Dravidian Studies
(A Quarterly Research Journal)

Vol.XIII Nos.1-4, Vol.XIV Nos.1-2

UGC Approved No.40721

ISSN 0976-5182

# DRAVIDIAN STUDIES

( A QUARTERLY RESEARCH JOURNAL )

Vol.XIII Nos.1-4, Vol.XIV Nos.1-2 June 2018

June 2018

DRAVIDIAN UNIVERSITY
Srinivasavanam, KUPPAM-517 425

DRAVIDIAN UNIVERSITY

# DRAVIDIAN STUDIES

*(A Peer Reviewed Quarterly Research Journal)*

**Vol. XIII Nos. 1-4 & Vol. XIV No.1-2** **June -2018**

*Honorary Editor*
**Prof. E.Sathyanarayana**

*Guest Editor*
**Prof. B. Ramakrishna Reddy**

**DRAVIDIAN UNIVERSITY**
**Srinivasavanam, Kuppam-517 426**

**DRAVIDIAN STUDIES**
(A Peer Reviewed Quarterly Research Journal)

**ISSN 0976-5182 ( UGC Approved No.40721)**

Vol. XIII, Nos.1-4 & XIX No.1-2, June-2018






*Published by*
**Director**
Centre Publications
Dravidian University
Srinivasavanam
Kuppam – 517 426

*For Copies*
**The Registrar**
Dravidian University
Srinivasavanam, Kuppam - 517 426

*Printed at*
Vijayavani Printers
Chowdepally–517 257, Chittoor Dt.

# A Message

"Dravidian Studies" which reflects Dravidian languages, Literatures and Culture is reemerging to serve the cause of language lovers and researchers. The Journal brings out the ongoing research and recent developments in Dravidian languages/Literatures, duly authenticated by learned researchers and peers.

I appreciate the efforts of the scholars behind this academic initiative and desire that the Journal should win acclaim from all the concerned.

Kuppam
15.10.2018

**Prof. E.Sathyanarayana**
**Vice-Chancellor**

# Editor's Note

The present volume consists of research papers pertaining to the areas of linguistics, philosophy, education and language variation. Contemporary linguistic theory concentrates on machine translation, parsing etc. among others. Case marking is a dynamic phenomenon in Dravidian. Drawing data from Telugu and Tamil the acquisitive case marking processes, similarities and differences between the two languages are well portrayed. The criterion of animacy (of the object NP) plays a crucial role in case marking. A detail account of parsing with supporting examples from Tamil stands as the model for further research in the area. Thus, the first two articles provide an exploration in to the contemporary dynamics of computational linguistics.

Telugu unique in exhibiting a deep historical as well as a wide geographical spread. The latter factor is dealt in detail concentrating on gender differences in kinship terms across Telugu dialects from the perspective of sociolinguistics. Unlike other lexical items, kinship terms are used both for address and reference, wherein certain gender specific features are noticed in the Rayalaseema and Telangana dialects.

Language learning and language teaching are two of the crucial areas of Applied Linguistics. Out of the four skills of listening, comprehending, reading and writing., the first two are acquired (naturally) while the latter two are taught. The article on educational technology concentrates on the performance of Telugu students in Upper Primary Schools in achieving the skills of reading and writing. These achievement patterns are displayed through supporting statistical tables.

The musical literature of sankeerthanalu by the great devotional poet Annamayya (and his followers) are engraved on copper plates, most of which are deciphered and used extensively. Preservation and propagation of these songs their importance in the devotional literature, the storage of undeciphered copper plates – all these features form the focal songs of this article. The world wide well known Philosophy of Tiruvalluvar forms the final article of this volume, wherein the authors selects important teachings of the poet - philosopher one-by-one and explains each in detail for the comprehension of the modern readers.

# CONTENTS



---

**Note:** Views expressed in the articles are personal opinions of the contributors. Neither the Dravidian University nor the Editor is responsible for them.

# ACCUSATIVE CASE MARKING IN TELUGU AND TAMIL:TRANSLATION DIVERGENCE IN TEANSFER-BASED MACHINE TRANSLATION

**- Parameswari, K.**

*Abstract*

*A systematic translation mapping for syntactic structures of a source language to a target language is one of the main goals in any machine translation (MT) systems. Building a transfer-based MT system between Telugu and Tamil is also a challenging task as it involves considerable amount of translation divergence which requires suitable handling strategies and proper translation mapping of linguistic patterns. This paper deals with the differences that arise between Telugu and Tamil in the (overt/ covert) marking of the accusative case marker and its divergences between these languages. It also attempts to quantify divergences and proposes solution to handle them in transfer-based MT.*

Key words: Accusative Case Marking, Machine Translation, Divergences

## 1. Introduction:

Telugu and Tamil being cognate languages (which belong to Dravidian language family) show a number of differences in their marking of the accusative case with nominals. Traditionally, it is viewed that the accusative case marking is optional with inanimate nouns in Telugu and Tamil (cf. Krishnamurti and Gwynn, 1985 for Telugu, Lehmann, 1993 for Tamil). However, in this article, we demonstrate the instances where the object marking with inanimate nouns is realized obligatory in certain constructions in Telugu and Tamil; and the differences between their marking when they are translated. In machine translation, mapping case markers with their suitable functions between languages are significant to obtain better results. In this paper, the divergence between the accusative case marker in Telugu and Tamil is studied in order to handle them in transfer-based MT.

## 2. Accusative Case Marking in Telugu and Tamil

The accusative case is marked by the suffix *-ai* in Tamil and *-ni/-nu* in Telugu. The direct object of a transitive verb is the accusative case marked either overtly or covertly with certain nouns in Telugu and Tamil. In Telugu, nouns denoting animate objects must take the accusative case suffix, whereas with inanimate nouns, its use is optional (cf. Krishnamurti and Gwynn, 1985:88). In Tamil, the accusative case suffix is used obligatorily with nouns denoting rational objects and with definite marked non-rational objects. It is used optionally with indefinite non-rational nouns (cf. Lehmann, 1993:29). In both the languages, the accusative case marker is syncretic to the nominative case marker(ø) when the direct object marker is not overtly case marked.

Major divergences in marking of the accusative case marker are observed in the following factors between Telugu and Tamil:

i. Differential Object Marking
ii. Causative Construction
iii. Verbs of Inherently Directed Motion
iv. Dative Subject Construction
v. Accusative case Marker as Complements of Postpositions

This paper explains the occurrence of accusative case marker in the above domains and its divergence between Telugu and Tamil.

### 2.1 Differential Object Marking

It is more common to find in Telugu and Tamil that the object in certain cases is not accusative case marked. The variation in the occurrence of case marking with direct object is termed as Differential Object Marking (DOM) (Bossong, 1991; Aissen, 2003; Subbarao, 2012 amongst others). The occurrence of accusative case marker on objects depends on prominence scales such as animacy and definiteness. The higher one in the scale is more prominent and tends to be marked by the accusative case than the lower one. This can be shown in the following scalar dimensions.

**Animacy scale:**

Human > Animate > Inanimate

It is rewritten as,

[+human] > [-human, +animate] > [-animate]

**Definiteness scale:**

Personal pronoun > Proper noun > Definite NP > Indefinite specific NP > Non-specific NP

(adopted from Aissen, 2003)

An overt marking of the accusative case on nominals is determined by the co-occurrence of features of definiteness scale with the features of animacy scale. The divergent behaviour of DOM in Telugu and Tamil is figured out by applying animacy and definiteness scale on objects. The ideas can perhaps be better illustrated with the following examples.

**A. Nouns of [+human] with Definiteness Scale**

Both in Telugu and Tamil, a noun[+human] on the definiteness scale of Aissen (2003) are obligatorily marked for the accusative case. Categories in the definiteness scale (i.e. personal pronoun, proper noun, definite NP, indefinite specific NP and non-specific NP) are overtly marked for the object marker when the animacy feature is

[+human]. Example for definite NP with the accusative case marked is shown in the example (1).

(1) Te. *nēnu* *nā* *snēhitui-* *ni* *cūs-* *ā-* *nu.*
I.NOM my friend- ACC see- PST- 1.SG.
Ta. *nāI* *eI* *naGpaI-* *ai.p* *pār-* *tt-* *ēI.*
I.NOM my friend- ACC see- PST- 1.SG.
'I saw my friend (male)'

However, in certain cases, the accusative case marker with the object noun[+human] is not marked in Tamil as in (2).

(2) Ta. *kumār* *paiyaI* *tēmu-* *ki_-* *āI.*
Kumar boy search- PRS- 3.SG.M.
Te. *kumāru* *abbāyi-* *ni* *vetuku-* *tunn-* *āu.*
Kumar boy- ACC search- PROG- 3.SG.M.
'Kumar is looking for a son-in-law.'

The absence of accusative case marking with nouns[+human] in the above example (2) in Tamil indicates that the object noun is incorporated into the verb. Here the predicate *paiyan* 'look for a son-in-law' exhibits an idiomatic sense due to the incorporation of the noun into the verb and it forms a complex predicate. Similarly, other predicates like *peI tēmu /pār* 'look for a bride', *ā7 tēmu* 'look for a person', *peI komu* 'give a girl (to marry)' exhibit the same idiomatic sense. In such cases, Tamil disallows the accusative marker since the object is a part of a complex predicate.

However, when the nouns mentioned above explicitly appear with the accusative, it can be interpreted that they are direct objects of transitive verbs *tēmu* 'search'/ *pār* 'look'/ *komu* 'give'. In this context, the object does not get incorporated with the verb in Tamil, hence it is treated as a distinct object and is accusative case-marked.

(3) Te. *kumāru* *abbāyi-* *ni* *vetuku-* *tunn-* *āu.*
Kumar son-in-law- ACC search- PROG- 3.SG.M.
Ta. *kumār* *paiyaI-* *ai.t* *tēmu-* *ki_-* *āI.*
Kumar son-in-law- ACC search- PRS- 3.SG.M.
'Kumar is searching the boy.'

**B. Nouns of [-human, +animate] with Definiteness Scale**

Similar to nouns[+human], nouns [-human, +animate] on the definiteness scale are also obligatorily marked for the accusative case both in Telugu and Tamil. Example:

(4) Te. *nēnu gurrān- ni cūs- ā- nu.*
I horse- ACC see- PST- 1.SG.
Ta. *nāI kutirai.y- ai.p pār- tt- ēI.*
I horse- ACC see- PST- 1.SG.
'I saw a horse.'

**C. Nouns of [-animate] with Definiteness Scale**

Though there are a few similarities in the marking of accusative case on nouns that are [-animate], it is also found that there are some dissimilarities in such marking between Telugu and Tamil. The dissimilarity occurs when the object[-animate] is a personal pronoun or modified by definite modifiers.

**(i) Personal pronoun[-animate]**

Pronouns *idi* 'it.PROX' and *adi* 'it.DIST' can be used as referential to nouns[-animate] (eg. computer) as well as nouns[-human,+animate] (eg. dog) in Telugu and correspondingly *itu* and *atu* in Tamil. When these pronouns are used as co-referential to nouns [-human,+animate] in its object position, the accusative is marked obligatorily in Telugu and Tamil.

Whereas, when the object co-referencing with [-animate] objects in singular number (see example(5)) and in plural number (see example (6)), they are optionally marked with the accusative case marker in Telugu and obligatorily marked in Tamil.

(5) Te. *nēnu idi cūs- ā- nu.*
I it.PROX see- PST- 1.SG.
Ta. *nāI it- ai.p pār- tt- ēI.*
I it.PROX- ACC see- PST- 1.SG.
'I saw it.'

(6) Te. *nēnu avi cūs- ā- nu.*
I they.N.DIST see- PST- 1.SG.
Ta. *nāI ava__- ai.p pār- tt- ēI.*
I they.N.DIST- ACC see- PST- 1.SG.
'I saw them (N.)'

The pronouns *idi* 'it.PROX' and *adi* 'it.DIST' may also refer to nouns that are [+human, +female] in Telugu, besides the pronoun *īme*'she. PROX' and *āme* 'she.DIST'. In such cases, either intimacy or derogatoriness of the subject towards the direct object is indicated. Here, the accusative case suffix is used obligatorily as in (7), since the pronoun [+human] has a higher status in the animacy scale.

(7) Te. *nēnu dīni- ni cūs- ā- nu.*
I [she/it].PROX- ACC see- PST- 1.SG.
'I saw her/it'

Whereas in Tamil, the pronouns *iva7* 'she.PROX' and *ava7* 'she.DIST' are only used to refer to [+human, +female] features and thus carry the accusative marker in their object position.

**(ii) Proper noun [-animate]**

Both in Telugu and Tamil, proper nouns[-animate] in their object position are marked for the accusative case as in (9). Similarly, nouns[-animate] which indicate higher cultural items are obligatorily marked by the accusative case in Telugu (cf. Ramarao, 1975) as well as in Tamil as in (10).

(8) Ta. *nāI intiyā.v- ai nēci- kki_- ēI.*
I India- ACC love- PRS- 1.SG.
Te. *nēnu Bāradadēsān- ni prēmis- tunn- ānu.*
I India- ACC love- PRS- 1.SG.
'I love India'

(9) Ta. *kutupcā kōlkoMmā.v- aik kamm- iI- āI.*
Qutub Shah Golkonda- ACC build- PST- 3.SG.M.
Te. *kumub cā gōlkoMa- ni kāmm- ā- u*
Qutub Shah Golkonda- ACC build- PST- 3.SG.M.
'Qutub Shah built the Golkonda.'

**(iii) Definite NP [-animate]**

In Tamil, when a direct object[-animate] is modified by definite modifiers, the presence of accusative case is obligatorily realized in contrast to Telugu where it is optional.

Possessives and determiners in a noun phrase indicate the definiteness or specificity of the referent. As mentioned above, Tamil

uses the accusative case compulsorily with the object modified by possessives (see example(11)) and demonstratives (see example(12)).

(10) Te. *vāu nā cemmu- (ni) cūs- ā- u.*
he my tree- (ACC) see- PST- 3.SG.M.
Ta. *avaI eI maratt- ai.p pār- tt- āI.*
he my tree- ACC see- PST- 3.SG.M.
'He saw my tree.'

(11) Te. *vāu ā cemmu- (ni) cūs- ā- u.*
he that tree- (ACC) see- PST- 3.SG.M.
Ta. *avaI anta maratt- ai.p pār- tt- āI.*
he that tree- ACC see- PST- 3.SG.M.
'He saw that tree.'

**(iv) Indefinite specific NP [-animate]**

The accusative case marker is also referred to as a specificity marker, since it denotes specificity When the predicate is [+transitive] in the nominative subject construction, the object is accusative case-marked when it denotes specificity/animacy (Subbarao, 2012:8). Example:

(12) Te. *nēnu oka pedda māmiicemmu- nicūs-ā- nu.*
I that big mango.tree-ACC see- PST- 1.SG.
Ta. *nāI oru periya māmaratt- ai.ppār- tt- ēI.*
I a big mango.tree- ACC see- PST- 1.SG.
'I saw a big mango tree.'

**(v) Non-specific NP [-animate]**

With inanimate non-specific object nouns, the accusative case is not marked or optionally marked in both the languages. Example:

(13) Te. *nēnu pāma pām- ā- nu.*
I song sing- PST- 1.SG.
Ta. *nāI pāmmu pām- iI- ēI.*
I song sing- PST- 1.SG.
'I sang a song.'

However, when the subject's animacy feature is also inanimate (S[-animate]), the object is obligatorily marked by the accusative to

avoid the ambiguity of the role of arguments in the sentence as in (13).

(14) Te. *piugu [*baMa/ baMa- ni] cīlci- M- di.*
lightening *rock/ rock- ACC split- PST- 3.SG.N.
Ta. *imi [*pā_ai/pā_ai.y- ai.p] pi7a- nt- atu.*
lightening *rock/ rock- ACC split- PST- 3.SG.N.
'The lightening breaks the rock'.

In certain cases, the accusative case is marked with an inanimate non-specific object (O[-animate,-specific]), though the subject is [+animate]. Especially, verbs which show cause-motion effects or verbs of impingement mark their object for the accusative both in Telugu and Tamil. The impingement verbs *kommu* 'to beat', *gillu* 'to pinch', *campu* 'to kill', *tannu* 'to kick' and etc., in Telugu and *ami* 'to beat', *ki77u* 'to pinch', *kol* 'to kill', *utai* 'to kick' *tammu* 'knock' and etc., in Tamil. Example:

(15) Te. *rāmuu kurcī- ni tann- ā- u*
Ram.NOM chair- ACC kick- PST- 3.SG.M.
Ta. *rāmaI nā_kāli.y- ai.t ta77- iI- āI.*
Ram.NOM chair- ACC kick- PST- 3.SG.M.
'Ram kicked the chair'

Besides, in the example (16) it is shown that the object in Telugu appears in nominative despite the fact that the predicate contains an impingement verb. The predicate in Telugu, i.e *talupu kommu* compositionally provides the meaning as 'knock at the door'. Here, it forms a noun+verb compound where the accusative is not allowed. If it is not treated as a compositional predicate, it means 'hit the door'. When the accusative is explicit as in (17), it means 'beat the door' in Telugu. Whereas in Tamil (18), the verb *tammu* 'to knock' has an exclusive interpretation as 'knock at' with [-animate] direct object, it obligatorily marks the direct object (DO) with the accusative case marker.

(16) Te. *rāmuu talupu komm- ā- u.*
Ram.NOM door.NOM beat- PST- 3.SG.M
Ta. *rāmaI katav- ai.t tamm- iI- āI.*
Ram.NOM door- ACC knock-PST- 3.SG.M
'Ram knocked at the door'

(17) Te. *rāmuu talupu- ni komm- ā- u.*
Ram.NOM door- ACC beat- PST- 3.SG.M.
'Ram beat the door'

The major divergence in the usage of accusative case marker between Telugu and Tamil occurs with the object which is realized in,

(i) the definiteness scale of [+pronoun] with the animacy scale of [-animate].

(ii) the definiteness scale of [+definite] with the animacy scale of [-animate].

In both the cases, Telugu optionally marks the DO with the accusative whereas in Tamil it is realized obligatorily.

It is also noted that the accusative case marker is obligatorily used with certain non-specific NP[-animate] when occurring with the S[-animate] and with the V[+impingement], both in Telugu and Tamil.

### 2.2 Causative Construction

A causative verb in predicate requires three arguments in the form of noun phrases, viz., causer agent, actor agent and object. The second agent is considered as the real actor, whereas the first agent causes the second agent to act. Here the second agent phrase is marked for the postposition *cēta* 'by means of' in Telugu, the accusative case marker is present in Tamil.

(18) Te. *nēnu vāi- cēta iMmi panu- lu cēy- iMc- ā-nu.*
I.NOM he- by house.OBL task- PL..ACC do- CAUS- PST- 1.SG.

Ta. *nāI avaI- ai vīmmu vēlai.(y-ai) ceyy- a- vai- tt- ēI.*
I.NOM he- ACC house.OBL task- ACC do- INF- CAUS- PST- 1.SG.

'I made him to do the household tasks.'

An object-complement double accusative is a construction in which one accusative is the direct object of the verb and the other accusative (either noun, adjective or participle) complements the object in that it predicates something about it (Wallace, 1985:93). In Telugu, the accusative case marker is assigned to the object irrespective of

the animacy and definiteness features. The grammatical case of object is also percolated/ copied to its complement in Telugu as in (21). Whereas in Tamil, the object complement do not take the accusative and thus realized in its nominative form.

(19) Te. *nēnu vāi- ni [rāyi- ni/ manici- ni] cēs- ā- nu.*
I he- ACC [stone- ACC/ human- ACC do- PST- 1.SG.
Ta. *nāI avaI- ai [kal/ maIitaI] ākk- iI- ēI*
I he- ACC stone.NOM/ human. NOM make- PST-1.SG.
'I made him into a stone/a human.'

In Telugu, the verb *ceyyi* 'do' functions here as a periphrastic causative verb and allows two accusative case marked noun phrases in a row. It is illustrated as below:

Te. NP(causer) NP(causee)-ACC NP(theme)-**ACC** VG <root="*ceyyi*">

Ta. NP(causer) NP(causee)-ACC NP(theme)-**NOM** VG<root=" *ākku*">

However, the swapping/scrambling of objects is not permitted, since only the preverbal object (i.e. theme) is interpreted as object-complement to the object (causee) of the verb *ceyyi* 'to do'.

(20) Te. **nēnu [rāyi- ni/ maniRi- ni] vāi-ni cēs- ā- nu*
I [stone- ACC/ human- ACC] he- ACC do- PST-1.SG.
Ta. **nāI [kal/ maIitaI] avaI- ai ākk- iI- ēI*
I [stone.NOM/ human.NOM he- ACC make- PST-1.SG.
'I made him into a stone/ a human.'

However, when the object and its complement share similar semantic features, the complement is not usually marked by the accusative in Telugu. In the example (23), the object *kukka* 'dog' is [-human, +animate] and the object complement *gurraM* 'horse' is also [-human, +animate], this situation precludes marking the object complement with the accusative case marker. Hence, causative constructions require the object and the object complement in Telugu belong to two different semantic categories which then enables a verb to cause one noun into another noun.

(21) Te. *nēnu kukka- ni gurrā- ni cēs- ā- nu.*
I dog- ACC horse- ACC/ do- PST- 1.SG.

Ta. *nāI nāy- ai kutirai ākk- iI- ēI*
I dog- ACC horse- make- PST- 1.SG.
'I made a dog into a horse'.

