</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>person</td>
<td>pronoun</td>
<td>construction</td>
</tr>
<tr>
<td>gen</td>
<td>putusā</td>
<td>son. Sg. M. 2: his son</td>
</tr>
<tr>
<td>oblique</td>
<td>huna khe:</td>
<td>beat him</td>
</tr>
<tr>
<td>loc</td>
<td>me: vijhìsi</td>
<td>insert (in) him</td>
</tr>
<tr>
<td>accusative</td>
<td>ma:risi</td>
<td>beat her, his son</td>
</tr>
<tr>
<td>instrumental</td>
<td>vijhìsi</td>
<td>insert (in) him</td>
</tr>
<tr>
<td>objective</td>
<td>vijhìsi</td>
<td>insert (in) him</td>
</tr>
</tbody>
</table>
Pronominal Suffixes and Syntax

• What are POS classes of following words?
• One word with multiple POS properties!
  • Pronoun+Postposition+Noun
  • Verb+Pronoun+Pronoun
  • Adverb+Pronoun
• One word with multiple grammatical functions!
• SUBJ, OBJ, VERB !!!

puțusi  
dițʰoːmãːva  
vată-si
Son.2P.Sg
Syntactic Analysis

• Lexical functional syntax analysis of verb “likhu” with 1P.SG.OBL suffix (likhiyami) causing SUBJ Pro-drop
Syntactic Analysis

• Analysis of verb ma:ru with dual pronominal suffixation causing SUBJ and OBJ Pro-drop
Syntax Analysis

• Analysis of noun putu (son) with pronominal suffix causing genitive pro drop
Syntax Analysis

• analysis of postposition khe: with pronominal suffix –si

• analysis of pronominal suffixed adverb poyā:
Conclusion

• Patterns found in Sindhi seem unique from its neighboring Indo Aryan languages Urdu/Hindi, Punjabi, and Kashmiri

• These patterns may appear with different grammatical classes

• with verbs they can also appear in pairs.

• Urdu/Hindi does not have pronominal suffixes, Kashmiri pronominal suffixes are part of its verbal morphology and Punjabi pronominal suffixes also appear with verbs.

• However; Saraiki has many pronominal suffix patterns identical to Sindhi.
Conclusion

• Analyses show that pronominal suffixation patterns cause different case formations, pro-drop, and postpositional drop which is sometimes not recoverable without going into the lexical semantics of a verb.

• Verbal pronominal suffixation analysis reveals that agreement, case and pro drop correlation do not necessarily hold for Sindhi like other south Asian languages (Butt and King, 1997).

• This may be considered areal feature of Sindhi along with Urdu and Punjabi.
Introduction: Particles

- Do not inflect

- do not easily fit into the established system of parts of speech

- Particles usually encode grammatical categories like negation, mood, tense, case, clitics, and fillers

(SIL, 2003), (Richards, Platt, and Weber 1985)
Sindhi Particles

- Sindhi particles "ee" (ئی), "ba" (بی), "na" (ئے), and "ta" (تا) are commonly used in everyday spoken and written language.

- Corpus analysis show that these particles are within top 30 most frequently occurring words in written language (Rahman, 2010).

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>جی</td>
<td>138904</td>
</tr>
<tr>
<td>2</td>
<td>تم</td>
<td>66930</td>
</tr>
<tr>
<td>3</td>
<td>آہی</td>
<td>61165</td>
</tr>
<tr>
<td>4</td>
<td>کی</td>
<td>60385</td>
</tr>
<tr>
<td>5</td>
<td>جو</td>
<td>56846</td>
</tr>
<tr>
<td>6</td>
<td>تی</td>
<td>39705</td>
</tr>
<tr>
<td>7</td>
<td>بی</td>
<td>31202</td>
</tr>
<tr>
<td>8</td>
<td>نم</td>
<td>27184</td>
</tr>
<tr>
<td>9</td>
<td>سکری</td>
<td>26021</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>27</td>
<td>جنھس</td>
<td>12573</td>
</tr>
<tr>
<td>28</td>
<td>گئی</td>
<td>12383</td>
</tr>
<tr>
<td>29</td>
<td>اہو</td>
<td>11372</td>
</tr>
<tr>
<td>30</td>
<td>ہو</td>
<td>11362</td>
</tr>
</tbody>
</table>
Sindhi Particles

- Part of speech status of these particles is debatable as they appear with different syntactic/grammatical functions at different places.

- For example, "na" (نا) can be used as a negation marker as well as reaffirmation "na ach" (نا اچ don't come) as in "acha na" (اچ نا come on).

<table>
<thead>
<tr>
<th>na</th>
<th>achu</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEG</td>
<td>come</td>
</tr>
<tr>
<td>Don’t come</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>achu</th>
<th>na</th>
</tr>
</thead>
<tbody>
<tr>
<td>come</td>
<td>Reaffirm / intensifier</td>
</tr>
<tr>
<td>Come on</td>
<td></td>
</tr>
</tbody>
</table>
Particle “ee” (ئﻴٰ)
Particle “ee” ےی

• “ee” ےی can follow any POS class except conjunction and interjection

• “ee” is used as emphatic particle (emphasizer)

• It emphasizes one thing and excludes the other one(s) which are either present in discourse or assumed/dropped

• Behavior of “ee” ےی in Sindhi is almost identical to Urdu/Hindi “hee” ہی (Koul, 2008)

<table>
<thead>
<tr>
<th>Noun ee</th>
<th>Pronoun ee</th>
<th>Adjective ee</th>
<th>Verb ee</th>
<th>Postpositon ee</th>
<th>Adverb ee</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHOkirO-ee</td>
<td>hUa ee</td>
<td>suhiNrO ee</td>
<td>chayO ee</td>
<td>jO ee</td>
<td>Ahiste ee</td>
</tr>
<tr>
<td>Boy-ee</td>
<td>She ee</td>
<td>Beautiful ee</td>
<td>Said ee</td>
<td>Genitive-ee</td>
<td>Slow-ee</td>
</tr>
</tbody>
</table>
Particle “ee” ﺃﻱ

- NP can have Noun, Pronoun, and Coordinated NP with any number of coordinated nominal elements
- ee is exclusive in nature it rules out other present / absent participants/events

**Ahmed=ee** kitAb PaRhyO asIN=ta veTHA huA-sIN

**Ahmed=emph** Book Read.past We=emph sit.PastPart Were.Aux-2pPl

Only Ahmed read the book, we were just sitting.
Ahmed was the only one who read the book, we were just sitting.

**Ahmed** kitAb=ee PaRhyO asIN=ta veTHA huA-sIN

**Ahmed** Book=emph Read.past We=emph sit.PastPart Were.Aux-2pPl

Ahmed read the book *(nothing else)*, we were just sitting.
Only Ahmed was in hurry, we were just sitting.

(It) was [Ahmed and Ali]’s book.
KWIC Snapshot of Corpus showing different patterns of “ee”
“ee” with Verbal Constructions

• ee as Verbal emphasizer

<table>
<thead>
<tr>
<th>mUN</th>
<th>kElO</th>
<th>khAdhO=ee</th>
<th>huO</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.Obl</td>
<td>Banana</td>
<td>Eat.past=ee</td>
<td>Was</td>
</tr>
</tbody>
</table>

I ate (did not throw away) banana.

• ee as Verbal emphasizer + event completion marker

<table>
<thead>
<tr>
<th>mUN</th>
<th>kElO</th>
<th>khAINdE=ee</th>
<th>pANrI</th>
<th>peetO</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.Obl</td>
<td>Banana</td>
<td>Eat.FutPart.Obl=ee</td>
<td>Water</td>
<td>drank</td>
</tr>
</tbody>
</table>

Having eaten banana I drank water (immediately)

• Ahmed AhE ee aBOJh (to be emphasizer)

Ahmed | AhE=ee | aBOJh |

Ahmed is (very)

simple=as soon as, an event or action which occurs immediately prior to ... Ref (Essentaila Urdu Grammar 904).