However, alternatively, the object complement nouns are marked by the adverbializer *-gā* when the predicate verb is *mārcu* 'change' in Telugu. Similar construction is also noticed in Tamil.

(22) Te. *nēnu kukka- ni gurraM- gā mārc- ā- nu.*
I dog- ACC horse- ADV change- PST- 1.SG.
Ta. *nāI nāy- ai kutirai.y- āka mā__- iI- ēI*
I dog- ACC horse- ADV change- PST- 1.SG.
'I made a dog into a horse.'

### 2.3 With Verbs of Inherently Directed Motion

Certain inherently directed motion verbs such as 'climb up', 'climb down' and 'cross' require their complements to be marked for the accusative in Telugu. Whereas in Tamil, the first two verbs mark their predicate nouns with the locative and the ablative marker respectively and the last verb marks the noun with the accusative case marker in Tamil.

(23) Te. *nēnu ēnugu- nu ekk- ā- nu.*
I.NOM elephant- ACC climb up- PST- 1.SG.
Ta. *nāI yāIai- mēl ē_- iI- ēI.*
I.NOM elephant.OBL- on climb up- PST- 1.SG.
'I climbed up the elephant.'

(24) Te. *nēnu ēnugu- nu dig- ā- nu.*
I.NOM elephant- ACC climb down- PST- 1.SG.
Ta. *nāI yāIai- mēl- iruntu i_aEk- iI- ēI.*
I.NOM elephant. OBL- on ABL climb down- PST- 1.SG.
'I climbed down the elephant.'

(25) Te. *nēnu nadi- ni dām- ā- nu.*
I river- ACC cross- PST- 1.SG.
Ta. *nāI ā__- ai.k kama- nt- ēI.*
I river- ACC cross- PST- 1.SG.
'I crossed the river.'

### 2.4 Dative Subject Construction

The dative subject construction is the most widespread in Dravidian (Subbarao, 2012:134). The dative marked subject acts as an experiencer subject and the verb agrees with the object. In Tamil,

stative predicates expressing the notion of mental, emotional and physical experience require the case marking pattern of DAT-ACC. (Lehmann, 1993:180).

**(i) With verbs of mental experience:**

Verbs such as *teri* 'know' and *puri* 'understand' etc., in Tamil and *telusu* 'know' and *arThamavvu* 'understand' etc., in Telugu refer to the mental experience. When they occur in predicate position, the theme is marked for the accusative in Tamil and the nominative in Telugu as in the following example. Though the theme is in the nominative in Telugu, the verb *telusu* 'know' gets default agreement since it is a defective verb.

(26) Ta. *eI- akku avaI- ai.t teriyum.*
I- DAT he- ACC know
Te. *nā- ku vāu telusu.*
I- DAT he.NOM know
'I know him.'

**(ii) With verbs of emotional experience:**

Verbs like *piti* 'like' etc., in Tamil and *naccu* 'like' etc., in Telugu express emotional experience. The theme is accusative case marked in Tamil, whereas it is in the nominative and agrees with the verb in Telugu.

(27) Ta. *eI- akku uII- ai.p pimi- kk- um*
I- DAT you- ACC like- FUT- 3.SG.N.(default)
Te. *nā- ku nuvvu nacc- ā- vu$_i$.*
I- DAT you. NOM like-PST- 2.SG.
'I like you.'

**(iii) With verbs of physical and biological experiences:**

Verbs such as *paci* 'be hungry', *vali* 'be painful', *ari* 'be itching' etc., in Tamil *and ākali veyyi/ ākali pummu/ ākali pemmu* 'be hungry', *noppi veyyi/ noppi pummu / noppi pemmu* 'be painful', *duradagā uMu*'be itching' etc., in Telugu convey psycho-semantic, physical and physiological experiences. These predicates may occur with DAT-ACC pattern in Tamil.

(28) Ta. *eI- akku talai.(y- ai) vali.k- ki_- atu.*
I- DAT head- (ACC) pain- PRS- 3.SG.N.(default)
Te. *nā- ku talanoppi$_i$ vēs- tuM- di$_i$*
I- DAT head.NOM put- PRS- 3.SG.N.
'I feel headache.'

**2.5 Accusative Case Marker as Complements of Postpositions**

Historically, most of the postpositions in Telugu and Tamil are grammaticalized forms of verbs and nouns. They lost their autonomous status as words and are frozen as postpositions (Cf. Krishnamurti, 2003:240). Postpositions derived from grammaticalized forms of transitive verbs in Tamil and Telugu govern their complements inflected for the accusative case marker. The accusative case marker here is an argument of the postposition but not of the verb. Postpositions such as *guriMci* 'about', *bammi* 'depend on/according to' and *pōlina* 'like' in Telugu and *ku_ittu/ pa__i* 'about', *po_uttu* 'depend on/ according to', *poI_a* 'like' in Tamil require the accusative case marked nouns. Whereas, postpositions such as *mātiri* 'like', *vima* 'than', *nōkki* 'towards' and *koGmu* 'with' in Tamil require the complement noun with the case marker *–ai* and the corresponding postpositions do not require their nouns to be marked with the accusative case marker in Telugu.

In Tamil, the postposition *tavira* 'except' characteristically occurs either with the accusative case or with the case which is copied from the following noun phrase. The postposition *tappa* 'except' in Telugu occurs with the complement marked for a specific case that is a copy of the following noun phrase. In the following example (29), the occurrence of postposition *tavira* in Tamil and *tappa* in Telugu with the following noun in its dative case marking is shown.

(29) Te. *nēnu$_i$ vāi- ki tappa evari-kī ivv-a- ledu*
I.NOM he- DAT except who-DAT.NPI give-INFNEG
Ta. *nāI [avaI-ai/avaI-ukku] tavira yāru-kk-um komu.kk-a.v- illai*
I.NOM he- ACC he-DAT except Who-DAT-NPI give-INFNEG
'I didn't give it to anyone except him.'

## 3. Quantification of Divergence

The divergence in the usage of the accusative case maker is quantified based on their different domains of occurrence. The following situations are counted for divergence.

Let languages be A and B.

1. If the language A and B uses the accusative marker invariably, then no divergence
2. If any one of the languages i.e. A or B differ in their usage, then counted as divergence

If any one of the languages i.e. A or B 'optionally' realized with the accusative case marker, then counted as divergence

| **Domains of Occurrence** | **No. of divergences** | **Examples** |
|---|---|---|
| Differential Object Marking | **07** | (2),(5),(6),(7),(10), (11) & (16) |
| Causative Construction | **02** | (18) & (19) |
| With Verbs of Inherently Directed Motion | **02** | (23) & (24) |
| Dative Subject Construction | **03** | (26), (27) & (28) |
| Complements of Postpositions | **05** | *mātiri* 'like', *viṭa* 'than', *nōkki* 'towards' and *koṇṭu* 'with' in Tamil and example (29) |
| **Total** | **19** | |

**Table 3.1 Quantification of Divergence between Telugu and Tamil**

As shown in the table (3.1), there are nineteen instances where Telugu and Tamil show divergences in their usage of accusative case marking which need special attention in building MT.

## 4. Handling Strategies in Machine Translation:

A number of NLP modules are used in handling divergences in transfer-based MT. To handle the case marking, modules such as dependency parser, transfer grammars (TG) and morphological generators (MG) are used (Parameswari, 2014). The dependency parser provides the relationship between nouns and a verb in a sentence. It disambiguates the functions of case marker by providing relationship tags. TG is equipped with performing certain tasks such as insertion, deletion, modification and re-ordering of words and

chunks. It also has the ability to handle files where it is possible to operate a single rule over a list of items. MG generates well formed wordforms of target language based on the grammatical feature.

## 5. Conclusion:

The marking of accusative case in Telugu and Tamil shows considerable amount of divergence and hence, the straightforward mapping between these languages is not possible in MT. NLP modules such as parsers, TG and MG are useful in handling these divergences. Quantification of divergence facilitates MT in proposing where to put efforts for the given language pair to attain a better result. A systematic study of divergences and proper handling of them are indispensable to get comprehensive output in MT.

## Acknowledgement:

I would like to acknowledge members of Indian Language to Indian Language (IL-IL) MT project and Prof. G. Uma Maheshwar Rao, the chief investigator of Telugu-Tamil and Telugu-Tamil bidirectional MT systems for giving me an opportunity in building Telugu-Tamil MT systems.

## References:

Aissen, Judith. (2003). 'Differential Object Marking: Iconicity vs. Economy'. *Natural Language & Linguistic Theory* 21(3). 435-483.

Bossong, Georg. 1991. Differential object marking in Romance and beyond. In *New Analyses in Romance Linguistics, Selected Papers from the XVIII Linguistic Symposium on Romance Languages 1988*, eds. D. Wanner and D. Kibbee, 143–170. Amsterdam: Benjamins

Krishnamurti, Bh. & J. P. L. Gwynn. (1985). *A Grammar of Modern Telugu.* Delhi: Oxford University Press.

Krishnamurti, Bh. (2003). *The Dravidian Languages*. Cambridge: Cambridge University Press.

Lehmann, Thomas. (1993). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

Parameswari, K. (2014). *Development of Telugu-Tamil Machine Translation: With Special Reference to Divergen*ce. Unpublished Ph.D. Thesis. Hyderabad: University of Hyderabad.

Ramarao, Chekuri. (1975). *Telugu vâkyam* [‘The Telugu Sentence’- in Telugu]. Hyderabad: AP Sahitya Academy.

Subbarao, K. V. (2012*). South Asian Languages: A Syntactic Typology*. Cambridge:Cambridge University Press.

Wallace, Daniel B. (1985). ‘The Semantics and Exegetical Significance of the Object-Complement Construction in the New Testament’. *Grace Theological Journal* 6(1). 113-160.

# PARSING IN INDIAN LANGUAGES WITH SPECIAL REFERENCE TO TAMIL: THE STATE-OF-THE-ART

**- Keerthana B. and Parameswari K.**

***Abstract***

*Parsing is the task of assigning syntactic/syntactico-semanticroles for sentences by segmenting sentences into relevant processing units. The tool used for the automation of such process is a parser, a major module in building Natural Language Processing (NLP) applications like Machine Translation systems. The paper aims to provide the current scenario of parsing in Indian languages in general and Tamil, in particular, focusing on the grammar formalisms and annotation schema available. It alsoattempts to study existingresearch works to gain an overall understanding of the current parsing techniques and the state-of-the-art of research in Indian NLP in the global scenario.*

## 1.Introduction

The parser is an automated NLPtool used for syntactic/ syntactico-semantic analysis of sentences. It processes the input sentences based on the grammar formalism followed in implementation and produces output as constructed parse trees. Building a parser is a challenging task as it involves in handling

structural ambiguities in languages. Structural ambiguities are realized due to two different factors; (i) attachment ambiguity and (ii) coordination ambiguity. They are illustrated in examples (1) and (2) respectively:

(1) I saw a girl with the telescope.

(2) Old men and women

In the example (1), the attachment of noun phrase (NP) (*the telescope)*shows an ambiguity as it can be potentially attached with the subject NP (*I)*or with the object NP (*the girl)*.In the example (2), the coordination ambiguity is shown where the adjective *old* may have interpreted to be coordinated with either *men* or *women.*

The parser attempts to resolve such structuralambiguities based on various factors such as morphological, syntactic, semantic, contextual and discourse knowledge of a language. Once ambiguities are identified, the parser attempts to choose rightly parsed output with the given knowledge.Hence, building a parser with an effective algorithm is desirable for an efficient disambiguation process.

The focus of this paper is to discuss the state-of-the-art of parsers in Indian languages with a special reference to Tamil, a Dravidian Language. It includes discussions ongrammar formalisms, annotation schema andtechniques followed in implementing parsers for the above-said languagesfocusing research taken place since 2009 till today.The paper also discusses the suitable tagset for Tamil by considering specific language features and grammar formalisms.

The paper is divided into six sections discussing the following topics:

(i) Grammar formalisms in Parsing
(ii) Parsing technique
(iii) Types of parser
(iv) Annotation schema
(v) Parsers: a review
(vi) A discussion on a suitable model for Tamil

## 2. Grammar Formalisms in Parsing

Linguistic understanding and selecting suitable grammar formalism are considered to be the base for building an effective parser for any language. A number of grammar formalisms areavailable for analysing languages and some of them are also adaptedin building computational parsers. This paper discussed formalisms such as (i) Generalized Phrase Structure Grammar (ii) Head-driven Phrase Structure Grammar (iii) Combinatory Categorial Grammar(iv) Lexical-FunctionalGrammar and (v) Dependency Grammar.This section also discusses general strategies of parsingsuch as top-down and bottom-up parsing and types of parsing in implementation.

### 2.1. Generalised Phrase Structure Grammar (GPSG)

GPSG, a constraint-based grammar was developed by Gerald Gazdar in the 1970s with Ewan Klein, Ivan Sag, and Geoffrey Pullum for English (Cf. Gazdar, G., et.al, 1985), deriving from constituency grammar. One of its main goals is to show that natural languages can be expressed in context-free grammars. The analysis of GPSGfor example (3) is given below.

(3) He gave the woman the gift

**GPSG output:**

((NP-he (N)))((VP-gave (V)))((NP- the (DET) woman (N)))((NP-the (DET) gift (N)))

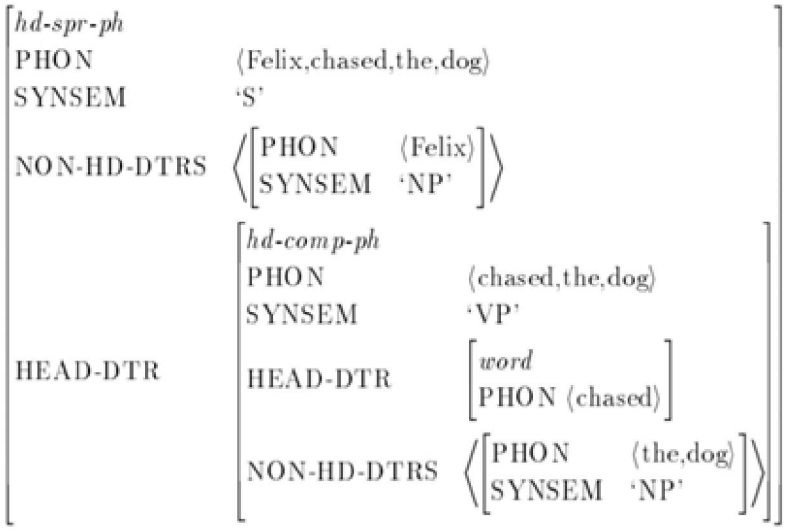
Later, Bahrani, M. et.al. (2011) claimed that GPSG, a unification-based formal theoryapplies to languages like Persian, French, Chinese and Arabic. It was provided with an instance of Persian output, which had84.5% parsed output out of 89% accepted sentences among1200 sentences (with varying contexts), using a hybrid approach and following X- bar theory. Philips J.D. (1992) reasoned out that the annotated representations were not enough for the parser to interpret several hundred rules (especially for constituent order languages) that were formulated by GPSG to analyse the natural language and as a result, a parsed output of above 90% was unachievable, even after revising the rules of GPSG.

## 2.2. Head-driven Phrase Structure Grammar (HPSG)

HPSG, the successor of GPSG is a lexical- based, constrained PSG developed by Carl Pollard and Ivan Sag (Cf. Pollard, C., and Sag, I. A. 1994).It deals with the *sign*, taking *words* and *features* as sub-types of the sign. Words are represented by PHON and SYSTEM. These signs and the formulated rules are together called feature structures. The structure of HPSG output for the example (4) is seen in figure 1 below:

(4) Felix chased the dog

**HPSG output (**extracted from Sag, I. A., 1995:15):



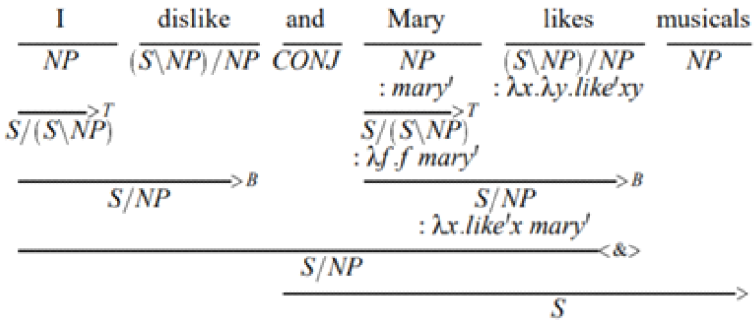

Levine, R. D. and Meurers, W. D. (2006) claimed that HPSG is suitable toapply for languages such as Romance languages, Slavic languages, German, Japanese, Welsh, English, Korean and Warlpiri .The theory was applied in 'Enju', an English probabilistic HPSG parser developed by Tsujii Laboratory (University of Tokyo), Japan, which was derived from Penn Treebank. Proudian, D. and Pollard, C. (1985) have stated that HPSG would be comparatively advantageous as it could give a good accuracy because importance is givento the heads of the phrases unlike GPSG.

## 2.1. Combinatory Categorial Grammar (CCG)

The onset of CCG in computational linguistics was in the late 1970s and early 1980s, where the categorial grammar was extended with functional operators like the functional composition, substitution, etc. This grammar focused on grammatical constituents which were differentiated by syntactic types, identifying them either as a function from arguments or as an argument(Steedman, M., and Baldridge, J. 2011).The example (5)is taken from Mark Steedman (1996:4):

(1) I dislike and Mary likes musicals

**CCG Parser Output:**




CCG is also used in Penn Treebank even though it has a huge number of categories. The trained corpus has 1207 categorial lexicons which are compared with 48 POS tags of Penn Treebank. On the other hand, CCG has a very few grammatical rules to accommodate such ahuge number of categories, resulting in a less over-generating grammar.The final recall of the system is 89.9% against the trained data(Hockenmaier, J., and Steedman, M., 2002:342).

## 2.1. Lexical Functional Grammar (LFG)

LFG is first published by Joan Bresnan (1982),which addresses the mechanisms to extract grammatical relations from a sentence in a positional language such as English. Neidle (1994)claims that the main focus of LFG is on syntax even though the grammar extends its relation with morphology and semantics.LFG is being represented in constituent and functional structureas seen below:

**LFG outputfor the example (3):**

| | | |
|---|---|---|
| SUBJ | PRED | 'he' |
| | NUM | sg |
| | PERS | 1 |
| | DEF | + |
| PRED | 'give <agent,goal,patient>' | |
| | SUBJ,OBJ2,OBJ1 | |
| TENSE | past | |
| OBJ2 | PRED | 'the woman' |
| | NUM | sg |
| | PERS | 3 |
| | DEF | + |
| OBJ1 | PRED | 'the gift' |
| | NUM | sg |
| | PERS | 3 |
| | DEF | + |

LFG is used in parsing Wall Street Journal (WSJ) by Stefan Riezler, et.al. (2002). The evaluation output is shown as 74.7% accuracy for full parses and 25.3% for fragment parses. In this process, Brown corpus isgold standardly annotated and trained. The major advantage found in LFG is that the mismatch between the surface structure and the deep argument structure as discussed in Chomskyan framework is not found here. However, LFGdoes not deal with lexical ambiguity, optional theta roles, adjuncts, and mapping from grammatical relations to theta- roles (Bharati, A., et.al, 1995).

### 2.1. Dependency Grammar (DG)

Dependency approach (both projective and non-projective) is one of the approaches to automatically parse the natural language, following the grammatical tradition of dependency gramm--ar, tracing back to Panini's grammar. While old school of thought is still in practice, the modern thought was proposed by aFrench linguist Lucien Tesnière during post-1950s, which attempted to capture the grammar of all typologies including Indian languages' typology.The dependency grammar structure represents the relation between the head and its dependents through directed arcs and the functional categories in the form of arc labels.Content words are marked by dependency relations;

functional words attach to the content words they modify and punctuation attach to the head of the phrase/ clause. The parse is a tree, where the nodes stand for the words in an utterance and the link between the words represent the relation between the pair of words. Such dependencies can either be argument dependencies (subject, object, indirect object, etc.) or modifier dependencies (determiner, noun modifier, verb modifier, etc.). The dependency parseroutputfor the example (3) is given below.

**Dependency parser output:**


In the above output, the verb root is considered as the head and other noun phrases are related with the verb as its dependents with various *kārakar*oles (i.e SUBject, OBJect, IndirectOBJect).The dependency grammar is widely usedin building parsers for Indian languages (*see* section 6.2 for more information).

### 3.Parsing Technique

The parsing systemis implemented in two different styles: (i) top-down parsing (ii) bottom-up parsing.

### 3.1. Top-down parsing:

Top-down parsing attempts to construct a parsed tree for the input from the root (top) to the leaves (bottom), where the transitions of tokens are seen from left to right, attempting to resolve ambiguities by changing the rules of right hand side. The major advantage of such systems is that it never wastes time in validating trees that would not lead to S (root) but the negative aspect is that it processes output before examining the input (Cf. Jurafsky, 2000: 356-359). This type of parsing uses Context-free grammar, which is a set of rewriting rules. Recursive descent parsers and $LL^1$ parsers are examples of such kind of parsing.

### 3.2. Bottom-up parsing:

Bottom-up parsing constructs a parsed tree from the leaves to the root, i.e. from bottom to top. The positive aspect of such systems is that it never suggests a tree that is not grounded to input but never reaches to the root, S. Grammar formalisms such asGPSG, HPSG, CCG, LFG, and DG are applied through bottom-up parsing.LR[2] or shift reducing parsers like MALT[3] are examples for such parsing.