<table>
<thead>
<tr>
<th>Ahmed</th>
<th>Lahore</th>
<th>puhachaNdE=ee</th>
<th>fone</th>
<th>kayO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed</td>
<td>Lahore</td>
<td>Reach.FutPart.Obl=ee</td>
<td>Phone</td>
<td>did</td>
</tr>
</tbody>
</table>

Ahmed phoned / called just after /as soon as he reached Lahore.
Relative – Corelative Constructions with ee

- jeeN ee, teeN ee
- jeeaN ee verb teeaN ee verb
- Relative adverb as corelative adverb ee
“ee” as suffix in quantifiers / adjectives

- **sabH-ee (All-ee, All.emph)**

- **haR-ee (bundle-ee, All.emph)**

- **Ba-ee (two-ee)**

- **kE-ee (son)**
“ee” Movement

- CHOkirO ee School AyO huO (boy came to school (others may also))
- CHOkirO school ee AyO huO (boy came to school too)
- (CHOkirO school AyO ee huO (came to school (and left as well)))
- CHOkirO school (AyO aIN parRhO) ee* (dual emphasis on coordinated events/verbs)
- CHOkirO school AyO huO ee*?
- Achu ee na (please come on)
Particle “ba” (به)
Particle “ba” (به)

• Syntactic use pattern of “ba” is identical to “ee”
• However, “ee” is exclusive and “ba” is inclusive. Consider following examples

<table>
<thead>
<tr>
<th>Ahmed=ee</th>
<th>kitAb</th>
<th>paRhyO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed=ee.emph</td>
<td>Book</td>
<td>Read.Past</td>
</tr>
</tbody>
</table>

Only Ahmed read a book. (Others didn’t)

<table>
<thead>
<tr>
<th>Ahmed=ba</th>
<th>kitAb</th>
<th>paRhyO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed=ba.emph</td>
<td>Book</td>
<td>Read.Past</td>
</tr>
<tr>
<td>Ahmed(also)</td>
<td>read a book. (Others may also)</td>
<td></td>
</tr>
</tbody>
</table>
Particle “ba” (به)

• “ba” is ephasizer following NPs with same syntactic pattern as “ee”

Not a single rupee could be spent.
ba with Verb

- Following verb, it emphasizes the event in respective tense

Now I have to say too?

He fought with them too.
ba with Adverb

• Following adverb
  • pOe ba (even so), aNjA ba (even more),

That thing is far away from our imagination, even then why we are living this life?

• “taDeNh ba” “taba” nevertheless

...
More “ba” Patterns

- Dependent/subordinate clause

<table>
<thead>
<tr>
<th>Ahmed</th>
<th>kitAb</th>
<th>PaRhyO=ee</th>
<th>na</th>
<th>par</th>
<th>likhyO=ba</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed</td>
<td>Book</td>
<td>Read.past=eph</td>
<td>na.affirm</td>
<td>but</td>
<td>Wrote=emph</td>
</tr>
</tbody>
</table>

Ahmed not only read the book but wrote as well.

- Infinitive-tE ba (despite of, even then)
- hUNdE-ba (be.FutPart-ba despite of)
- kEr-ba (who-ba, whoever, anyone),
- Kujh-ba (some-ba, whatever)
- Kujh-ba na (some-ba-NEG, nothing)
ba Movement

- CHOkirO ba School AyO huO (boy came to school (others may also))
- CHOkirO school ba AyO huO (boy came to school too)
- (CHOkirO school AyO ba huO (came to school (and left as well))
- CHOkirO school (AyO aIN parRhO) ba (dual emphasis on coordinated events/verbs)
- CHOkirO school AyO ba? (Imperative)
- Achu ba (please come on)
- na ba, ta ba (even then)
Particle “ta” (تَهَ)
Particle “ta” (ته)

• “ta” is used as subordination conjunction ta=then here

(When) I read/studied the book then I understood the thing.
Particle “ta” (تہ)

• “ta” as an emphasizer
  • hU ta paRhE thO (contrast others may not)
    - hU=ta paRhE thO
    - He=ta.Emph read be.Pres
    - He reads. (Contrast others may not)

• hU English ta paRhE thO (contrast may not read other subjects)
  - English=ta.Emph read be.Pres
  - He reads/studies English. (Contrast may not read/study other subjects)
Subordination with ta

• “na ta” (otherwise) Links a conditional clause with consequence
  • vEha na ta mArENdO sANe (sit otherwise I will punish you)
Subordination with ta

- Time hujE ta AUN likhAN (ta marks resulting clause with conditional statement) If I had time I would write.
Subordination with ta, jadhen ta