## 4. Types of Parser

The parser is implemented in various ways with the suitable grammar formalism. In this section the four major types of parsers are discussed.

### 4.1.Rule-based Parser

Rule-based parsing uses pre-written rules to describe the data. Using rule-based parsing, the main functor-argument relations are obtained. Grammar-driven dependency parsing is a type of rule-based parsing, which is formed from the combinations of context-free and constraint -basedDependency Parsing (DP), which is well defined by the grammar of the language. An input sentence of the language is validated, only when it is accepted by the grammar of the language as the formal language is vital in this approach (Cf. Kübler, S., Ryan McDonald, Joakim Nivre and Graeme Hirst, 2009: 64-70). Weighted Constraint Dependency Grammar (WDCG)is another example of rule-based parsing (Krivanek, J., and Meurers, D., 2013).

### 4.2. Statistical Parser

In Statistical parser, grammar rules are associated with probability of a complete parse of a sentence. In building statistical parsing, the commonly used grammar formalism is probabilistic context-free grammars (PCFG). Data-driven dependency parsers (Nivre, 2006) follow machine learning approach and thus, it is widely used in statistical models. Any sentence or phrase given as input is considered as a valid grammatical sentence and parsed. Data-driven dependency parsing is sub-categorized into two types, transition-based dependency parsing and graph-based dependency parsing. MALT is the best example forstatistical parser. However, statistical parser has

drawbacks like lack of lexical conditioning and poor independence assumptions but can be improved by annotating bigger data (Jurafsky, 2000).

### 4.3. Hybrid Parser

A combination of rule-basedand statistical parsing results in hybrid parsing, where the rules are applied to sentences after the machine learning. It gives better accuracy than the rule-based and probabilistic/ stochastic models as both these models are inbuilt in this system. The requirement includes fully annotated treebank for probabilistic parsing and fully developed rules for the second phase of implementation (Cf. Kilian A. Foth, Wolfgang Menzel, 2006).

### 4.4. Neural Network Based Parser

Neural network based parser works in dependency approach in both transition-based and graph-based dependency parsing. Yet, commonly found neural network parsers use transition based dependency parsing, where the parser is powered by neural network[4]. The input word embeddings are represented in vectors as shown in Figure (1).

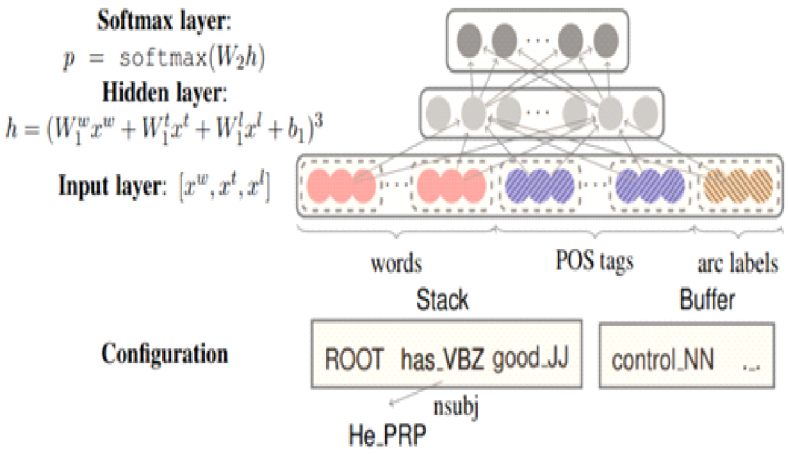



**Figure 1:** Neural Network Schema (Danqi Chen and Christopher D. Manning, 2014).

## 4. Annotation Schema

Annotation schema is used in parsing to have a uniform pattern in marking features of a big data. There are many annotation schemawith tagset and guidelinesare available in building parsers. In this section, the tagsets such asPenn tagset (Santorini, B., 1990), UCREL parsing tagset (ucrel.lancs.ac.uk), Prague tagset(ufal.mff.cuni.cz), Stanford tagset(De Marneffe, M. C., and Manning, C. D., 2008), Chinese Dependency tagset(Liu, H., and Huang, W., 2006), Anncorra tagset (Bharati, A., et al. 2009), Universal Dependency (UD) tagset (universaldependencies.org) are discussed. Among these tagsets, Anncorra tagset (Pāninian framework) and Universal Dependency tagset are taken into a detailed discussion, as it is mostly used in implementing parsers for Indian languages.

### 5.1. Penn Tagset

Penn Treebank (1989-1996), developed by University of Pennsylvania contains the POS tagged and syntactically bracketed forms of Brown corpus and Wall Street Journal. The Treebank has annotated 7 million words of part-of-speech tagged text, 3 million words of skeletally parsed text, over 2 million words of text parsed for predicate argument structure, and 1.6 million words of transcribed spoken text annotated for speech disfluencies (Cf. Taylor, A., M. Marcus, and B. Santorini, 2003). It has 42 fine grained POS tags, 8 chunk tags, and 9coarse grained relation tags, which are used in parsing. The major aim in introducing the tagset was to reduce the lexical and syntactic redundancy (Cf. Santorini, B., 1990).

### 5.2.UCREL Parsing Tagset

The UCREL tagset, developed by Lancaster University is used in semantic analysis systems of English. It has 21 coarse-grained discourse tags and 232 fine-grained semantic tags (Paul Rayson, et.al., 2004). The accuracy of the manually tagged system developed by Paul Rayson, Dawn Archer, Scott Piaoand Tony McEnery had a precision of 91%.

### 5.3. Prague Dependency Tagset

Prague Dependency Treebankwasdeveloped by Prague School of Functional and Structural Linguistics. The project began in 1995

with the notion of following Praguian dependency tradition and building a Treebank similar to Penn Treebank. They have collected the database from the Czech National Corpus (Charles University, under the guidance of F. Čermák co-joint with other research centres/ institutions), and developed a three-layer system of tags: morphemic, syntactic at analytical level, and syntactic at tectogrammatical level (The Prague Dependency Treebank 3.0.). They had developed 68 fine-grained POS tagsets.(https://ufal.mff.cui.cz/).

### 5.4. Stanford Dependency Tagset

Stanford Dependency tagset was developed by a group of people from Linguistics and Computer Science as a part of AI lab in 2005 for English. It was later extended to Chinese, Italian, Bulgarian and Portuguese.The main was to have a simple representation of the analysed sentences which could be used by commons to extract word relations. The present Stanford Dependency Treebank has an approximate count of 50 relation tags (Cf. De Marneffe, M. C., and Manning, C. D., 2008). Most of these tags are also seen in Universal Dependency tagset.

### 5.5. Chinese Dependency Tagset

Chinese Dependency Treebank 1.0 was released on May 2012 in Harbin Institute of Technologys Research Center for Social Computing and Information Retrieval (HIT-SCIR)for Mandarin Chinese, Chinese. It was developed by Wanxiang Che, Zhenghua Li, Ting Liu. It contains 49,996 Chinese sentences with 902,191 words, which were sourced from Peoples Daily newswire stories (1992-1996) and annotated with syntactic dependency structures.The data is provided in the format of CoNLL-X and in UTF-8 and has 13 word class tags and 34 fine grained dependency tags (Cf. Liu, H., and Huang, W., 2006).

### 5.6. Anncorra Tagset

AnnotatedCorpora (Anncorra) is developed based on the Pāninian Dependency grammar with *kāraka* and non-*kāraka* relations aiming at a uniform representation of annotated corpus of Indian languages(Bharati, A., et.al, 2002). It is developed for parsing

Hindi sentences and the tags are given accordingly. Later, the same guidelines were adapted for other Indian languages (Marathi, Urdu, Bengali, Kannada, Telugu, Tamil, and Malayalam) (Cf. Tandon, J. and Sharma, D. M., 2017). It was even used by Amita, A. J. (2015) for English, in which HyDT annotation scheme and hybrid approach (statistical+ rule based) were used for parsing 2000 words.

The 19 fine-grained *kāraka* relations that are included in the Anncorra tagset are k1 (*karta* 'doer/agent/ subject'), pk1 (*prayojakakarta*'causer'), jk1 (*prayojyakarta*'causee'), mk1 (*madhyasthakarta*' mediator-causer'), k1s (*kartasama nadhikarana*- 'noun complement of *karta*'), k2 (*karma* 'object/ patient'), k2p (Goal, Destination), k2g (secondary *karma*), k2s (*karma samanadhikarana*'object complement'), k3 (*karana* 'instrument'), k4 (*sampradana* 'recipient'), k4a (*anubhavakarta* 'Experiencer'), k5 (*apadana* 'source'), k5prk (prakruti apadana 'source material'), k7t (*kAlAdhikarana* 'location in time'), k7p(*deshadhikarana* 'location in space'), k7(*vishayadhikarana* 'location elsewhere'), k7a (according to) and k*u (*sAdrishya* 'similarity/comparison').

The 25 fine-grained non-*kāraka* relations include genitive case, adverbial and adjectival relations. It includes, r6 (*shashthi* 'genitive/ possessive'), r6-k1, r6-k2 (*karta* or *karma* of a conjunct verb (complex predicate)), r6v (*kA* 'relation between a noun and a verb), adv (*kriyAvisheSaNa* 'manner adverbs'), sent-adv (Sentential Adverbs), rd (direction), rh (*hetu* 'reason'), rt (*tadarthya* 'purpose'), ras-k* (*upapada sahakArakatwa* 'associative'), ras-neg (Negation in Associative), rs (noun elaboration), rsp (address terms), nmod__relc, jjmod__relc, rbmod__relc (relative clauses, *jo-vo* constructions), nmod (participles etc. modifying nouns), vmod (verb modifier), jjmod (D-Rel modifiers of the adjectives), pof (part of units such as conjunct verbs), ccof (co-ordination and sub-ordination), fragof (Fragment of), enm (enumerator), rsym (ag for a symbol) and psp_cl (relation between clause and postposition following that clause).

### 5.7. Universal Dependency(UD) tagset

Universal Dependency is across-linguistic project, built with the goal of facilitating multilingual parser development, cross-lingual

learning, and parsing research from a language typology perspective (http://universaldependencies.org). The idea of annotation scheme has been taken from Stanford dependencies, Google universal part-of-speech tags and the Interset interlingua for morphosyntactic tagsets in 2013 (McDonald et al., 2013). The parser has a lesser number of modules (Preprocessing (transliteration, sentence segmentation, tokenization); M-layer annotation (positional tagging) and A- layer annotation (dependency annotation), making it universal for any language typology. Indian languages like Hindi, Marathi, Sanskrit, Tamil, Telugu, and Urdu are included in the existing UD Treebank and Kannada and Pnar are upcoming languages listed in the UD website.

| | Nominals | Clauses | Modifier words | Function Words |
|---|---|---|---|---|
| Core arguments | nsubj<br>obj<br>iobj | csubj<br>ccomp<br>xcomp | | |
| Non-core dependents | obl<br>vocative<br>expl<br>dislocated | advcl | advmod*<br>discourse | aux<br>cop<br>mark |
| Nominal dependents | nmod<br>appos<br>nummod | acl | amod | det<br>clf<br>case |
| **Coordination** | **MWE** | **Loose** | **Special** | **Other** |
| conj<br>cc | fixed<br>flat<br>compound | list<br>parataxis | orphan<br>goeswith<br>reparandum | punct<br>root<br>dep |

**Figure 2: Universal Dependency tagset**(http://niversaldependencies.org)

The figure 2 lists 37 coarse grained UD tags including*acl* (clausal modifier of noun (adjectival clause)), *advcl*(adverbial clause modifier), *advmod* (adverbial modifier), *amod* (adjectival modifier), *appos* (appositional modifier), *aux* (auxiliary), *case* (case marking), *cc* (coordinating conjunction), *ccomp* (clausal complement), *clf* (classifier), *compound* (compound), *conj* (conjunct), *cop* (copula), *csubj* (clausal subject), *dep* (unspecified dependency), *det* (determiner), *discourse* (discourse element), *dislocated* (dislocated

elements), *expl* (expletive), *fixed* (fixed multiword expression), *flat* (flat multiword expression), *goeswith* (goes with), *iobj* (indirect object), *list* (list), *mark* (marker), *nmod* (nominal modifier), *nsubj* (nominal subject), *nummod* (numeric modifier), *obj* (object), *obl* (oblique nominal), *orphan* (orphan), *parataxis* (parataxis), *punct* (punctuation), *reparandum* (overridden disfluency), *root* (root), *vocative* (vocative) and *xcomp* (open clausal complement).

Apart from these listed universal tags, there are 198 language specific tags that are used in various languages parsing system. For instance, in Tamil, experiencer subjectsin dative subject construction require a different tag as it is inflected with the non-nominative case and verbs do not agree with the so-called subject. The experiencer subjects are given with the tag 'nsubj:nc'i.e.. non-canonical subjects/ dative subjects.

(6) *eIa-kku* *pamam* *pār.kk-a* *vēGmum*
I-DAT movie see-INF want
'I want to see a movie'

In the example (1), the dative-marked subject *eIa-kku* 'I-DAT' is given with the tag 'nsubj:nc'.

The table (1)provides the statistics of Indian language in UD project.

| S.No. | Language | No. of annotated sentences | No. of UD tags | No. of language-specific tags |
|---|---|---|---|---|
| 1 | Hindi | 17,647 | 13 | 3 |
| 2 | Urdu | 5130 | 16 | 1 |
| 3 | Telugu | 1328 | 14 | 11 |
| 4 | Tamil | 690 | 14 | 3 |
| 5 | Marathi | 466 | 15 | 8 |
| 6 | Sanskrit | 225 | 14 | 11 |

Table (1): Statistics of Indian language in UD project.

## 5. Parsers: A Review

Parsers are being implemented worldwide for various languages. This section deals with a review of parsing in the following languages:

(i) World languages

(ii) Indian languages

(iii) Tamil

### 5.1. WorldLanguages

A parsing algorithm recorded to be the earliest was proposed by Yngve (1955). Yet, most of the parsers were developed in early 1990s. Such implemented parsers for world languages include:

- Collins's (1999) statistical parser for Czech using Prague Dependency Treebank
- Eugene Charniak's (2000)maximum-entropy parser for English
- Bikel and Chiang's (2000 first statistical model on Chinese Treebank
- DeSR, developed by Yamada and Matsumoto(2003) for English
- Dubey and Keller's (2003) proposal of a probabilistic parsing for German
- Stanford parser (2003), a statistical parser, using lexicalized PCFG (Probabilistic Context-Free Grammar), developed by Dan Kleinbeing for English and further extended to Arabic, Chinese, French, German and Spanish
- A probabilistic parser with supervised learning based on PCFG for English (Collins, 1997)
- Robust Accurate Statistical Parsing (RASP) System, a hybrid domain independent English parser (Cf. Briscoe, et.al,2006)
- MALT (2007) and MST (2005) (developed by Johan Hall, Jens Nilsson and Joakim Nivre at Växjö University and Uppsala University, Sweden), a transition based parser
- ISBN (Incremental Sigmoid Belief Networks), a trainable dependency parser (Cf. Titov, I. and Henderson, J., 2010)

- Carnegie-Mellon's Link Grammar parser, built for English, Arabic, Russian and Persian
- Seraji, M., Jahani, C., Megyesi, B., and Nivre's (2014) work on Persian by obtaining the data from large-FARSDAT
- Universal Dependencies (UD) (McDonald et al., 2013), a project developed by Joakim Nivre, which is involved in developing a cross-linguistic study, maintaining a treebank annotation for 60 languages with 102 treebanks

### 5.2. Indian Languages

Parsing is one such area, which has to be explored in depth for Indian languages. Some of the Indian languages including, Hindi, Urdu, Telugu, Kannada, Tamil, Bengali, Marathi, and Assamese have delved into area of parsing, whichare still work in progress. (Cf. Monika T. Makwana and Deepak C. Vegda, 2015). In fact,Tandon, J., and Sharma, D. M. (2017)hascome up with a unified strategy for parsing Indian languages usingPāninian framework. Researches related to cost-effective methods of building dependency parser for Indian languages are also in thecurrent trend (Cf.Tammewar, A. 2015). The list of implemented parsers in Indian languages between 2009 and 2018is discussed below:

- Nivre (2009) optimizied MALT parser for Hindi, Bengali and Telugu. With coarse-grained tagset, the respective accuracies are 81.1%, 79.6% and 63%. But, when fine-grained tagset is used, it had lower accuracies, i.e. 75.3%, 72.9% and 58.5%.
- Hindi, Bengali and Telugu sentences are tested with MALT and MST (data-driven parsers) by Bharat Ram Ambati, et.al.(2009), where MALT has a better performance than MST. The report has a final average score as 88.43%, 71.71% and 73.81% respectively.
- A bidirectional dependency parser for Hindi, Bengali and Telugu is proposed by Prashanth Mannem (2009), which shows the accuracy of 71.63%, 59.86% and 67.74% respectively when run with test data. The same data has better accuracies with

the coarse-grained tagset, 76.90%, 70.34% and 65.01% respectively.

- A constraint based dependency parsing system for Bengali with Pāninian Grammar formalism is proposed by Sankar De, et.al. (2009), which is trained with 1000 annotated sentences, and evaluated with 150 sentences. Ithas the accuracy of 79.81%, 90.32% and 81.27% for labelled attachments (LAS), unlabelled attachments (UAS) and label scores (LS) respectively.
- Aniruddha Ghosh, et.al. (2009) trainsBengali data using CRF and was implemented using rule-based algorithm. It results in 74.09% (LAS), 53.09% (UAS) and 61.71% (LS).
- Sanjay, et.al. (2009) has run Bengali sentences on a data-driven parser and hybrid parser.The wrongly annotated sentences are given rules to improve the accuracy. A special look at subject, object, location and relation is observed.
- Rahman, Mirzanur, et.al. (2009) analyse the issues in areas of parsing Assamese sentences when tagged with 7 tags based on CFG formalism. Later, rules are developed accordingly and algorithms are modified from Earley's Algorithm to solve those issues.
- Aconstraint-based Hindi dependency parsing system with the accuracy of 62.20% (LAS)and 85.55% (UAS) is implemented by Meher Vijay Yeleti and Kalyan Deepak (2009).
- Bharat Ram Ambati, et.al. (2010)analyse the role of linguistic features in data-driven dependency parsing for Hindi and found that accuracy gain is seen when adding morphosyntactic features like case and TAM features. They had finally gained 2% accuracy (76.5% in total) after combining morph features from two different parsers.
- Antony P.J. (2010) has developed a statistical syntactic Kannada parser using Penn Treebank with 1000 POS tagged sentences using SVM POS tagger. It is implemented using

supervised machine learning and is evaluated using SVM algorithms. As a result, they claim to have good accuracy.

- B.M.Sagar (2010) has developed a CFG for Kannada parser and finally proposes that top-down parser is best suited for Kannada.
- Navanath Saharia, et.al. (2011) have used CFG to parse the simple sentences of Assamese, which is not implemented.
- B.Venkata S. Kumari, et.al. (2012) use a combination of MALT and MST parserswhich shows LAS 90.66% for gold standard and 80.77% for automatic tracks.
- Karan Singh, et.al. (2012) propose a two-stage approach for Hindi Dependency Parsing using MALT parser. Their system has a record of 90.99% (LAS) for thegold standard.
- Uma Maheshwar Rao G., K. Rajya Rama, A. Srinivas (2012) has worked on Dative case towards building a parser. Various functions of the dative marker is discussed and a flowchart is developedto build a robust parser for Telugu.
- Sambhav Jain, et.al. (2013) has added the ontological features to Hindi dependency parser which added the accuracy improvement of 1.1% (LAS) for 1000 sentences and 0.2% (LAS) for 13371 sentences.
- A Lexicon parser for Devanagiri script (Hindi)is developed by Swati Ramteke, et.al.(2014), which generated semantic parsed trees with an accuracy of 89.33% when run with unambiguous sentences.Rule-based approach was used to resolve the lexical ambiguities.
- Arpita Batra, Soma Paul, and Amba Kulkarni (2014) had worked on the constituency analysis for Hindi using four approaches. Adjacency global, adjacency greedy, dependency global and dependency greedy were applied for 2322 sequences of words. Applying all these approaches, 92.85% (using global dependency algorithm and syntactic rules) accuracy was obtained.
- Dhanashree Kulkarni, et.al. (2014) has taken up CFG as the grammar formalism and used the same in Top-Bottom and

Bottom-Top parser for Marathi. The final outcome of the paper was to develop (computerized) grammar checking for Marathi text from CFG perspective.

- A Combinatory Categorical Grammar (CCG) Telugu treebank is created using CCG lexicon and dependency Treebank and it is tagged with CCG supertags as features to Telugu dependency parser. An improvement of 1.8% in UAS and 2.2% in LAS (especially on verbal arguments) was observed when implemented using MST parser (Cf. Kumari, B. and Rao, R. R., 2015).

- Telugu Dependency parser, developed by Nagaraju, G. et.al. (2016) have used bottom-up parser and parsed 200 Telugu sentences using *kāraka* relations. Out of 200 sentences, they have obtained 178 correct parsed sentences. As a whole, 880 words were correctly tagged and 140 were incorrect and thus, they claim the precision to be 99.

- 'Improving Transition-Based Dependency Parsing of Hindi and Urdu by Modeling Syntactically Relevant Phenomena', by Bhat, R. A., Bhat, I. A., and Sharma, D. M. (2017) have used *kāraka* and non-*kāraka* relations and annotated the inter-chunk dependencies manually. They have implemented in transition-based dependency parser with syntactic features and obtained an accuracy of 87.82% in trained set and 87.72 in test data of LAS.

### 5.3. Tamil

Tamil, belonging to Dravidian language family, is morphologically rich. It has a (S)OV word order with agglutinative morphology. Hence, building a parser for Tamil is a challenging task. This section lists the Tamil parsers with grammar formalisms and techniques used in their respective parsers.

- Hybrid approach combining PSG and DG with Lexicalized and Statistical Parsing (LSP) is used by Selvam, M., Natarajan, A. M. and Thangarajan, R.(2008) with 500 tags and 31 dependency relations on Tamil. 3261 sentences with 51026 words are used and as a result, 73% accuracy in trained data and 65% accuracy

in test data with just 600 trained sentences were obtained.The lacuna is seen in their choice of their tagset which had 500 tags.

- Loganathan Ramasamy and ZdenkŽabokrtský (2011) have done initial experiments with Tamil dependency parsing using rule-based approach (with an accuracy of 79% (LAS) in the trained data and 61% with the test data) and corpus based approach (with an accuracy of 75%. Finally, it is concluded that both the approaches have failed in identifying coordination nodes.
- Universal Dependency has extended its system to Tamil, by developing a Tamil Treebank (from Prague dependency Treebank) (Cf. Ramasamy, Loganathan and Zdenk Žabokrtský,2012), which has universal tagsets and just involves three processes: Pre-processing (transliteration, sentence segmentation, and tokenization); M-layer annotation (positional tagging) and A- layer annotation (dependency annotation) with 217 distinct tags (including all 9 positions). 96% of the test data was unambiguous; 3% was ambiguous with 2 tags and tokens with 3-4 tags were just 1% which is negligible. Altogether, 21 dependency relations are used for labeling edges. It has an accuracy of 69% when trained with 690 sentences.
- A Tamil syntactic parser, proposed by K. Sureka, Dr. K. G. Srinivasagan and S. Suganth (2014) works on dependency grammar and follows hybrid approach, with clause boundary identifier. After adding the module, the result obtained is that out of 150 sentences, 120 are parsedcorrectly.
- Vigneshwaran (2017) hasworked on Tamil parsing based on cognitive grammar as the theoretical grammarand Pāninian framework as the computational grammar. The main argument revolves around parsing Tamil sentencesat discourse level, as it claims that sentential analysis is not enough to get an idea of the complete context of the text.

## 7. A Discussion on a Suitable Model for Tamil

Tamil, a Dravidian language is morphologically rich, when compared to Indo-Aryan languages, hencebuilding a syntactic parser becomes a complex task. A range of language-specific annotation

schema is required to cover a good range of sentence structures for a better accuracy.

The dependency approachis considered to be the best suitable model for Tamil over other formalisms available. It has a better reach than the Constituency approach as it accommodates maximum typologies of languages, making it universal by accommodating a maximum number of languages. Phrase Structure Grammar and other formalisms are related to constituency parsing have failed to look at the semantics, which has led to multiple drawbacks including no mechanisms for resolving structural ambiguities, not able to map theta rolesetc. Moreover, constituency parsing's average number of nodes is twice as the number of dependency nodes. Thus, automated DP is faster than constituency parsing.

When Indian languages with different typologies try to adapt the Pāninian framework which was actually designed for Sanskrit, adding the language specific featuresmaygain good accuracy. For instance, Anncorra tagset, designed based on*kāraka* systems may not accommodate a range of sentence structures that are available in Tamil. Similarly, Universal Dependencies have a coarse grained tagset for all languages, though some language-specific tags are introduced to some languages. Whereas, Tamil needs more tags to develop a well-annotated corpus.

As seen in reviews of parsers in different languages, hybridparsing is considered to have advantageous than other methods as it combines multiple parsing strategies like rule-based and supervised probabilistic parsing. Here, the rules and machine learning algorithms are blended to ensure the effectiveness of the output. Thus, it would give a better output than the standalone rule-based or probabilistic models for any language.

## 8. Conclusion

The paper has listed a number ofgrammar formalisms in parsing, parsing techniques, types of parser, and annotation schema that are available for languages. A number ofimplemented parsers in Indian

languages with a special reference to Tamil are discussed. The conclusion inferred from a discussion on a suitable model for Indian Languages is that both UD and Anncorra guidelines have its respective drawbacks. Developing language specific tags would improve the accuracies in both the cases. Implementing well-annotated corpus in Hybrid system would give better accuracies for Tamil.

## References


Ambati, B. R., Gadde, P., & Jindal, K. 2009. Experiments in Indian Language Dependency Parsing. In the *Proceedings of the ICON09 NLP Tools Contest: Indian Language Dependency Parsing*, pp. 32-37.

Amita, A. J. 2015. An Annotation Scheme for English Language Using Paninian Framework. *IJISET*, Vol. 2, pp. 616-619.

Bahrani, M., Sameti, H., and Manshadi, M. H. 2011. *A Computational Grammar for Persian Based on GPSG*. Retrieved on 21st November, 2017. https://link.springer.com/article/10.1007/s10579-011-9144-1

Bresnan, J. 1982. Control and complementation. In *Linguistic inquiry*. Vol. 13. No. 3, pp. 343-434.

Bharati, A., Sangal, R. 1993. Parsing Free Word Order Languages in the Paninian Framework.*In: Proceedings of the 31st Annual Meeting on Association for Computational Linguistics*, pp. 105–111.

Bharati, A., Sangal, R., Chaitanya, V., Kulkarni, A., Sharma, D. M., and Ramakrishnamacharyulu, K. V. 2002. AnnCorra: Building Tree-banks in Indian Languages. In *Proceedings of the 3rd Workshop on Asian language Resources and International Standardization:Association for Computational Linguistics*. Vol 12 (pp. 1-8).

Bharati, A., Sharma, D. M., Husain, S., Bai, L., Begam, R., and Sangal, R. 2009. *Anncorra: Treebanks for Indian Languages, Guidelines for Annotating Hindi Treebank.*Retrieved on 1st December, 2016. http://aclweb.org/anthology/W17-63

Bharati, A., Vineet Chaitanya and Sangal, R. 1995. *Natural Language Processing: A Paninian Perspective.* New Delhi: Prentice-Hall of India.

Bhat, R. A., Bhat, I. A., and Sharma, D. M. 2017. Improving Transition-Based Dependency Parsing of Hindi and Urdu by Modeling Syntactically Relevant Phenomena. In *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, Vol 16, No. 3, pp 17.

Bikel, D. M., and Chiang, D. 2000. Two statistical parsing models applied to the Chinese Treebank. In *ACL*, Vol 12, pp. 1-6.

Briscoe, E, Carroll, J and Watson, R. 2006. *The Robust Accurate Statistical Parsing (RASP). Retrieved on 20th February, 2018.* http://sro.sussex.ac.uk/25679/

Charniak, E. 2000. A maximum-entropy-inspired parser. In *ACL*, pp. 132-139.

Collins, M. (1997, July). Three Generative, Lexicalised Models for Statistical Parsing. In *Proceedings of the 8th Conference on European Chapter of theACL*, pp. 16-23.

Collins, M., Ramshaw, L., Hajič, J., and Tillmann, C. 1999. A statistical parser for Czech. In *ACL*, pp. 505-512.

Dan Klein and Christopher D. Manning. 2003a. Accurate Unlexicalized Parsing. In *Proceedings of the 41st ACL*, pp. 423-430.

Dan Klein and Christopher D. Manning. 2003b. Fast Exact Inference with a Factored Model for Natural Language Parsing. In *Advances in Neural Information Processing Systems 15 (NIPS 2002)*, Cambridge: MIT Press, pp. 3-10.

Danqi Chen and Christopher D. Manning 2014. A Fast and Accurate Dependency Parser using Neural Networks. In *EMNLP*, pp. 740-750.

De Marneffe, M. C., and Manning, C. D. 2008. *Stanford Typed Dependencies Manual*. California: Stanford University. Technical report, pp. 338-345.

Dhivya, R. 2011. *Dependency Parser for Tamil Using Machine Learning Approach.* Retrieved on 15th November, 2016. http://nlp.amrita.edu:8080/project/mhrd/ms/Tamil/DependencyParser_Dhivya_CEN09009.pdf

Dubey, A., and Keller, F. 2003. Probabilistic Parsing for German Using Sister-Head Dependencies. In *Proceedings of the 41st ACL*. Vol 1, pp. 96-103.

Foth, K. A., and Menzel, W. 2006. Hybrid parsing: Using Probabilistic Models as Predictors for a Symbolic Parser. In *ACL*, pp. 321-328.

Garapati, U. R., Koppaka, R., and Addanki, S. 2012. Dative Case in Telugu: A Parsing Perspective. In *Proceedings of the Workshop on Machine Translation and Parsing in Indian Languages*, pp. 123-132.

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. 1985. *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press.

Hockenmaier, J., and Steedman, M. 2002. Generative Models For Statistical Parsing with Combinatory Categorial Grammar. In *Proceedings of the 40th ACL*, pp. 335-342.

Husain, S. (2009). Dependency Parsers for Indian Languages. In *Proceedings of ICON09 NLP Tools Contest: Indian Language Dependency Parsing*.

Jurafsky, D. 2000. *Speech and Language Processing*. London: Pearson.

Klein, D., and Manning, C. D. 2003. Accurate Unlexicalized Parsing. In *Proceedings of the 41st ACL*. Vol 1, pp. 423-430.

Krivanek, J., and Meurers, D. 2013. Comparing Rule-Based And Data-Driven Dependency Parsing of Learner Language. In *Computational Dependency Theory*, pp. *258*-207.

Krivanek, J., and Meurers, D. 2011. *Comparing Rule-Based and Data-Driven Dependency Parsing of Learner Language*. Retrieved on 1st May, 2018. http://www.depling.org/proceedingsDepling2011/papers/krivanekMeurers.pdf

Kübler, S., Ryan McDonald, Joakim Nivre, Graeme Hirst. 2009. Dependency Parsing. In *Synthesis Lectures on Human Language Technologies*. California: Morgan and Claypool Publishers. Vol 1. No.1.

Kumari, B., and Rao, R. R. 2015. Improving Telugu Dependency Parsing Using Combinatory Categorial Grammar Supertags. In *ACM Transactions on Asian and Low-Resource Language Information Processing*, Vol 14. No.1, pp. 3.

Levine, R. D., and Meurers, W. D. 2006. Head-Driven Phrase Structure Grammar. In *Encyclopedia of Language and Linguistics*. Vol.5, pp. 237-52.

Liu, H., and Huang, W. 2006. A Chinese Dependency Syntax for Treebanking. In *Proceedings of the 20th Pacific Asia Conference on Language, Information and Computation*, pp. 126-133.

Makwana, M.T. and D.C.Vegda. 2015. *Survey: Natural Language Parsing for Indian Languages*. Retrieved on 12th January, 2018. https://arxiv.org/ftp/arxiv/papers/1501/1501.07005.pdf

Mannem, P. 2009. Bidirectional Dependency Parser for Hindi, Telugu and Bangla. *Proceedings of ICON09 NLP Tools Contest: Indian Language Dependency Parsing*.

Mates, Benson. 1961.*Stoic Logic*. Berkeley: University of California Press.

McDonald, R., Crammer, K., and Pereira, F. 2005. Online Large-Margin Training Of Dependency Parsers. In *Proceedings of the 43rd ACL*, pp. 91-98.

McDonald, R., Nivre, J., Quirmbach-Brundage, Y., Goldberg, Y., Das, D., Ganchev, K., … andBedini, C. 2013. Universal Dependency Annotation for Multilingual Parsing. In *Proceedings of the 51st ACL*.Vol. 2, pp. 92-97.

Nagaraju, G., Mangathayaru, N., and Rani, B. P. 2016. Dependency Parser for Telugu Language. In *ACM,Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, pp. 138.

Neidle, C. 1994. Lexical-Functional Grammar (LFG): In *RE Asher, editor, Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press,pp. *9*-25.

Nivre, J. 2006. Inductive Dependency Parsing.In *Text, Speech and Language Technology*. Netherlands: Springer. Vol 34.

Nivre, J. 2009. Parsing Indian Languages with Maltparser. In the *Proceedings of the ICON09 NLP Tools Contest: Indian Language Dependency Parsing*, pp. 12-18.

Nivre, J., Hall, J., Nilsson, J., Chanev, A., Eryigit, G., Kübler, S., Marinov, S.and Marsi, E. 2007. MaltParser: A Language-Independent System for Data-Driven Dependency Parsing. *Natural Language Engineering*, Vol. 13. No. 2, pp. 95-135.

Phillips, J. D. 1992. A Computational Representation for Generalised Phrase-Structure Grammars. *Linguistics and Philosophy*. Vol.15. No.3, pp. 255-287.

Pollard, C., and Sag, I. A. 1994. *Head-Driven Phrase Structure Grammar*. US: University of Chicago Press.

*Prague Tagset*. Retrieved on 1st June, 2018.http:// ufal.mff.cuni.cz

Proudian, D., and Pollard, C. 1985. Parsing Head-Driven Phrase Structure Grammar. In *Proceedings of the 23rd ACL*. pp. 167-171.

Ramasamy, L., and Žabokrtský, Z. 2011. Tamil Dependency Parsing: Results Using Rule Based And Corpus Based Approaches. In *International Conference on Intelligent Text Processing and Computational Linguistics*. Berlin: Springer, pp. 82-95.

Ramasamy, L., and Žabokrtský, Z. 2012. Prague Dependency Style Treebank for Tamil. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*.

Rayson, P., Archer, D., Piao, S., and McEnery, A. M. 2004. *The UCREL Semantic Analysis System*. Retrieved on 21st January, 2018. http://eprints.lancs.ac.uk/1783/

Richard Socher, John Bauer, Christopher D. Manning and Andrew Y. Ng 2013. Parsing with Compositional Vector Grammars. In *ACL*. Vol 1, pp. 544-465.

Riezler, S., King, T. H., Kaplan, R. M., Crouch, R., Maxwell III, J. T., and Johnson, M. 2002. Parsing the Wall Street Journal Using a Lexical-Functional Grammar and Discriminative Estimation Techniques. In *Proceedings of the 40th ACL*, pp. 271-278.

Ryan McDonald, Joakim Nivre, Yvonne Quirmbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló, and Jungmee Lee. 2013. Universal Dependency Annotation for Multilingual Parsing. In *Proceedings of ACL*.

Sag, I. 1995. HPSG: Background and Basics, Stanford, CSLI; Retrieved on 1st November 2017. http: //hpsg.stanford.edu/hpsg.

Santorini, B. 1990. Part-Of-Speech Tagging Guidelines for the Penn Treebank Project (3rd revision). In *Technical Reports (CIS)*, pp. 570.

Seraji, M., Jahani, C., Megyesi, B., & Nivre, J. 2014. A Persian Treebank with Stanford Typed Dependencies. In *the 9th International Conference on Language Resources and Evaluation (LREC)*,pp. 796-801.

*Stanford Parser*. Retrived on 1st June, 2018. https:// nlp.stanford.edu/software/srparser.html.

Steedman, M. 1996. A Very Short Introduction to CCG. *Unpublished paper.*Retrieved on 21st September, 2018. http:/ /www.coqsci.ed.ac.uk/steedman/paper.html

Steedman, M., and Baldridge, J. 2011. Combinatory Categorial Grammar. In *Non Transformational Syntax: Formal and Explicit Models of Grammar*, pp. 181-224.

Stolcke, A. 1995. An Efficient Probabilistic Context-Free Parsing Algorithm that Computes Prefix Probabilities. In *Computational linguistics*, Vol.21. No.2, pp. 165-201.

Sureka, K., Srinivasagan, K. G., and Suganthi, S. 2014. An Efficiency Dependency Parser Using Hybrid Approach for Tamil Language. Retrieved on 20th December, 2017. https://arxiv.org/abs/ 1403.6381

Tammewar, A. 2015. *Cost Effective Dependency Parsing for Indian Languages.*Hyderabad: IIIT.MS dissertation.

Tandon, J., and Sharma, D. M. 2017. Unity in Diversity: A Unified Parsing Strategy for Major Indian Languages. In *Proceedings of the Fourth International Conference on Dependency Linguistics*,pp. 255-265.

Taylor, A., Marcus, M., andSantorini, B. 2003. The Penn Treebank: An Overview. In *Treebanks*. Dordrecht:Springer,pp. 5-22.

Titov, I., and Henderson, J. 2010. A Latent Variable Model for Generative Dependency Parsing. In *Trends in Parsing Technology.* Netherland: Springer,pp. 35-55.

*UCREL Parsing Tagset*. Retrieved on 1st June, 2018.http:// ucrel.lancs.ac.uk.

*Universal Dependency (UD) Tagset*. Retrieved on 1st June, 2018. http://universaldependencies.org

H. Yamada and Y. Matsumoto. 2003. Statistical Dependency Analysis with Support Vector Machines. In *Proceedings of the 8th International Workshop on Parsing Technologies(IWPT)*, pp. 195–206.

**(Footnotes)**

[1] LL- Left to right, Left most derivation type

[2] LR- Left to right, Right most derivation type

[3]MALT- MaltParser, developed by Johan Hall Jens Nilsson and Joakim Nivre at Växjö University and Uppsala University, weden is a system for data-driven dependency parsing, which can be used to induce a parsing model from treebank data and to parse new data using an induced model. (http://www.maltparser.org).

[4]Neural Network is an information processing paradigm, which is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems(https://www.doc.ic.ac.uk).

# GENDER DIFFERENCES IN KINSHIP TERMS OF TELUGU DIALECTS: A SOCIOLINGUISTIC STUDY

***- K. Balu Naik***

***Abstract***

*Research on language and gender studies reveals that gender differences are reflected through language and certain words are used with specific gender only. In this paper an attempt is made to study the kinship terms used for address and reference to different relatives in Telugu dialects. It will also examine the relationship of these relatives to ego and to each other.*

## Introduction

Kinship terms are classified on the basis of abstract relationship rules. These may be genealogical relations such as Son to Father, Daughter to Mother, and Daughter to Father. But, the classes bear no overall relation to genetic closeness. If a stranger marries into a society, for example, he may be placed to an appropriate class opposite to his spouse. Therefore, a chosen kinship term that merge or equate relatives genealogically is distinct from one another. But the same term is used for the different kin.

Studies on Dravidian Kinship terms (1964) revealed that it is of a classificatiory type of kinship system. Different languages organized these distinctions differently. Kin terms and terminologies are either descriptive or classificatory in nature. Six basic patterns of kinship terminologies are: (i) Hawaiian kinship, (ii) Sudani Kinship, (iii) Eskimo

Kinship, (iv) Iroquois kinship, (v) Crow kinship and (vi) Omaha kinship. Floyd Lounsbury (1964) proposed a seventh one for Dravidian type of terminological system, which in fact can go with Iroquois of Morgan type of kinship system. Because, both systems distinguish relations by marriage from relative by descent, which are classificatory categories rather than based on biological descent. The Dravidian kinship system involves selective "cousinhood". One's father's brother's children and one's mother's sister's children are not cousins but brothers and sisters i.e. one step removed.

They are considered consanguineous and marriage with them is strictly forbidden as it is incestuous. But one's father's sister's children and one's mother's brother's children are considered cousins and marriages are allowed between such cousins. Therefore, there is clear distinction between cross cousins and parallel cousins. Cross cousins are one's true cousins and parallel cousins are in fact siblings. There is a similarity between Iroquois kinship system and Dravidian kinship system referring to Father's sister as Mother-in-law and Mother's brother as Father-in-law. In some cultures the role of cross cousin is very important, because marriage is allowed in between them. On the other hand parallel cousins are considered the subject of promoted marriage, such as the preferential marriage of a father's brother's daughter which is common among some pastoral community. Such type of marriage helps to keep property within a lineage, whereas parallel cousin unions or parallel cousin marriages in some cultures would fall under incest taboo, because such type of parallel cousins are part of the subjects unilineage but cross cousin marriages are not. With this background, let us look into the gender differences.

Sociolinguistic studies since 1950's has identified sex as an important variable in language use. The authors made an attempt to explore the manifestation of Gender differences in address and reference of kinship terms and the semantics behind them in the Telugu dialects across different social classes, i.e. forward class (FC), backward class (BC), and scheduled class (SC)belonging to three regions, namely Telangana, Rayalaseema and Coastal regions. The data has been collected and analyzed keeping in mind the regional and social variables. The following kin terms illustrate the gender differences.

## 2.Earlier studies on kinship terms in Telugu

There are a couple of studies on Telugu kinship terminology.

The prominent ones are:

Rajagopal's (1970) 'Semantic analysis of Telugu kinship terminology'

Narasimha Reddy (1977) 'Telugu kinship terminology: A propositional perspective"

Trautmann (1981) "Dravidian Kinship"

Literature on feminism reveals that there are three waves: The first wave feminist's writers like Robin Lakoff's (1972) article has created lot of debate and interest in the study of language and gender, in which she presented that the language of women is tentative, powerless, and trivial; as such it disqualifies them from positions of power and authority. Therefore language functions as a tool of oppression and it is learned as part of learning to be a woman, imposed on women by societal norms, and in turn it keeps women in their place (Eckert and McConnell-Ginet, 2003).

The second wave of feminist linguists like Dale Spender (1980); Deborah Tannen's (1991) focus is on 'homogeneous women's language', which they assume is the result of either oppression of women or of the different socialising of women and men. In other words, this is also referred to as dominance and difference. Their analysis is concerned at the utterance level and concentrates on the production of individual speaker.