• “jadhen ta” corelative phrases (mUN kEIO khAdhO, jadehn ta ahmed pArNI pItO (I ate a banana while ahmed drank water))

• “ta” also pose new subjects in the discourse
Other “ta” Patterns

• hUNa ta, chokiro sutHo AhE (with manner adverbs) by that way, by this way etc.
Other “ta” Patterns

• achO ta, achO ta sahi

  كا ضميرتاني شي با آهي وقت هن نمي تدوتم سمي هوکهن جا پگهاردار آهنت. بذایو
  جا پگهاردار آهن. بذایت سمي هوکهن نمي تدوتم سمي اهي پيرا کيدانم بيا وجن. باستو
  ذي پتي ثي تهجاري هن کي نو اهي ونمي تیدندو آهي تدمحمي سهتيه. (آزي) کي کن

• hU halyo vayo ta moon ali khe ghurayyo (ta=tadhen)

  كان بيجوته توالين چوھن مون. تا مكی ضروريت آمي، مختار هنوتم فياجن جبن هون. بند
  لينان کي چي؟ آهن. پویوی ان هنیار مون. تچیو اس مرامتابار مارکس کي مکی تور مویو
  پری شوکت ترين گي چی ناکه کر مون. تچیو شوکت ترين چي طکردار جین وکار گرکه کي. هن
  ناتیر ناهي، جهن کنکري پيدي تپمي وت مون. با بلیاپو کيکا آهن. رابطي سکنن کي جيف انجهنچر کوکری براج

• hin wat paisa aahin ta ba mani natho khae

  لکه دیو چنجي؟ جونس بحال رکی لودکی پدآ آهن. یوا آهن. جینکه دنوتن ابتریتون تي گھمش زراب
  وقت کي بھي مان وژون بووقوفی هکری پدآ آهن. پایه تار شنما دشمن یو سنتی وآپرنا کفی هکچا دشمن
  گرھیمون مختل پر آهن. اسلامی رنگ برگھئیمون پر دآ آهن. یچیو آهن. گچی رتویون یوسیون هندو کلجر پاکی جون
“ta” movement

- hU ta Pakistani AhE (*he* is Pakistan)
- hU Pakistani ta AhE (*he* is *Pakistan*)
- Ahe ta hU Pakistani (*he is* Pakistani)
- hU Pakistani AhE ta? (*he is Pakistani* (*so*?))
- mUN ta CHOkirE khE mAr dinI (*I* gave punishment to boy)
- mUN CHOkirE khE ta mAr dinI (*I* gave punishment *to boy*) (exclusive)
- mUN CHOkirE khE mAr ta dinI (*I* gave *punishment* to boy) (exclusive event)
- dinI ta mUN CHOkirE khE mAr ((*Though*) *I* punished the boy (but...))
- mUN CHOkirE khE mAr dinI ta? (*I* punished the boy *so*?)
Particle “na” as Negation Adverb

• Usually used as Negation / adverb
• tUN na achu (you) don’t come.

• Affirmative marker
• achu na (come on)
“na” movement

• Movement of “na” particle changes the role of na
• CHOkirO School na AyO (as negation adverb)
• CHOkirO (na school AyO), (na paRhyO) dual negation for every verb
• (CHOkirO na) school AyO (Emphasizer on noun)
• na CHOkirO school AyO na parRhO (neither, nor)
• CHOkirO school AyO na (Affirmative)
• Achu na (please come on)
• na ba, ta ba (even then)
Conclusion

- Sindhi particles "ee" (ئﻲ), "ba" (ﺑﮫ), "na" (ﻧﮫ), and "ta" (تھ) are used in a number of different grammatical functions.
- These particles neither inflect nor hold any agreement.

- Part of speech status (specially for POS tagging) of these particles is debatable as they appear with different syntactic functions at different places.
- POS tagging of these particles is an important design issue while designing POS tag set of Sindhi and particles like affirmative “na” and negation “na” must have separate POS tags.
References


• Depending on the context:
  • “ee” emphasizes different elements of a sentence exclusively (Just, Merely, Only)
  • “ba” emphasizes different elements of a sentence inclusively (Too, Also or Even)
  • “ta” emphasizes different elements of a sentence inclusively/exclusively (Although, But, However, Indeed)

• Sometimes ba and ta are used as mood markers