The third wave of feminist linguists view is of anti-essentialist analysis and they are critical of the second wave feminist linguists. They are concerned with the analysis of role of gender in language production and interpretation. It is more concerned with the contextual use of language and interlocutors concerned in the interaction (Mills, 2003).

The term community of practice (Wenger, 1998) is concerned with and groups of people who are drawn together in the performance of a particular task. This notion of community practice has been influential in feminist linguistics.

Eckert and McConnell-Ginet's (1998 & 1999) studies aim to produce more context based model of gender, where gender construction is constrained by its negotiations with suppositions of community rules of appropriacy and stereotypes.

Gender is dispersed into contextual elements rather than being located at the level of the individual. Gender may be described at the level of discourse rather than only at an individual and utterance level, where by settings strategies, discursive moves are normatively gendered by communities of practice and negotiated, contested affirmed and crucially changed by community members.

Gender is performed but within constraints established by communities of practice and our perceptions of what is appropriate within those communities of practice. There is an assumption that men are powerful and women as powerless, therefore, the gender difference reflected through language.

Mills (2003) tries to explain with a model of the relation between speakers and their communities, which is more concerned with the discourse level. Her model is concerned with the production and interpretation of language as rule governed at a level which is not within the control of individual speakers or of institutional forces.

## I.Methodology

### 3.1. Questionnaire

Structured questionnaire, interviews and by observer participant method data were collected. Apart from the above, data are also collected from our own fieldwork data on Sociolinguistic survey of Telugu dialects at and Telugu dialect dictionary preparation at Telugu University.

### 3.2. Selection of Informants

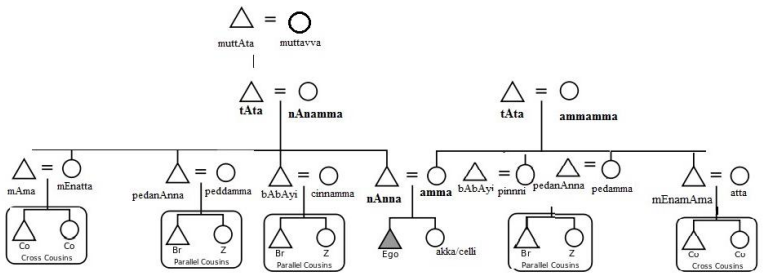
Informants were selected randomly from three regions of undivided Andhra Pradesh (i.e. Coastal, Rayalaseema and Telangana). They belong to different social classes background (SC, BC, and FC). The total number of informants are18, 9 male and 9 female. The informants selected fall into three age groups and there are:

15 to 30

31 to 45

46 to 65

The following diagram depicts the Telugu kinship terms:




## V.Discussion

The change of address system between wife and husband reflects the socio-cultural factors over the last fifty years. By and large earlier generations the husband would use the familiar address term *nuvvu* 'you sg.' *mi:ru* 'you pl.' Whereas the wife would use the polite form of honour

There is no more asymmetrical relationship between wife and husband among the educated young men and women, therefore, addressing by name is quite common, whereas in the older generation it is not so. In the present generation the address system between the couple has switched over to practice of using personal names *Ramana* and *Saroja* etc.

While addressing some terms for example *ra:* 'hey (third person sg. impolite)' which are meant for men being used for the opposite sex. This shows intimacy, attachment and endearment to the addressee. For example, the father would address his daughter as *e:ra:, na:jana, endukura:* etc.,

On the contrary there are also instances where, an elder person may address a young man with the terms meant for women for

example, *endukumma: anta ko:pam'* why are you so angry, my dear', which is typically meant for a girl is extended to a boy.

Similarly, among the boys and girls (friends) address patterns *Ore:j* 'hey man' and *Ose:j* 'hey girl' reflects the socio-cultural dynamics and also the impact of western and filmy culture. *adevaru na:ku tseppaVa:niki* 'who is she to tell me' and *va:dendi na:ku tseppe:di'* who is he to tell me'

In Rayalaseema *ammi/me:j* is used while addressing the girl younger to the ego and similarly *abbi/be:j* is used for addressing the boy. (*adiga:d*)*e:* 'hey' is actually a feminine form used for addressing women but nowadays very commonly used for endearment among men in Telangana dialect, while addressing close male relatives as well as across caste groups among friends. Similarly *abba:* 'hey' and *adiaa:dura*: 'it isn't my dear' are masculine forms used for men but, are used by female members in Telangana area. These forms are used to express endearment and solidarity with the ego.

*ma:ma:, ka:ka:, tsitsts:* 'uncle' are used to address friends among the men, and *mummy* 'mother' for daughters or for non relatives to express intimacy and endearment in Telangana dialect.

The use of double honorific forms for addressing certain kin relations is prevalent in Coastal dialect upper caste people. For ex. *na:nnaaa:randi* 'respected father', *ma:waaa:randi* 'respected father-in-law', *attaaa:randi* 'respected mother-in-law'. But not for addressing mother **ammaaa:randi* 'respected mother'. If the children do not use these forms they are considered as impolite or rude and are cautioned to be careful with their language.

*mit na:nna* ' a friendly father who is not a real father and also not a relative. By friendship they call themselves as *mitna:nna* (Srikakulam).

The following tables give the comparative statement of kinship terms with gender differences in Telugu dialects of the three classes:

| **F** *na:nna, tanḍri* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *ajja/ na:jina* | *ajja/ na:jina* | *na:nna/ ba:bu/ na:nnaga:ru* |
| **Rayalaseema** | *ajja* | *na:jina/ba:bu/ba:pu* | *na:jina* |
| **Coastal** | *na:jina* | *na:jina* | *na:nna ga:ru* |
| ***Reference Terms*** | | | |
| **Telangana** | *ma: na:jina* | *ma: na:jina/ ma: na:na* | *ma: na:nnaga:ru/ ma: na:na* |
| **Rayalaseema** | *ajja:* | *na:jina:/ ma: ba:pu* | *na:nna* |
| **Coastal** | *na:jina* | *na:nnaga:ru* | *na:nnaga:ru* |

| **M** *amma:* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *amma/avva* | *amma/talli* | *amma/talli* |
| **Rayalaseema** | *amma* | *amo:v* | *amo:v* |
| **Coastal** | *amma:* | *amma* | *amma* |
| ***Reference Terms*** | | | |
| **Telangana** | *amma/avva* | *amma* | *amma ga:ru* |
| **Rayalaseema** | *amma:* | *amma* | *amma* |
| **Coastal** | *amma:* | *amma* | *amma* |

| **S** *koḍuku* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *By Name/are:j/ ore:* | *By Name/ are:j/ ore:/ po:raga: /pollaga:* | *By Name/are:j/ ore:* |
| **Rayalaseema** | *By Name/ore/ are:j/ abajo:/ abbi* | *By Name/ are:j/ o: abbi /abbo:ḍa:* | *By Name* |
| **Coastal** | *By Name/ are:j* | *By Name/ abba:j* | *By Name /abba:j* |
| ***Reference Terms*** | | | |
| **Telangana** | *ma:vo:ḍu/ koḍuku/ na:biḍḍa /pilaga:ḍu/ poraḍu/ poṭṭega:ḍu* | *By Name/ ma: vo:ḍu/ koḍuku/ pilaga:ḍu/ pora:ga:ḍu* | *By Name/ ma: vo:ḍu/ koḍuku/ ma: abba:ji* |
| **Rayalaseema** | *By Name/ ma: poṭṭega:ḍu:* | *By Name/ma: poṭṭega:ḍu:* | *By Name* |
| **Coastal** | *By Name/ ma: guṇṭaḍu/ ma:kurro:ḍu* | *By Name/ ma: guṇṭaḍu/ ma: ba:bu/ ma: kurro:ḍu* | *By Name/ ma: abba:ji/ ma: pillava:ḍu/ ma: kurro:ḍu* |

In Rayalaseema *ammi/me:j* is used while addressing the girl younger to the ego and similarly *abbi/be:j* is used for addressing the boy.

| **D** *ku:turu* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *By Name /po:ri/ polla/ gada* | *By Name/ polla/ gadda:* | *By Name* |
| **Rayalaseema** | *By Name/ o:mmi/ ammaḍu* | *By Name/ o:mmi/ ammaḍu* | *By Name/ o:mmi/ ammaḍu* |
| **Coastal** | *By Name/ ammi/ pa:pa* | *By Name/ pa:pa/ bujji:* | *By Name/ bujji:/ amma:j* |
| ***Reference Terms*** | | | |
| **Telangana** | *By Name/ ma: amma:j/ ma: pilla/ ma: po:ri/ ma: biḍḍe:/ ma: bo:je* | *By Name/ ma: pilla/ ma: biḍḍa /ma: amma:ji* | *By Name/ ma: amma:ji/ ma: ku:turu/ na: biḍḍa* |
| **Rayalaseema** | *By Name/ ma: ammaḍu/ a: jammi* | *By Name/ ma: ammaḍu/ ma: biḍja/ a: jammi* | *By Name/ ma: ammaḍu/ ma: biḍḍa/ ma: ammaḍu:/ a: jammi* |
| **Coastal** | *By Name/ ma: ammi/ pa:pa/ guṇṭadi* | *By Name/ ma: ammi/ pa:pa/ guṇṭa* | *By Name/ ma: amma:ji/ pa:pa* |

| **H** *bharta/ moguḍu* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *e:majja/ igo: / idigo:/ ma:ma/ e:majja/ ajjo:v* | *e:majja /ba:va/ e:malla* | *e:manḍi/ o:j/ ba:va* |
| **Rayalaseema** | *e:majja /igo:/ idigo:* | *e:majja/ e:mi* | *emanḍi* |
| **Coastal** | *idigo:* | *iṭra:/ igo:* | *e:manḍi* |
| ***Reference Terms*** | | | |
| **Telangana** | *na: moguḍu/ ba:va/penmiṭi* | *ma: a:jana/ na: magaḍu/ penimiṭi* | *ma: va:ru/ ma a:jana* |
| **Rayalaseema** | *a: maniṣi* | *a: maniṣi/ ma: a:jana* | *ma: va:ru/ ma: a:jana* |
| **Coastal** | *ma: vo:ru* | *ma: a:jana* | *ma: va:ru* |

| **W** *bha:rja/ pe[[a:m* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *By Name /e:me:/ ose:/ e:m/ e:mmamma:/ o:mmo:* | *e:me: /ose: / By Name* | *e:me: /By Name* |
| **Rayalaseema** | *e:me:* | *By Name/ e:me:* | *By Name/ e:me:* |
| **Coastal** | *e:me:* | *e: amma:j* | *e: amma:j* |
| ***Reference Terms*** | | | |
| **Telangana** | *na: pella:m/ ma: a:ɖo:llu/ ma: imɳidi/ ma: a:ɖamaniṣi / ma: vo:[[u/ ma:ja:me/ i:me/ ba:ji* | *na: ba:rja/ ma: a:ɖo:llu/ By Name/ i:jamma* | *ma: a:viɖa/ na: pella:m/ na: bha:rja/ By Name* |
| **Rayalaseema** | *By Name/ ma: inɳo:di/ ma: inɳa:me* | *na: penɖ[[a:m* | *na: bha:rja/ ma: a:viɖa* |
| **Coastal** | *pe[[a:m* | *ma: inɳidi* | *amma:j/ ma: a:viɖa/ na: bha:rja* |

| **FFF** *mutta:ta* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *mutta:ta* | *mutta:ta* | *mutta:ta* |
| **Rayalaseema** | *musilo:ɖu* | *ta:ta* | *ta:ta* |
| **Coastal** | *ta:ta* | *ta:ta* | *ta:taga:ru* |
| ***Reference Terms*** | | | |
| **Telangana** | *mutta:ta* | *mutta:ta* | *mutta:ta* |
| **Rayalaseema** | *dže:džĩna:jina* | *dže:džĩna:jina* | *dže:džĩna:jina* |
| **Coastal** | *mutsts̲anna /muttanna* | *muttanna/ mutta:ta* | *mutta:ta* |

| **FeB** *pedda:na:nna* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *pedaba:pu/ pedana:na:nna* | *peddana:na* | *pedana:nnaga:ru* |
| **Rayalaseema** | *peddana:na/ peddajja* | *peddana:na/ peddajja* | *peddana:na* |
| **Coastal** | *peddappa* | *peddana:na* | *peddana:na* |
| ***Reference Terms*** | | | |
| **Telangana** | *peddajja/ pedana:na* | *peddana:na/ peddajja* | *pedana:nnaga:ru* |
| **Rayalaseema** | *peddana:na* | *peddana:na* | *peddana:na* |
| **Coastal** | *peddana:na* | *peddana:na* | *peddana:na* |

| **FyB** *tšinna:nna* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *tšitsts̲a/ ka:ka/ ba:ba:ji/ tšinna:na* | *tšinna:nna* | *ba:ba:ji/ tšinna:na/ tšinna:ba:pu* |
| **Rayalaseema** | *ba:ba:j* | *ba:ba:j* | *ba:ba:j* |
| **Coastal** | *tšiɳ[iba:bu/ tšinna:nna/ dada/ ba:bajja* | *tšiɳ[iba:bu/ ba:ba:j/ tšinna:na* | *ba:ba:j /tšinna:nna* |
| ***Reference Terms*** | | | |
| **Telangana** | *ba:ba:ji/ kakkajja/ tšinnajja/ tšinnappa* | *ba:ba:ji* | *ba:ba:ji* |
| **Rayalaseema** | *ba:ba:j/ tšinnajja* | *ba:ba:j/ tšinnajja* | *ba:ba:j* |
| **Coastal** | *ba:ba:j* | *ba:ba:j* | *ba:ba:j* |

| **FeBW** *peddamma* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *peddamma* | *peddamma/ dobbamma* | *peddamma/ pedda:ji/ peddavva* |
| **Rayalaseema** | *peddamma* | *peddamma* | *peddamma* |
| **Coastal** | *amma/ peddamma/ doɖɖa:mma* | *peddamma/ doɖɖa:mma* | *peddamma* |
| ***Reference Terms*** | | | |
| **Telangana** | *peddamma* | *peddamma* | *Peddamma* |
| **Rayalaseema** | *peddamma* | *peddamma* | *peddamma* |
| **Coastal** | *peddamma* | *peddamma* | *peddamma* |

| **FyBW** *pinni* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *tšinnamma/ tšinnavva/ tšinni* | *tšinnamma/ tšinnavva/ tšinni* | *pinni/ tšinnamma/ pinni* |
| **Rayalaseema** | *tšinnamma/ pinni* | *pinni* | *pinni* |
| **Coastal** | *pinnamma:/ tšinnamma:* | *pinnamma:/ pinni* | *pinnamma: /pinni/ tšinnamma* |
| ***Reference Terms*** | | | |
| **Telangana** | *tšinnamma/ pinni:/ tšinnavva* | *tšinnamma* | *pinni/ tšinnamma/ pinna:j* |
| **Rayalaseema** | *pinni* | *pinni* | *pinni* |
| **Coastal** | *tšinnamma:/ pinni /tšinnamma* | *pinni* | *pinniga:ramḍi* |

| **MB** *me:nama:ma* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *ma:va/ ma:ma* | *ma:ma* | *ma:majja* |
| **Rayalaseema** | *ma:ma* | *ma:ma* | *ma:majja* |
| **Coastal** | *ma:ma* | *ma:ma* | *ma:majja* |
| ***Reference Terms*** | | | |
| **Telangana** | *me:nama:ma* | *me:nama:ma* | *me:nama:ma* |
| **Rayalaseema** | *me:nama:ma* | *me:nama:ma* | *me:nama:ma* |
| **Coastal** | *me:nama:ma* | *me:nama:ma* | *me:nama:ma* |

| **DHSI** *a:ḍabiḍḍa* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *odina* | *odina* | *odinaga:ru* |
| **Rayalaseema** | *odina* | *odina* | *odinaga:ru* |
| **Coastal** | *odina* | *odina* | *odinaga:ru* |
| ***Reference Terms*** | | | |
| **Telangana** | *a:ḍapilla* | *a:ḍabiḍḍa:/ odina* | *a:ḍabiḍḍa:/ vodina* |
| **Rayalaseema** | *a:ḍapilla* | *a:ḍabiḍḍa:/ odina* | *a:ḍabiḍḍa:/ vodina* |
| **Coastal** | *a:ḍapilla* | *a:ḍabiḍḍa:/ odina* | *a:ḍabiḍḍa:/ vodina* |

| **SWM/DHM** *vijjamkura:lu* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *akka:/ vadine* | *akka:/ vadine* | *akka:/ vadine/ amma:ji* |
| **Rayalaseema** | *odine/ akka* | *vodina/ akka* | *vodina/ akka* |
| **Coastal** | *akka:/ vo:dina* | *akkaga:ru/ vodinaga:ru* | *akkaga:ru/ vodinaga:ru* |
| ***Reference Terms*** | | | |
| **Telangana** | *ijjamkura:lu/ ijjalsa:ni/* | *vijjamkura:lu* | *vijjapura:lu/ vijja:mkura:lu/ akka/ vodina* |
| **Rayalaseema** | *bigura:lu/ ijjapura:lu* | *bigura:lu /ijjapura:lu* | *akka/ tšelle/ vo:dina /bigura:lu/ vijjapura:lu* |
| **Coastal** | *ijjapura:lu/ ejjapura:lu* | *vijjapura:lu* | *vijjapura:lu/ vijjampulu* |

| **SWF/DHF** *vijjaMkuḍu* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *ba:va/ anna* | *ba:va/ anna* | *ba:va/ anna* |
| **Rayalaseema** | *ba:va/ anna* | *ba:va/ anna* | *ba:vaga:ru/ annaga:ru* |
| **Coastal** | *ba:va /anna* | *ba:va/ anna* | *ba:vaga:ru/ annaga:ru* |
| ***Reference Terms*** | | | |
| **Telangana** | *ijjamkuḍu/ ijaal ka:pu* | *vijjamkuḍu* | *vijjamkuḍu/ annavga:ru/ ba:vaga:ru* |
| **Rayalaseema** | *bi:guḍu* | *bi:guḍu* | *bi:guḍu/ vijjaMkulu* |
| **Coastal** | *ijjamkuḍu* | *ijjamkuḍu/ ba:vaga:ru* | *ba:vaga:ru/ annaga:ru/ vijjamkulava:ru* |

| **WZH** *saḍḍakuḍu* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *By Name/ tammi/ anna* | *By Name/ tammi/ anna/* | *By Name/ tammi/ anna/ sodara:* |
| **Rayalaseema** | *anna* | *annaga:ru* | *annga:ru* |
| **Coastal** | *anna:/ tammuḍu* | *anna/ tammuḍu* | *annaga:ru/ tamnḍu* |

| **MM** *ammamma* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *ammamma:/ mukkavva:/ mutta/ a:ji/ avva* | *ammamma: /mutta/ a:ji* | *ammamma:* |
| **Rayalaseema** | *ammamma* | *ammamma* | *ammammaga:ru* |
| **Coastal** | *ammamma* | *ammamma* | *ammamma* |
| ***Reference Terms*** | | | |
| **Telangana** | *ammamma/ musaldi* | *ammamma* | *ammamma* |
| **Rayalaseema** | *ammamma* | *ammamma* | |
| **Coastal** | *ammamma* | *ammamma* | *ammammaga:ru* |

| **FF** *nanamma* | SC | *BC* | *FC* |
|---|---|---|---|
| ***Address Terms*** | | | |
| **Telangana** | *na:namma/ na:jinamma* | *nanamma/ na:jinamma* | *na:nammaga:ru* |
| **Rayalaseema** | *na:namma/ na:jinamma* | *na:namma/ na:jinamma* | *na:nammaga:ru* |
| **Coastal** | *na:namma/ na:jinamma* | *na:namma/ na:jinamma* | *na:nammaga:ru* |
| ***Reference Terms*** | | | |
| **Telangana** | *ammamma/ musaldi* | *nanamma/ na:jinamma* | *na:nammaga:ru* |
| **Rayalaseema** | *na:namma/ na:jinamma* | *na:namma/ na:jinamma* | *na:nammaga:ru* |
| **Coastal** | *na:namma/ na:jinamma* | *na:namma/ na:jinamma* | *na:nammaga:ru* |

| **FF/MF** *ta:ta* | SC | *BC* | FC |
|---|---|---|---|
| **Address Terms** | | | |
| **Telangana** | *ta:ta/ musalio:ɖa:* | *ta:ta* | *ta:taga:ru* |
| **Rayalaseema** | *ta:ta* | *ta:ta* | *ta:taga:ru* |
| **Coastal** | *musalina:na/ ajjamma* | *musalina:na* | *ta:ta/ musalina:na* |
| **Reference Terms** | | | |
| **Telangana** | *ta:ta/ musalo:ɖu* | *ta:ta* | *ta:taga:ru* |
| **Rayalaseema** | *ta:ta/ dže:džina:jina/ abba* | *ta:ta/ dže:džina:jina* | *ta:taga:ru/ džʌ:džina:nna* |
| **Coastal** | *ta:ta* | *ta:ta/ muttanna* | *ta:taga:ru/ ta:tajja* |

## III. Conclusion

The use of polite forms for addressing the affinal kin relation is very important in the coastal dialect when compared with the Rayalaseema and Telangana dialect. The breakdown of joint families and the emergence of nuclear families are also having some impact in declining the use of kinship terms for addressing by the children. The prevalent use of uncle and aunty has eaten away all our kinship relations across the families. Similarly *abba:* 'hey' and *adiga:dura*: 'it isn't my dear' are masculine forms used for men but, are used by female members in Telangana area.

In Telugu kinship terms generally replaced by a person name, both as address and reference.


## Acknowledgements

We are grateful to our teacher Prof. B. Ramakrishna Reddy for his comments and suggestions on the earlier version of this paper.

## REFERENCES


Brown, P and Levinson, S. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Brown, R and Gilman, A. 1960. The pronouns of power and solidarity. I*n Style in Language*, Ed. By Thomas Sebeok, 253-276 Cambridge, MA: MIT Press.

Eckert, P and Penelope. 2003. *Language and Gender*. Cambridge: Cambridge University Press.

Eckert, P. and McConnell-Ginet, S. 1998. 'Communities of practice: where language, gender and power all live,? pp.484-494, in Coates, J. ed. *Language and Gender: A Reader,* Oxford, Blackwell.

Eckert, P. and McConnell-Ginet, S. 1999. 'New generalisations and explanations in language and gender research', pp.185-203, in *Language in Society,* 28/2.

Eckert, P. and McConnell-Ginet, S. 2003. *Language and Gender.* Cambridge University Press, Cambridge.

Lakoff, R. 1972, "Language in Context", in *Language*, vol. 48, no. 4, pp. 907—927. Linguistic Society of America.

Lewis Henry Morgan. 1871. Systems of Consanguinity and Affinity of the Human Family. *Smithsonian Contributions to Knowledge*, Vol. XVII. Washington: Smithsonian Institution.

Lounsbury, Floyd G. 1964. "A Formal Account of the Crow- and Omaha-Type Kinship Terminologies", in Ward H. Goodenough (ed.), E*xplorations in Cultural Anthropology: Essays in Honor of George Peter Murdock,* New York: McGraw-Hill, pp. 351–393.

Mills, Sara. 2003. *Gender and Politeness*. Cambridge: Cambridge University Press.

Ryali, Rajagopal. 1970. Semantic Analysis of Telugu Kinship Terminology. Ph.D. dissertation. Duke University.

Spender, Dale, 1980 [1985]. *Man Made Language*, second edition, New York: Routledge.

Tannen, Deborah. 1990. *You just don't understand: women and men in conversation.* New York, NY: Morrow.

Trautmann, Thomas R.1981. *Dravidian Kinship*. Cambridge: Cambridge University Press.

Wenger, Etienne. 1998. *Communities of practice: learning, meaning, and identity*. Cambridge, U.K.: Cambridge University Press.

# INVESTIGATING READING AND WRITING PERFORMANCES AND LANGUAGE USE AT UPPER PRIMARY SCHOOL LEVEL OF ESL STUDENTS WITH TELUGU AS MOTHER TONGUE

**- Nageswara Rao Chelli**

***Abstract***

*This study aimed at investigating the challenges and problems in the development of Reading and writing skills such as Read the text carefully, Read the silent letter words aloud, match the parts of the sentences to make meaningful sentences, and choose the correct synonym of underlined words. Under Writing select the correct choice of the words, select correct preposition from the given in the brackets, formation of plural nouns, select the correct word from the choice given in brackets, and provide the verb form with corresponding sense required (proper use of Grammar laid in the curriculum) in upper primary level students. For the purpose of the study three hundred students from different schools of Andhra and Telangana has been selected. Initially for pilot study the schools of Andhra Pradesh and Telangana areas were listed, within that among these two areas were selected on the basis of the high frequency literacy rate of education competitively to the state literacy rates linearly*

*Andhra 67% and Telangana 64% (Current status), If we see United Andhra Pradesh 67% according to 2011. Based on that ten schools were selected for the study, each study area the study keeping in mind the convenience and administration support. A questionnaire with ten variables was utilized as the instruments of the study. The findings revealed that Telugu ESL students have problem in Reading and writing tasks especially in language use. The study reveals some problems and practical methods in order to cope with Reading and writing difficulties.*

**Keywords**: Reading, Writing, ESL, Curriculum, Performances

## 1.Introduction

The main emphasis of the paper is to understand the English Second Language (ESL) Reading and writing problems and language use at Upper Primary School level. In the recent national curriculum review popularly known as National Curriculum Framework-2005 (NCF) brings out the issue of English language as a subject of study and medium of instruction at 1st class onwards. It further says that "The level of introduction of English has now become a matter of political response to people aspirations rendering almost irrelevant an academic debate on the merits of very early introduction." State of United Andhra Pradesh introduced English subject as second language from 3rd class in the government schools of United Andhra Pradesh from the academic year 2008-2009.

The government of Andhra Pradesh introduced English subject as second language from 1st class in the government schools in the state of United Andhra Pradesh from the academic year 2011-2012. This subject of the study is based on Language as an increasingly important area in applied linguistics is often regarded as a skill rather than knowledge itself. It is generally understood to be a matter of doing than of knowing. While absorbing the mother tongue, the first skill that a child acquires is the ability to understand the spoken word, involving the skill of listening. As a progress, the child tries to reproduce these sound sequences to express his/her own desires and needs and thereby acquires the skill of speaking. Hence, these two basic skills

can be said, constitute one's language ability. On the other hand, though qualified as secondary, it is worth noting that the abilities reading and writing are significant, illustrating matters of literacy.

It is a systematic process of recoding speech sounds through a symbol system (alphabets). It is a learned skill, not an acquired one. It requires training in the art of Reading and writing. In both an individual's life and in the life of mankind, Reading and writing comes after a speech. In human history, these came on later stage while speech was the first medium of communication. Most of us have difficulty in writing because it seems to require more efforts in terms of care, and in terms of thought, than speaking does. Speaking is spontaneous in most cases, whereas Reading and writing always carries with it the notation of correctness of grammar uses, of appropriate expression and comprehension on the reader's side, which are aspects that make it difficult for us in terms of effort and time.

## 2.Investigating Reading and Writing performances and Language Use at Upper Primary School Level

Reading is getting meaning from the printed page. Actually there are no meanings on the printed page. There are only symbols that stand for meanings. Printed symbols are such can only stimulate recall of familiar concepts. New meanings come from the manipulation concepts recalled by the reader. According to the Bond and Thinker (1957), reading involves the recognition of written or printed symbols which serve as stimuli for recall of meanings built up through the readers past experience. New meanings are derived through manipulation of concepts already with reader's possession.

Writing has been regarded as an alternative medium of language, as it gives permanence to utterances. Applied linguistics inherited the view of language as speech and writing as an orthographic. Many people would say that writing is an inaccurate representation of speech. Though there can never be one satisfactory answer for writing, the following points will be provided with some idea. Lado (1971: 222) points out "writing is a graphic representation of a language. Pictures or symbols do not constitute writing unless they form a system representing the units of language and those patterns

can be grasped by the reader". The message is conveyed through the written medium by the use of conventional graphemes. An authentic communication takes place through a universal activity.

Language is a storehouse of knowledge having many dimensions of production and reception, so a standard system is needed to record a language in coded form. Writing is a form of encoded symbols in the form of print or impression.

## 3.Research Questions

Firstly we ask what is the empirical relationship between Telugu ESL students Reading and writing problem with English as a subject? This question allows us to test whether there is more congruence of Telugu as mother tough than English as a subject of learning. We ask: has the congruence between ESL students Reading and writing problems and origins of Telugu students in increased more in rural areas than urban areas? Finally, to analyse the interaction of ESL students and origins of government upper primary schools with reference to the parental education and income on specific outcomes of students' performance in Reading and writing skills, We ask do Telugu ESL students find it harder to learn English as a subject at upper primary level in government run schools.

## 4.Data and Measurement

For the final try out of the study three hundred students from different schools of Andhra Pradesh and Telangana are selected. Initially the schools of the Andhra Pradesh and Telangana areas were listed, within that among these Two areas were selected randomly for pilot study on the basis of the high frequency literacy rate of education competitively to the states literacy rate linearly 67% and 64% (United Andhra Pradesh 2011). Based on that ten schools were selected for the study, each study area the study keeping in mind the convenience and administration support.

## 5. Methods

In the study multi stage cluster sampling approach[1] has been adopted. In the first stage of sampling selection is based on schools located in urban and rural with high proficiency rate of English

language. To ensure there would be enough students for the each type of school to collect data and compare by groups, all public upper primary schools of Andhra and Telangana were divided into two groups. Schools were selected ascending and descending order. In the Second stage a structured questionnaire has been distributed to the students to collect data related to English Reading and writing skills. Then a sample of students in each selected section has been chosen to conduct written test by applying random sampling method. Under Reading and writing, there are five indicators are used for understanding their ability of Reading and writing process. Further, the researcher also makes clear picture between rural and urban school students' performance from the selected indicators.

## 6.Results and Analysis

We employ two techniques to analyze classified tables to study the association between student in schools and parental background in learning English as a subject at government run Upper Primary Schools in Andhra and Telangana. To do this systematically to address the problems of Telugu ESL Reading and writing performance, the researcher carefully designed his questionnaire and followed two methods to obtain information. For the purpose schematic representation of the reading and Writing tests was conducted by using the questionnaire method. The following table gives you a clear picture of schematic representation of reading and writing skills of students.

**Table 1.1** A schematic representation of the reading tests used in the questionnaire to obtain the students' reading performance through the following indicators is given below.

| **Tools used** | **Purpose served** |
|---|---|
| 1.Read the text carefully | Fluency and accuracy test skill and Exactness |
| 2. Read the words aloud | The ability to spell words correctly |
| 3. Complete the sentence by given options | The ability to comprehension |
| 4. Match the parts of sentences to make meaningful sentences | The ability to make complete sense |
| 5. Choose the correct synonym of underlined words | The ability to locate similarity |

**Table 1.2** A schematic representation of the writing tests used in the questionnaire, to obtain the students' writing performance through following indicators

| | |
|---|---|
| 1. select the correct choice of article | The ability to comprehend the Meaning of the test |
| 2. Spelling test | The ability to fill the Missing letters |
| 3. Opposite word test | The ability to write Opposite word |
| 4. Sentence comprehensive test | The ability to arranges the words it to meaningful sentence |
| 5. Test of abstract | The ability recognizes the miss-fit words logically |

**Source: Field Data Repots**

**TABLE 1-3 Class wise Distribution of the Students**

| Age | Class wise Distribution of Students | | | | | | Total |
|---|---|---|---|---|---|---|---|
| | 5th class Rural | 5th class Urban | 6th class Rural | 6th class Urban | 7th class Rural | 7th class Urban | |
| 10 | 50 | 50 | 0 | 0 | 0 | 0 | 100 |
| 11 | 0 | 0 | 50 | 50 | 1 | 1 | 101 |
| 12 | 0 | 0 | 1 | 0 | 30 | 33 | 64 |
| 13 | 0 | 0 | 0 | 0 | 19 | 16 | 35 |
| **Total** | 50 | 50 | 50 | 50 | 50 | 50 | 300 |

Source: Based on Field Study

The Table, 5-1 displays the area, age and the number of students involved in the research study. This study was conducted in different schools situated in urban and rural regions AP and Telangana. The samples of the students were the same in both the areas and ten students were randomly selected from each class of upper primary schools of Andhra Pradesh and Telangana.

The above table displays the age range of the students according to their classes. It can be observed that the total number of students is same in 5th, 6th, and 7th classes. However, the age range of students is different from each class and within the class.

**TABLE 1-4 Gender Wise Distribution of the Students**

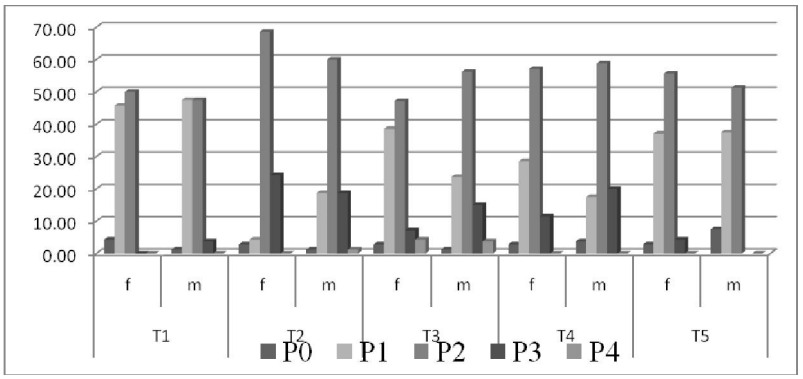
| Area | Gender | | Total |
|---|---|---|---|
| | Female | Male | |
| 5th class Rural | 30 | 20 | 50 |
| 5th class Urban | 20 | 30 | 50 |
| 6th class Rural | 30 | 20 | 50 |
| 6th class Urban | 25 | 25 | 50 |
| 7th class Rural | 25 | 25 | 50 |
| 7th class Urban | 25 | 25 | 50 |
| Total | 155 | 145 | 300 |

Source: Based on pilot Study

Table 5-2 displays the distribution of students of different classes and their participation in the investigation/ research. The above table shows that an equal number of the students from rural and urban, where 145 boys and 155 girls have participated in this research. The distribution is maintained in accordance with the gender aspect and in rural-urban ratio, as shown strictly by the numbers: 150 students from rural and 150 students from urban randomly selected for the study.

**TABLE 1-5 Rural Writing Test Scores in all Classes**

| Rural writing test performance in all classes | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender/ Exercise | | 0 | % | 1 | % | 2 | % | 3 | % | 4 | % | Total |
| Test-1 | f | 1 | 1.18 | 34 | 40.00 | 37 | 43.53 | 8 | 9.41 | 5 | 5.88 | 85 100 |
| | m | 5 | 7.69 | 28 | 43.08 | 19 | 29.23 | 9 | 13.85 | 4 | 6.15 | 65 100 |
| Test-2 | f | 14 | 16.47 | 39 | 45.88 | 28 | 32.94 | 4 | 4.71 | 0 | 0.00 | 85 100 |
| | m | 5 | 7.69 | 28 | 43.08 | 25 | 38.46 | 5 | 7.69 | 2 | 3.08 | 65 100 |
| Test-3 | f | 8 | 9.41 | 26 | 30.59 | 39 | 45.88 | 8 | 9.41 | 4 | 4.71 | 85 100 |
| | m | 3 | 4.62 | 19 | 29.23 | 31 | 47.69 | 8 | 12.31 | 4 | 6.15 | 65 100 |
| Test-4 | f | 3 | 3.53 | 46 | 54.12 | 36 | 42.35 | 0 | 0.00 | 0 | 0.00 | 85 100 |
| | m | 2 | 3.08 | 28 | 43.08 | 31 | 47.69 | 3 | 4.62 | 1 | 1.54 | 65 100 |
| Test-5 | f | 2 | 2.35 | 29 | 34.12 | 34 | 40.00 | 14 | 16.47 | 6 | 7.06 | 85 100 |
| | m | 1 | 1.54 | 20 | 30.77 | 30 | 46.15 | 9 | 13.85 | 5 | 7.69 | 65 100 |




Histogram 1-5

As we observe the above table (1-5) display the rural students writing test scores which are four sub-writing tests include boys and girls. Out of (150)them 1.18% girls and 7.69% boys scored zero, 40% girls and 43.8% boys scored one, 43.53% girls and 29.23% boys scored 2 and 9.41% girls and 13.85% boys scored 3, 5.8% girls and 6.15% boys scored 4 in the test. Out of 150 students 85 (56.6%) girls and 65 (43.3%) boys scored marks in the writing test-1. Comparatively girls scored higher than boys.

Writing test – II of rural students secured scores of boys and girls. Out of them 16.47% girls and 7.69% boys scored zero, 45.88% girls and 43.08% boys scored one, 32.94% girls and 38.46% boys scored two, 4.71% girls and 7.69% boys scored three marks, 0% girls and 3.8% boys scored four in the test. Out of 150 students 85 (56.6%) girls and 65 boys (43.3%) secured marks in the writing test-2. Comparatively girls secured higher than boys.

Writing test - III scores of test include boys and girls. Out of them 9.41% girls and 4.62% boys scored zero, 30.59% girls and 29.23% boys scored one, 45.88% girls and 47.69% boys scored two, 9.41% girls and 12.31% boys scored three, 4.71% girls and 6.15% boys scored four marks in the test. Out of 150 students 85 (56.6%) girls and 65 (43.3%) boys scored marks in the writing test. Comparatively girls scored higher than boys.

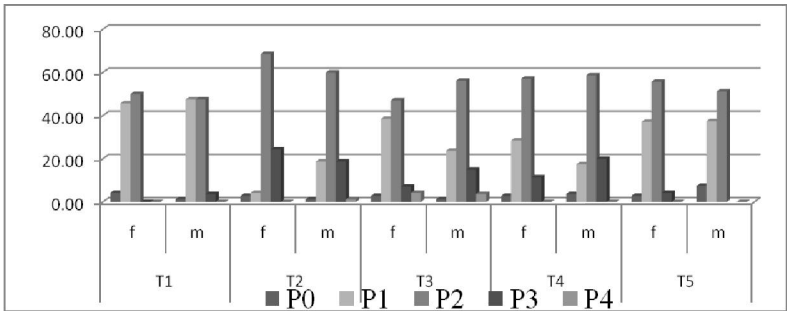
Writing test - IV scores include boys and girls. Out of them 3.53% girls and 3.08% boys scored zero mark, 54.12% girls and 43.08% boys scored one mark, 42.35% girls and 47.69% boys scored two marks, 0% girls and 4.62% boys scored three marks, 0% girls and 1.54% boys scored four marks in the test. Finally, Out of 150 students 85 (56.6%) girls and 65 (43.3%) boys scored marks in the writing test. Comparatively girls scored higher than boys.

Writing test - V test scores include girls and boys. Out of them 2.35% girls and 1.54% boys scored zero, 34.12% girls and 30.77% boys scored one, 40% girls and 46.15% boys scored two, 16.47% girls and 13.85% boys scored three, 7.06% girls and 7.69% boys scored four in the test. Finally, Out of 150 students 85 (56.6%) girls and 65(43.3%) boys scored marks in the writing test-5. Comparatively, girls scored higher than boys.

**TABLE 1-6 Urban Writing Test Scores in all Classes**

| Urban writing test performance in all classes | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender/EXC | | 0 | % | 1 | % | 2 | % | 3 | % | 4 | % | Total | % |
| Test-1 | f | 1 | 1.43 | 2 | 2.86 | 47 | 67.14 | 17 | 24.29 | 3 | 4.29 | 70 | 100 |
| | m | 1 | 1.25 | 11 | 13.75 | 49 | 61.25 | 11 | 13.75 | 8 | 10.00 | 80 | 100 |
| Test-2 | f | 0 | 0.00 | 19 | 27.14 | 35 | 50.00 | 16 | 22.86 | 0 | 0.00 | 70 | 100 |
| | m | 1 | 1.25 | 19 | 23.75 | 43 | 53.75 | 15 | 18.75 | 2 | 2.50 | 80 | 100 |
| Test-3 | f | 3 | 4.29 | 12 | 17.14 | 38 | 54.29 | 12 | 17.14 | 5 | 7.14 | 70 | 100 |
| | m | 0 | 0.00 | 10 | 12.50 | 50 | 62.50 | 16 | 20.00 | 4 | 5.00 | 80 | 100 |
| Test-4 | f | 1 | 1.43 | 16 | 22.86 | 47 | 67.14 | 6 | 8.57 | 0 | 0.00 | 70 | 100 |
| | m | 0 | 0.00 | 18 | 22.50 | 53 | 66.25 | 8 | 10.00 | 1 | 1.25 | 80 | 100 |
| Test-5 nt | f | 1 | 1.43 | 12 | 17.14 | 30 | 42.86 | 19 | 27.14 | 8 | 11.43 | 70 | 100 |
| | m | 0 | 0.00 | 21 | 26.25 | 37 | 46.25 | 18 | 22.50 | 4 | 5.00 | 80 | 100 |

Source: Based on Field Study




Histogram 1-6

As we observe above tables 1.6 display the urban students writing test scores include boys and girls. Out of them 1.43% girls and 1.25% boys scored zero, 2.86% girls and 13.75% boys scored one, 67.14% girls and 61.25% boys scored two, 24.29% girls and 13.75% boys scored three, 4.29% girls and 10% boys scored four in the test. Finally, Out of 150 respondents' girls 70 (46.6%) and boys 80 (53.3%) scored marks in the writing test-1. Comparatively boys scored higher than girls.

Writing test-II, where the urban respondents writing test scores of second test include boys and girls. Out of them 0% girls and 1.25% boys scored zero, 27.14% girls and 23.75% boys scored one, 50% girls and 53% boys scored two, 22.86% girls and 18.75% boys scored three, 0% girls and 2.50% boys scored four marks in the test. The respondents among 150, 70(46.6%) girls and 80 (53.3%) boys scored marks linearly in the second test. Comparatively in the second test boys scored higher than girls.

Writing test-III, where urban respondents scored marks in percentage among girls and boys. Out of them 4.29% girls and 0% boys scored zero, 17.14% girls and 12.50% boys scored one, 50.29% girls and 62.50% boys scored two, 17.14% girls and 20% boys scored three, 7.14% girls and boys 5% girls scored four in the test. The respondents among 70 (46.6%) girls 80 (53.3%) boys scored marks in third test. Comparatively in third test boys scored higher than girls.

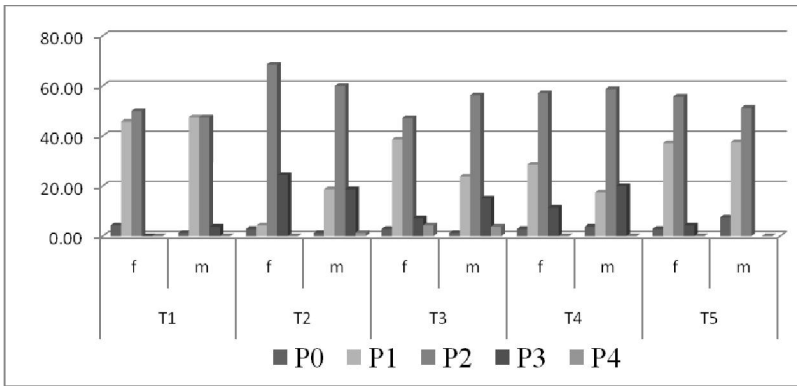
Writing test-IV, where urban respondents scored marks in percentage among girls and boys. Out of them 1.43% girls and 0% boys scored zero, 22.86% girls and 22.50% boys scored one, 67.14% girls and 66.25% boys scored two, 8.57% girls and 10% boys scored three, 0% girls and 1.25% boys scored four in the test. The respondents among 70 (46.6%) girls and 80 (53.3%) boys scored marks in the fourth tests. Comparatively in fourth test boys scored higher than girls.

Writing test-V, the urban respondents scored marks in the reading tests percentage include girls and boys. Out of them 1.43% girls and 0% boys scored zero, 17.14% girls and 26.25% boys scored one, 42.86% girls and 46.25% boys scored two, 27.14% girls and 22.50% boys scored three, 11.43% girls and 5% boys scored four. Finally, Out of 150 respondents' girls 70 (46.6%) and boys 80 (53.3%) scored marks in the writing test-V. Comparatively, boys scored higher than girls.

**TABLE 1-7 Rural Reading test performance in all classes**

| Rural Reading test Scores in all classes | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender/EXC | | 0 | % | 1 | % | 2 | % | 3 | % | 4 | % | Total | % |
| Test-1 | f | 19 | 22.35 | 42 | 49.41 | 24 | 28.24 | 0 | 0.00 | | 0.00 | 85 | 100.00 |
| | m | 11 | 16.92 | 34 | 52.31 | 18 | 27.69 | 2 | 3.08 | | 0.00 | 65 | 100.00 |
| Test-2 | f | 9 | 10.59 | 29 | 34.12 | 37 | 43.53 | 10 | 11.76 | 0 | 0.00 | 85 | 100.00 |
| | m | 7 | 10.77 | 17 | 26.15 | 32 | 49.23 | 7 | 10.77 | 2 | 3.08 | 65 | 100.00 |
| Test-3 | f | 8 | 9.41 | 48 | 56.47 | 23 | 27.06 | 3 | 3.53 | 3 | 3.53 | 85 | 100.00 |
| | m | 4 | 6.15 | 27 | 41.54 | 29 | 44.62 | 4 | 6.15 | 1 | 1.54 | 65 | 100.00 |
| Test-4 | f | 7 | 8.24 | 47 | 55.29 | 26 | 30.59 | 5 | 5.88 | | 0.00 | 85 | 100.00 |
| | m | 5 | 7.69 | 36 | 55.38 | 16 | 24.62 | 8 | 12.31 | | 0.00 | 65 | 100.00 |
| Test-5 | f | 15 | 17.65 | 41 | 48.24 | 29 | 34.12 | 0 | 0.00 | | 0.00 | 85 | 100.00 |
| | m | 10 | 15.38 | 35 | 53.85 | 17 | 26.15 | 3 | 4.62 | | 0.00 | 65 | 100.00 |

Source: Based on Field Study




Histogram 1-7

As we observe, Tables from 1.7 display the rural respondents reading test scores include boys and girls. Out them 22.35% girls and 16.92% boys scored zero, 49.41% girls and 52.31% boys scored one, 28.24% girls and 27.69% boys scored two marks, 0% girls and 3.08% boys scored three. No girls and boys were able to score four marks in the test. Finally, Out of 150 students' 85 (56.6%) girls and 65 (43.3%) scored marks in the reading test-I. Comparatively girls scored higher than boys.

Reading test-II displays students' scored marks in percentage include girls and boys. Out of them 10.59% girls and 10.77% boys scored zero, 34.2% girls and 26.15% boys scored one, 43.53% girls and 49.23% boys scored two, 11.76% girls and 10.77% boys scored three, 0% girls and 3.8% boys scored four in the test. Out of 150 students' 85 (56.6%) girls and 65 (43.3%) scored marks in the reading test-II. Comparatively, girls scored higher than boys.

Reading test-III, where students scored marks in percentage include girls and boys. Out of them 9.41% girls and 6.15% boys scored zero, 56.48% girls and 41.54% boys scored one, 27.06% girls and 44.62% boys scored two, 3.53% girls and 1.54% boys scored three, 3.53% girls and 1.54% boys scored four in the test. Out of 150 students' 85 (56.6%) girls and 65 (43.3%) scored marks in the reading test-III. Comparatively, girls scored higher than boys.

Reading test-IV, where students scored marks in percentage include girls and boys. Out them 8.24% girls and 7.69% boys scored

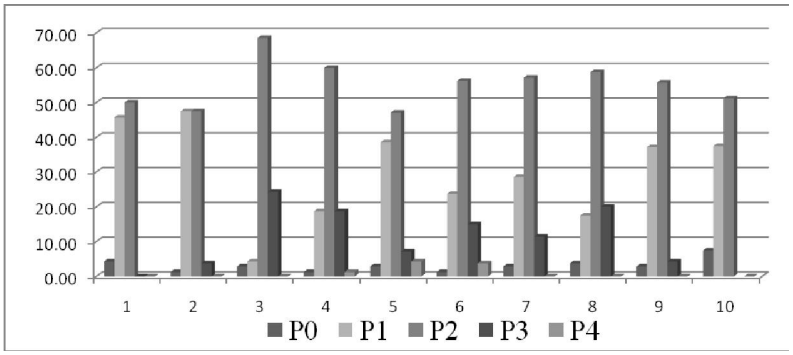
zero, 55.29% girls and 55.38% boys scored one, 30.59% girls and 24.62% boys scored two, 5.88% girls and 12.31% boys scored three marks. However, no students' were able to score four in the test. Finally, out of 150 students' 85(56.6%) girls and 65(43.3%) scored marks in the reading test-IV. Comparatively, girls scored higher than boys.

Reading test-V, the students scored marks in percentage include girls and boys. Out of them 17.65% girls and 15.38% boys scored zero, 48.24% girls and 53.85% boys scored one, 34.12% girls and 26.15% boys scored two, 0% girls and 4.62% boys scored three in the test. However, no students were able to score four marks in the test. Finally, out of 150 students' 85 (56.6%) girls and 65 (43.3%) scored marks in the reading test-V. Comparatively girl students scored higher than boy students.

**TABLE 1-8 Urban Reading Test Scores of Marks in all Classes**

| Urban Reading test performance in all classes | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender/EXC | | 0 | % | 1 | % | 2 | % | 3 | % | 4 | % | Total | % |
| Test-1 | f | 3 | 4.29 | 32 | 45.71 | 35 | 50.00 | 0 | 0.00 | | 0.00 | 70 | 100.00 |
| | m | 1 | 1.25 | 38 | 47.50 | 38 | 47.50 | 3 | 3.75 | | 0.00 | 80 | 100.00 |
| T-2 | f | 2 | 2.86 | 3 | 4.29 | 48 | 68.57 | 17 | 24.29 | 0 | 0.00 | 70 | 100.00 |
| | m | 1 | 1.25 | 15 | 18.75 | 48 | 60.00 | 15 | 18.75 | 1 | 1.25 | 80 | 100.00 |
| T-3 | f | 2 | 2.86 | 27 | 38.57 | 33 | 47.14 | 5 | 7.14 | 3 | 4.29 | 70 | 100.00 |
| | m | 1 | 1.25 | 19 | 23.75 | 45 | 56.25 | 12 | 15.00 | 3 | 3.75 | 80 | 100.00 |
| T-4 | f | 2 | 2.86 | 20 | 28.57 | 40 | 57.14 | 8 | 11.43 | | 0.00 | 70 | 100.00 |
| | m | 3 | 3.75 | 14 | 17.50 | 47 | 58.75 | 16 | 20.00 | | 0.00 | 80 | 100.00 |
| T-5 | f | 2 | 2.86 | 26 | 37.14 | 39 | 55.7 | 3 | 4.29 | | 0.00 | 70 | 100.00 |
| | m | 6 | 7.50 | 30 | 37.50 | 41 | 51.25 | 3 | 3.75 | | 0.00 | 80 | 100.00 |

Source: Based on Field Study




Histogram 1-8

As we observe the above table 1.8 display the urban students reading test scores. Four sub-reading tests include girls and boys. Out of them 4.29% girls and 1.25% boys scored zero, 45.71% girls and 47.50% boys scored one, 50% girls and 47.50% boys scored two, 0% girls and 3.75% boys scored three in the test. However, no students were able to score four marks in the current test. Finally, out of 150 students' 70 (46.6%) girls and 80 (53.3%) boys scored marks in different percentages. Comparatively boys scored higher than girls.

Reading test- II, where respondents scored marks in percent include girls and boys. Out of them 2.86% girls and 1.25% boys scored zero, 4.29% girls and 18.75% boys scored one, 68.57% girls and 60% boys scored two, 24.49% girls and 18.75% boys scored three, 0% girls and 1.25% boys scored four in the test. Finally, out of 150 students' 70 (46.6%) girls and 80 (53.3%) boys scored marks in the reading test-II. Comparatively, boys scored higher than girls.

Reading test- III, where respondents scored marks in percent among girls and boys. Out of them 2.86% girls and 1.25% boys scored zero, 38.57% girls and 23.75% boys scored one, 47.14% girls and 56.25% boys scored two, 17.14% girls and 15% boys scored three, 4.29% girls and 3.75% boys scored four in the test. Finally, out of 150 students' 70 (46.6%) girls and 80 (53.3%) boys scored marks in the reading test-III. Comparatively, boys scored higher than girls.

Reading test - IV, where respondents scored marks in percentage include girls and boys. Out of them 2.86% girls and 3.75% boys scored zero, 28.57% girls and 17.50% boys scored one, 57.14% girls and 58.75% boys scored two, 11.43% girls and 20% boys scored three in the test. However, no respondents were able to score four in the test. Finally, out of 150 students' 70 (46.6%) girls and 80 (53.3%) boys scored marks in the reading test-IV. Comparatively, boys scored higher than girls.

Reading test -V, where respondents scored marks in percent include girls and boys. Out of them 2.86% girls and 7.50% boys scored zero, 37.14% girls and 37.50% boys scored one, 55.7% girls and 51.25% boys scored two, 4.29% girls and e 0% boys scored

three in the test. However, no respondents were able to score four in the test. Finally, out of 150 students' 70 (46.6%) girls and 80 (53.3%) boys scored marks in the reading test-4. Comparatively, boys scored higher than girls

**Conclusion**

The above results are computed based on the data collected in terms of rural and urban, girls and boys, the performance is shown in the tables based on the data findings. The study is based on case study for which survey and questionnaires have been used as data gathering techniques. The study is conducted to cover 300 students in AP and Telangana. During the study, different units i.e., Rural and Urban, Gender, and students class-wise performance have been examined to know or investigate Language 'Problems of Reading and writing skills in English among the students of upper primary classes'.

It is known earlier that, not much emphasis was laid upon studies like, the performance of English language skills in areas like the Government schools. This study involves government schools both in urban and rural areas of A.P. The performance shown by the students depicted by the data was quite interesting and demands explanation. When we skim through the overall responses it is found that the rural respondents have scored 0, 1 marks, which is the highest among them, whereas the urban respondents have scored 2, 3 marks, which is high in percent, compared to rural. However the overall percentage of marks scored by the respondents from both the areas is low in percentage.

According to the data, reading and writing skills of Telugu ESL students in government run Upper Primary Schools are gaining class-by-class momentum slowly.

When it comes to the performance between government schools in urban and rural selected areas of A.P and Telangana the performance shown by the data was quite competitive and overall performance of the both the areas is very low percentage. The performance of the rural students competing to get zero marks. The zero mark secured rural students' are very high in both girls and

boys. The findings of the study show that the zero marks secured students are very less in urban locality than rural.

The performance of the rural students one mark secured students' are better than urban, and urban students' performance is better than rural in 2, 3 marks and hardly 4. Significantly urban students' performance is better than rural, they might be due to better facilities in urban schools, getting inspiration by neighbor students, awareness of the parents and their better environments.

There was no significant difference between boys and girls, and rural and urban in the mean below average performance of reading and writing. If we observe the data in certain areas in reading namely reading the test, match the words, and circle the misfit word, and simultaneously. In writing test there was close interrelation among vocabulary, those filling the gaps, and missing letters, in areas the students' performance is average.

The grade (5th, 6th, and 7th) differences were significant with respect to reading and writing tests of students. The test score increased with the increase of the grades level. The students of higher classes' students performed better than the students of the lower classes. This is due to the number of years in schooling and the increasing difference in maturity level. Positive relationship between tests of the students and the type of area they lived in.

The performance of the pupil who lived in an urban area with surrounding environments and facilities available in schools was better understanding leads to better performance of the students. But final results are rural students overall performance is below average and urban students performance is quite average.

**Reference**

Bloomfield, L. 1993 *Language,* George Allen of Unwin Ltd: New York

Byrne, D. (1998) *Teaching Writing skills, London Longman*: Oxford University Press

Chaudron, C (1988) Second Language class room: Research on Teaching and Learning, Newyork:CUP

Cooper, J (1979) Think and link an Advanced course in Reading and Writing Skills. Malaya: University of Malaya.

Daly, J. A (1978) "Writing apprehension and writing competency". Journal of ducational Research. 72: 10-14.

Peregoy, S., & Boyle, O. (2001). Reading, writing, and learning in ESL: A resource book for K–12teachers. New York, NY: Longman.

Richards, J. C., & Rodgers, T. S. (2001). Approaches and methods in language teaching (2nd ed.). New York, NY: Cambridge University Press.

Sharma, J.C. Multilingualism in India. Language in India. Vol: 1:6 October 2001 www.languageinIndia

Shashikala. A (1995) A Study of fostering writing skills and Ability for creative Writing in Kannada among the students of Ninth Grade. (Unpublished Theses, University of Mysore).

Tan, E.K., & Miller, J. (2008). Writing in English in Malaysian High Schools: The Discourse of Examinations, England: Routledge

UNESCO (2003) Education in a multilingual world. UNESCO Education Position Paper. Paris. www.worldbank.org/ieg/education (accessed on 09.05.08)

**(Footnotes)**

[1] Cluster sampling is a sampling technique used when natural groupings are evident in a statistical population. In this technique, the total population divided into these cluster/groups and a sample of the groups is selected. The cluster should be jointly exclusive and collectively full. In single stage cluster sampling, all the elements from each of the selected cluster are used. In multi-stage cluster sampling, a random sampling technique is applied to the elements from th e each of the selected cluster

# PRESERVATION AND PROPAGATION OF MUSICAL LITERATURE ON COPPER PLATES

***- Dr. Veturi Anandamurthy***

The present endeavour is on aspects of preservation and propagation of musical literature engraved on copper plates during the first half of the 16th century. Indeed, the discovery of these copper plates from inside the temple cellar at Tirumala by the temple authorities at the beginning of the 19th century and the rediscovery of their real worth, extent and content in the year 1947, by the indefatigable endeavours of my revered father, the Late Scholar Poet PrabhakaraSastri, opened new vistas in the fields of literary and linguistic research and musicology.

These copper plates as available today, are a little less than 3000, contain inscribed songs in Telugu – ranging around 18000 – on an average of six per plate, though in fact we learn with regret that what remains today forms only a negligible part of the numbers originally inscribed. What we gather from internal evidences, we have to assume that we have lost nearly 5000 plates containing some 30,000 songs on an average. Ever since PrabhakaraSastri initiated systematic study in this field in 1947 after classifying these copper plates, publication of some 27 Volumes has so far been accomplished by the Tirumala Tirupati Devasthanams, Tirupathi.

Before going into statistical details of these copper plates it will be relevant to briefly sketch the family antecedents of the Tallapaka Poets who flourished during the reign of three dynastic rulers (namely

the Saluva, Tuluva and the Aravidu dynasties) of the Vijayanagar lineage. *Padakavita Pitamaha Annamayya*, the progenitor of the Tallapaka family of Sankirtana Poets, a senior contemporary of the King SaluvaNarasimharaya and the popular Karnataka HaridasaPurandaradasa, was born in 1424 AD. Owing to a divine vision and command in his sixteenth year he started composing songs in the name of Sri Venkateswara, the Lord of the Seven Hills. Since that time he composed at the rate of not less than one song a day till his death in 1503 A.D - all put together exceeding 32,000 songs in his life time. The unique feature with this family of Poets was that not only Annamayya, but also his sons and grandsons made valuable contributions by continuing the Sankirtana tradition with equal zeal, besides creating an astounding variety of other literary forms and enriching the contemporary literature.

TallapakaPedaTirumalayya, second son of Annamayya and China Tirumalayya the grandson are the other two in the family who also composed *padams* like their progenitor from their sixteenth year enjoying a full span of their lives. We are, however, unable to assess today, the actual number of compositions of these two Poets in the same way as we could know, with certainty, about the numbers composed by Annamayya from his biography in metre, which indicated that he alone composed more than 32,000 songs during his lifetime. It may not be a wild guess to assume that the compositions of the three Poets put together would have crossed a hundred thousand. There is a traditional account recorded in Ekamranatha Charita relating to Krishnamacharya of the Kakatiya period who is believed to have composed *Chaturlaksha Sankirtans* i.e. 4 lakhs. Tradition has it that he got these compositions engraved on copper plates though unfortunately, we do not see today, even a single copper plate of this Poet and could not recover or trace more than a hundred of his Talagandhi Vachanas from all the Palm Leaf manuscripts available in the state.

If the tradition is founded on fact, Saint composer Krishnamacharya of the 13$^{th}$ century, stationed at Simhachalam, the Vishnushrine found in North Andhra region, would be the first composer to have attempted to inscribe devotional musical literature

on copper plates. If that was not so, then, what happened perhaps could be that the episode of inscribing the *Sankirtanas* of Annamacharya on copper plates might have been superimposed at a later period on the life story of the earlier Krishnamacharya. However, these are the references of the earliest records we have of musical literature inscribed on copper plates. But for these, all the other copper plates we usually come across contain invariably inscriptional content of either donative or commemorative nature with occasional versification but certainly not containing exclusive literary masterpieces. The same was the case with stone edicts with the exception, of course, of what were found in the Omkara-Amareswara Jyothirlinga Kshetra in Nemad district of erstwhile Central Provinces in which literary compositions like the famous *Mahimnastava* and the *Halayudhastava* were inscribed during the 11th century, on a huge stone slab near the Mandhatrumantap in the shrine. These were perhaps the earliest records we have on lithic medium of exclusive devotional literature composed in *Vritta Chandas*.

Next to these are only the huge (7' x 4') wall stone slabs of the Vijayanagar period found at Tirumala on which were inscribed some Sanskrit *kirtans* of the Tallapaka Poets, together with the musical score and notation. These are yet to be fully deciphered and published. I have with me the dia-positive transparencies of these two chiselled slabs. It would be a wonderful contribution to musicology if these are deciphered and published. Reverting to the subject of musical compositions of the Tallapaka Poets on copper plates, it should be said that they are unique and rare in every respect. All these copper plates were preserved for a long time in a temple cellar which was popularly known as "Tallapakavarikottu" and which was also labelled by the Poets themselves as the *'Sankirtana Bhandagara'*. We learn from inscriptional evidence that this *Sankirtana Bhandagara* was established during the time of Annamayya, first as a mere repository of stylus worked Palm leaf Manuscripts. Annamayya in one of his lyrics sung thus "One song loaded with devotion is enough to protect me forever. Let the rest be at the repository". Obviously there existed the nucleus structure of this *Bhandagara* wherein piles of Palm leaf manuscripts were preserved.

Later when time was propitious Annamcharya's son PedaTirumalayya and his grandson China Tirumalayya seemed to have organised the transcription of all these manuscripts again on copper plates with the intention of making the exercise immortal for the dual purpose of preservation and propagation. In this exercise, Trial and error method eventually proved successful. After working at it for over 20 years (between 1525 and 1545) utilising tonnes of copper metal and employing more than half a dozen scribes skilled in the art of wielding the stylus on treated copper, the Tallapaka Poets finally achieved what they wanted. In the process, as we can perceive, the problems faced by them were many, not only in view of the massive nature of these compositions but from the point of view of classification of the plates content wise – depending on the erotic or philosophical content of the songs, or on the basis of authorship as to whether the songs were of Annamayya, PedaTirumalayya or China Tirumalayya. In fact they appeared to have struggled to devise ways and means to solve these problems, but, finally succeeded in differentiating the plates by bringing in variations in their sizes and thickness and also by other means like numbering the plates, labelling them by the author's name – initials and also by the round or the square shaped punch marks given on the plates for securing them together as we do in the case of library index cards today. In the first instance they tried to engrave the songs on lengthy plates (L.P.) of 33" x 2.5" size of comparatively minimal thickness in an attempt to copy the large Palm leaves; but evidently discontinued the programme after some time when they found them inconvenient on several counts. As a result, some 120 full plus 5 broken pieces of these LPs are available to us today. But later they seem to have tried the following variation to distinguish the author and the content as well.

| Authors's Label | Content | Size | Recovered Plates |
|---|---|---|---|
| A.ca | Erotic | 15.5" x 7" | 2002 |
| A.ca | Philosophical | 15.5" x 7" | 391 |
| Pe-Ti-ca | Erotic | 16" x 7" | 100 |
| Pe-Ti-ca | Philosophical | 15.75" x 6.75" | 77 |
| Ci-Ti-ca | Erotic | 15.5" x 7.2" | 10 |
| Ci-Ti-ca | Philosophical | 15" x 7.5" | 10 |
| Total Number of Plates | | | 2590 |

Considering an average of 6 songs per plate, the total number of songs would be 2590 x 6 = 15,540.

TheTallapaka poets seem to have standardised the size of 15" or 16" x 6.5" or 7.5" approximately as they have chosen only this standard size to inscribe several other minor works like the *SringaraManjari, Vairagya Vachana Malika geetalu, Ashtabhasha-dandakamu* etc, including Sankirtana Lakshanamu the grammar of music. The exact date of inscribing the *Ashtabhasha-dandakamu* is recorded on the copper plate as 'HevilambisamvatsaraMargasirasuddha Panchami Budhavaram' corresponding to 7-11-1537. (These standard plates are labelled as – S.Ps).

Besides these standard size copper plates discovered in the temple cellar at Tirumala, several big size plates (labelled as BPs) of dimension 27" x 16" and 36" x 15" were also observed at Vaishnava shrines like Ahobhilam and Srirangam. Obviously, they got these plates duplicated thus in those sizes for the exclusive purpose of propagation and easy transportation.

The nucleus structure of the *Sankirtana Bhandagara*, which was established as just a repository of MSS and copper plates containing Sankirtanas, later developed into a Religious-cum-Socio-cultural institution which catered to the needs and growing demands of the period by employing musicians and dancers and also by providing training facilities to beginners. *"SankirtanaArulappadu"* or the worship of the Sankirtanas at several Vishnu shrines became a part of the temple festivities. Consequent to this development in the organisation the Sankirtana poets had to prepare copies of copper plates in big sizes which were easy to transport through human agency to various inaccessible hill shrines and other places throughout South India like Simhachalam, Ahobhilam, Mangalagiri, Kanchi, Srirrangam etc. Hence the need for the big plates. Five such large sized plates inserted into a sealed ring on the top supported by a thick Banyan stalk pole called the 'Tandu' passing through the ring, used to be carried on human shoulders. Many of these plates were either lost with the passage of time or got converted into water containers and ablution vessels in the temples. In 1949 at Ahobhilam 35 large sized plates were procured and in 1962 another set of 40 plates from Srirangam were secured by the TTD.

The Madras epigraphical report which records in 1915, the existence of such plates at Ahobhilam reads thus – "Many huge inscribed copper plates are kept in the underground cellars in the temple at upper Ahobhilam. These are the same type as those found in the Tallapakavarikottu on the temple hill and naturally belong to that temple and not to Ahobhilam. This is because the Telugu songs, both erotic and philosophical, which were recorded on them in the various ragas and talas are all addressed to Venkatesa or TiruVengalanatha by their authors Annamacharya and his son Tirumalayya of Tallapakam" (P.P 96).

We do not know since when and for how many centuries these copper plates were lying idle in those temple cellars – they must have been there positively from the end of the 16th century till the beginning of the nineteenth when the British took over the administration of the Indian temples. Though these were noticed by the officials of the Survey department, none seem to have taken any interest to bring them to light. It was only in 1922 or a little later that the Devasthanammade an attempt, though haphazard, to publish some of those songs and the other minor works of the Tallapaka poets. But the publications did not come out till 1935 and when they were published, in three volumes, the quality was far from satisfactory. Later in 1947, due to the indefatigable efforts of the Late Sri PrabhakaraSastri, a clear and thorough picture of the nature of those compositions emerged. As a result, the 4th volume in the series was released in 1947 with a critical introduction by Sastriji giving a detailed analysis of all the available copper plates. Sastriji took stock of the entire collection by personal supervision; classified them with ability, insight and immense patience and presented his findings in his valuable introduction. A year later he came across another MSS, depicting the biography of Annamayya, in metre composed by Chinnanna who the grandson of the composer himself. Publication of this work, by Sastriji, in 1949, along with his exhaustive and learned introduction to the text brought into clear focus, for the first time, the genius of the Tallapaka Poets. As providence would have it, in the same year another incident occurred. Sastriji wanted to probe into the temple cellar to make sure that nothing else remained there unsalvaged. This idea of Sastriji was taken as a command by his disciple and Research Assistant, Sri.

A.V.SrinivasaCharyulu who was slim enough to slip into the dark and long neglected cellar unmindful of what would befall him. There he dug the layers of dusk deposited in the centuries old cellar and happily fished out two more copper plates, on which were embossed the replicas of two human figures holding veenas in their hands.These resembled those sculptures on either side of the entrance to the cellar. The embossed figures, on the copper plates, tallied with the description of the Saint Composer found in the biographical work. Sastriji promptly identified one with an aged face as that of Annamayya and the other with a comparatively younger look with that of Peda Tirumalayya. These plates which were deposited in the cellar some four centuries ago saw the light of the day again only in 1949. Is it not a wonder that these two plates escaped the notice of the temple authorities who salvaged from the same cellar hundreds of other inscribed plates years before? It was as if the two plates had waited all the time, in the dark caverns, only to be discovered now by Prabhakara, the Sun God. Of course, it was a wonder to everyone who knew the incident. These plates though wrapped in the husk which generated enough heat to absorb any moisture in the cellar showed signs of bronze-copper diseaseon them with deep green patches. Obviously, the figures on them were shedding tears all these years. So Sastriji, after intimating the authorities of these findings, hurriedly rushed to Madras to get the plates chemically treated by competent archaeological chemists. Promptly they were cured of the disease and preservative metal varnish was applied. These came at a time when Sastriji was organising, with the help of the TTD, the first Music festival commemorating the 446$^{th}$Vardhanthi of the Saint Composer Annamayya. The function was a grand success. Thus, after giving the needed impetus and initiating fundamental research in the field, Sastriji died in 1950. However, the projects he initiated gathered momentum. Since that period the vardhanthi celebrations are being organised regularly by the TTD without any break and till date 27 volumes, containing Sankirtanas, have been published. With the issue of a few more volumes, the publication work on sankirtanas would be completed. No doubt, the publication of all these songs is almost nearing completion but the problem of preserving the original plates for further study and verification still exists.

The general editor, Prof P.V. Ramanuja Swami rightly pointed out in his note to the fourth volume, in 1947 that "The main work of the editor has been to change the peculiar Orthography of the plates to suit the modern times keeping the text faithful to the original in other respects". As these copper plates are in the custody of TTD, it becomes the duty of TTD to ensure that all the plates are properly treated against corrosion which has already set in as could be seen from the editor's observations which are given in the footnotes to the texts. The plates must be properly treated tp arrest further corrosion and preserved in a clean and salt free atmosphere where there is less humidity. Further, I suggest to the TTD to take immediate steps to get all the plates microfilmed or copied using microfiche techniques to facilitate easy access for reference to Scholars and Researchers interested in the field as verification from the plates today is practically impossible, in view of the problems involved. If these facilities existed today, probably we could have brought one copy and projected some of the plates during this seminar. I am sure the need for this would be felt more when we discuss next week about editing texts on copper plates.

When we realise that the TTD, with all their funds, manpower and materials, took nearly 35 years to publish what little was available on the copper plates, we can well imagine the enormity of the task which the Tallapaka Poets must have encountered four centuries ago. It is needless to emphasize any further the immense value of the literary heritage handed over to us by the 15$^{th}$ century family of Tallapaka poets. It is a treasure which is rare, unique and unparalleled in the entire history of the world of letters.

Today a lot is being done at the TTD for the preservation and propagation of this unique literary heritage by sponsoring several projects and much more is expected from them in the coming days and years. Thanks to the interest evinced by the TTD and the pioneering effort put in by Scholars like PrabhakaraSastri and RallapalliAnantha Krishna Sarma, we could review atleast this much of what little remained of this immortal golden treasure.

# PHILOSOPHY OF THIRUVALLUVAR IN THIRUKKURAL

**- Prof. P. Chinnaiah**

Thiruvalluvar is a great Philosopher and poet. Thiruvalluvar expressed philosophical thoughts excellently with his wisdom and insight in Tirukkural relating to human life and promotion of good living in the world with peace and happiness. The Tirukkural is a classic Tamil text consisting of 1,330 couplets or *Kurals* dealing with the everyday virtues of an individual. It is considered one of the greatest works ever written on ethics and morality, chiefly secular ethics, it is known for its universality and non-denominational nature. It was authored by Thiruvalluvar.

Kural means something that is short, concise, and abridged. The text has been dated variously from 300 BCE to 7th century CE. The traditional accounts describe it as the last work of the third Sangam, but linguistic analysis suggests a later date of 450 to 500 CE. ThiruKkural emphasizes on the vital principles of non-violence, vegetarianism, casteless human brotherhood, absence of desires, path of righteousness and truth, and so forth, besides covering a wide range of subjects such as moral codes of rulers, friendship, agriculture, knowledge and wisdom, sobriety, love, and domestic. The Thirukkural has influenced several scholars across the ethical, social, political, economical, religious, philosophical, and spiritual spheres.

Thiruvalluvar said that God is the primary force in the world. Reasoning with a drunkard is like going under water with a torch to seek for a drowning man. —Folded hands may conceal a dagger.

Likewise a foe's tears. Great wealth, like a crowd at a concert, gathers and melts. Those who have wisdom have all:Fools with all have nothing. When the rare chance comes, seize itto do the rare deed. Wants and hatreds are the sources of unhappiness and misery. The world survives because of the rains and therefore rain is known to be the nectar of immortality. Unless drops of precious rain fall down, not even grass will sprout its head. Even the vast ocean will start shrinking, if the clouds that took the water away from it don't pour it back. Not even Gods will get their due of prayers and rituals, if the skies dry up. Thiruvalluvar says, "Charity and penance will both cease to exist in this world if the clouds are not charitable. World cannot survive without water and morality cannot exist without rains."[1]

Thiruvalluvar said that righteousness yields good reputation and wealth. True moral integrity lies in being flawless in your thoughts. Righteousness is all about removing the four flaws – envy, desire, anger and harsh words. True joy blossoms only due to righteous deeds; all else cause unhappiness and disrespect. The only way to live is to live righteously. True virtue is to lead a harmonious family life. Virtue alone is happiness. Loveless people hold everything to themselves. Those, who are filled with love, bear even their bones for others. The ignorant think that love is needed only for righteous deeds; they know not that love is a friend for bravery too. Like life is linked to the flesh, love is integral to life. Joy in this world is the result of a life of love for all. Thiruvalluvar says, "Love yields affection for all, which leads to invaluable friendships."[2]

Thiruvalluvar said in Thirukkural that life without love will be stung by the righteous conscience like a body without bones that is burnt by the sun. A life without love in the heart is as futile as a dried-up tree blossoming in a desert. What useful things can external body parts do, if the heart is devoid of love? Only a person, who follows the path of love, is in a state of living; else it is just a skeleton dressed with skin. Thiruvalluvar expressed that the words uttered by enlightened scholars will only be kind words carrying love and no hatred. To be someone who speaks sweet words with a smile is even better than being humanitarian with a happy heart. To look at someone with kindness, to smile at them and to say pleasant words, is a nice

virtue. Painful poverty will not afflict someone who speaks only pleasant words to everyone. Humility and pleasant words are the true jewels for anyone and not anything else. If one seeks and speaks pleasant words that cause good all around, righteousness resides and harm recedes. Speech that is inseparable from sweetness will yield just-consequences and will reflect right virtues. A sweet word, said without the slightest spite, will provide delight in this life and the next. To be nasty when you can say nice words is like tasting an unripe fruit when you have a ripe one. Thiruvalluvar says, "To say harsh words when you have nice words is like plucking an unripe fruit when there are ripe ones."[3]

Thiruvalluvar expressed that neutrality is a great virtue when one practices impartiality, unfailingly, towards all sections including friends, foes and strangers. Position of power is good to occupy, when one practices impartiality, dependably towards all sections. Responsible business is when a business person, caringly, deploys other people's money as one's own. The world will not think ill of one who has stumbled into poverty because of being fair and righteous. Humility is a good quality in everyone; in particular for the wealthy, it is like their wealth. The greatness of one, who is wise and follows the right path by being restrained, will be recognized and appreciated. Thiruvalluvar says, "A wound caused by fire will heal inside a scar caused by the tongue never heals."[4]

A person who has envy has no wealth; a person who has no decorum has no growth. One attains eminence through decorum; one attains unprecedented bad name due to indecorum. Good conduct becomes the seed for goodwill. Bad behavior always yields agony. Class is determined by propriety of conduct; impropriety will lead to being considered part of an dishonorable class. One, who, in his envy, doesn't appreciate the wealth of others, is known not to value virtue and his own wealth. Envy is a damned ill that will destroy one's wealth. Sridevi, goddess of wealth, will despise a jealous person and direct him to her sister, Moodevi, goddess of poverty. The foremost among all wise deeds, is to refrain from doing harmful deeds even to those who hate you. Sinners do not dread, while great men dread the delusion of evil deeds. Thiruvalluvar says, "Vile deeds yield vile results; and hence, vile deeds are more fearsome than fire."5

Thiruvalluvar expressed that doing one's duty without desiring any favors is like rain; what can the world do in return for the rains. Giving to the poor is charity; all else has the quality of anticipating a return. Give, and live with fame; there is no other gain for lives. The gains of compassion are most precious; material wealth is possessed even by the despicable. The compassionate, who care for all other lives, do not fear for their own lives. The heart of one who has eaten and relished flesh, is like the heart of one leading an army: it cannot be compassionate. The heart of one who has eaten and relished flesh, is like the heart of one holding a deadly weapon: it cannot be compassionate. Compassion is exemplified by not killing; and the lack of it, by killing: to eat meat so obtained, is not virtuous. Swindled wealth, will seem to swell, but will breach its limits and bust. The excessive love for swindling renders a person unable to abide by the limits of reason. Thiruvalluvar says, "The love for swindling, when it yields, yields indestructible distress."[6]

Wisdom is the shield against suffering; and the fortress, impregnable and indestructible for the foes. Whatever whoever may say, listen to it, and wisdom is to comprehend the true meaning of it. Whatever whoever may say, wisdom is to comprehend the true meaning of it. Wisdom is to keep it simple for your audience, and to comprehend complexities in others' speech. The wise anticipate the outcomes unwise are those who can't. The wise who are perceptive and take preventive measures, will not be staggered by any harm. The wise, who are perceptive, and are prepared, will not be staggered by any harm. The wise have everything; the unwise have nothing, irrespective of whatever they may possess. Thiruvalluvar says, "Wisdom, it is to embrace the world, and not to be overtly delighted or dismayed. What the world conforms to, wisdom lies in conforming to those."[7]

Great wealth accumulates like a crowd assembled at a theatre; when it goes, it vanishes like the crowd at the end of the play. Wealth is transient; when it comes about, use it to do a lasting righteous deed. A bird breaks away from the egg, leaving the shells alone; so is life related to body. To renounce all that one desires, the five senses should be overcome. Miseries will not near those who desist from

desire. Desire is the seed which, for all lives, unfailingly, leads to an endless cycle of birth. Desire is a deceptive trap; fearing it, is a virtue. Purity is absence of desire; which, is obtained through the quest for truth. One can do flawless deeds the way one wants to do, when desire is destroyed completely. When desire, the misery of miseries, is decimated, unending joy abounds. Fate turns favorable at one's beckoning, when desire is destroyed completely. Those who are devoid of desire are free from misery; for others, it keeps piling up. There is no wealth here better than lack of desire. Thiruvalluvar says, "Eternal bliss ensues, when insatiable desire ends." [8]

Thiruvalluvar says that the military, citizenry, resources, advisers, friends and fortresses who owns these six is a lion amongst kings. Fearlessness, generosity, wisdom and vitality: these four are persistent characteristics expected of a king. Vigilance, erudition and boldness: these three never abandon a competent ruler. A righteous king who renders justice, and nurtures his people, will be hailed as an almighty leader. Charity, love, justice and safeguarding the people: who has these four is the guiding light for all leaders. Learning is the indestructible, and significant, wealth; other riches are not true wealth. Learning is the harmless, and significant, wealth; other riches are not true wealth. Listen to the good; doesn't matter, how much; whatever be the quantum, immense greatness, it will bring. Thiruvalluvar says, "Avarice, ignominious pride, and repulsive gloating, are fatal flaws in a leader"[9.]

Clinging to wealth, out of avarice, is worse than the any other flaw. Water assumes the quality of the land; wisdom depends on the company one keeps. Senses are governed by the brain; reputation is determined by one's company. Wisdom seems to stem from the mind is shaped by people around us. Purity of mind, and deeds, depends on the quality of one's associations. Thiruvalluvar says, "Hard work, the wrong way, will be futile even if the world applauds."[10]

Losing the capital in pursuit of profit is not a task the wise will undertake. The life of one with no intimate relationship is like water flowing into a pond with no bounds. Just as lack of rain is to the world, so is to mankind – rulers' lack of benevolence.

Being wealthy is worse than being poor, under an unjust ruler. When the king doesn't take counsel of his advisers, and tends to explode with rage, his wealth starts decreasing. Tongue creates and destructs; hence ensure there is never a blemish in your speech. Speak words that befit your capabilities and those of the listener; there is no greater virtue and wealth than that. The world will instantly pay heed to those who can speak coherently and pleasingly. Thiruvalluvar says, "Good companions lead to valuable gains, and good deeds lead to everything needed. Being content with ill-gotten wealth is like storing water in a fresh unbaked mud pot."[11]

Good deeds, even if begun with losses, will yield fruits later on. The means to the end, the obstacles and the fruits of the deed, anticipate all these, and then act. Love, wisdom and sagacious speech are three qualities essential for an envoy. Funds generated in the normal course, internal taxes and acquisitions from enemies are sources of wealth for the king. A smile on the face does not make a friendship; a smile in the depths of the heart does. A friend, who abandons in times of adversity, will torment the heart even at death's door. Discern it not merely as ignorance, but as a sign of intimacy, when a friend does an undesired act that hurts. Lack of wisdom is the worst thing in the life. Food and deed, both in excess and deficit, cause disease. Thiruvalluvar says, "Being good within, is nobility; All else is no quality."[12]

Thiruvalluvar says that Love, compunction, beneficence, compassion and truthfulness: On these five columns rests nobleness. The virtue of the ascetic is to not kill any being; The virtue of the noble is to never resort to blaming. Poverty is no disgrace if you possess the strength of nobleness. Humility is the vital capability of the capable: to convert foes, the weapon of the noble. Thiruvalluvar says, "Discord is the misery of all miseries; when it goes the joy of all joys ensues"[13]

Thiruvalluvar expressed in Thirukkural that a kind heart and good heredity are considered to make the quality of being well-

Thiruvalluvar expressed in Thirukkural that a kind heart and good heredity are considered to make the quality of being well-mannered. Wealth of a person, who is loved by none, Is a toxic fruit tree in the middle of a town. The conformity of the vile is forced by fear and perhaps, a small bit by desire. A word is enough to move the noble to help. Good milk spoiled by the vessel that is stained. Wealth is spoiled in the hands of uncultured persons. Deceitful women, drink and dice are friends of those deserted by the deity of wealth. Be wise among the wise but pretend to be dull among fools.

Love on one side is bad; like balanced load by porter borne, love on both sides is good. Love is tender as an open flower. Water is pleasant in the cooling shade; so coolness for time with those we love. Thiruvalluvar says," Love without hatred is ripened fruit."14 Salvation comes when one wants it to, when desire is destroyed completely. Real kindness seeks no return.

**REFERENCES**


1. Thiruvalluvar, Thirukkural, P. 4.
2. Thiruvalluvar, Thirukkural, P. 16.
3. Thiruvalluvar, Thirukkural, P. 20.
4. Thiruvalluvar, Thirukkural, P. 26.
5. Thiruvalluvar, Thirukkural, P. 42.
6. Thiruvalluvar, Thirukkural, P. 58.
7. Thiruvalluvar, Thirukkural, P. 88.
8. Thiruvalluvar, Thirukkural,P. 88.
9. Thiruvalluvar, Thirukkural, P. 90.
10. Thiruvalluvar, Thirukkural, P. 96.
11. Thiruvalluvar, Thirukkural, P. 134.
12. Thiruvalluvar, Thirukkural, P. 200.
13. Thiruvalluvar, Thirukkural, P. 174.
14. Thiruvalluvar, Thirukkural, P. 266.

## CONTRIBUTORS


1. **Prof. P Chinnaiah**
   Professor and Chairman BoS
   Department of Philosophy
   Sri Venkateswara University, Tirupati - 517501

2. **Dr. Veturi Anandamurthy**
   Formar Professor of Telugu
   Osmania University, Hyderabad - 500 007

3. **K. Balu Naik**
   Research Scholar
   Department of Linguistics
   Osmania University
   Hyderabad - 500 007

4. **K. Parameswari**
   Centre for Applied Linguistics and Translation Studies
   University of Hyderabad
   Central University Post
   Gachibowli, Hyderabad - 500 046

5. **Dr. Chelli Nageswara Rao**
   Post Doctoral Fellow (ICSSR)
   University of Hyderabad
   Central University Post, Gachibowli
   Hyderabad - 500 046

6. **B. Keerthana**
   Centre for Applied Linguistics and Translation Studies
   University of Hyderabad
   Central University Post
   Gachibowli, Hyderabad - 500 046