

## Annotation guidelines : Dialectal differences between Standard and Telangana dialect

For each rule there is a tag associated in square brackets. All the phonemes are written in written in WX text.

1	అ	ఆ	ఇ	ఈ	ఉ	ఊ	ఋ
	a	A	i	I	u	U	q
2	ఎ	ఏ	ఐ	ఒ	ఓ	అం	అః
	eV	e	E	oV	o	aM	aH
3	క	ఖ	గ	ఘ			
	ka	Ka	ga	Ga			
4	చ	ఛ	జ	ఝ			
	ca	Ca	ja	Ja			
5	ట	ఠ	డ	ఢ	ణ		
	ta	Ta	da	Da	Na		
6	త	థ	ద	ధ	న		
	wa	Wa	xa	Xa	na		
7	ప	ఫ	బ	భ	మ		
	pa	Pa	ba	Ba	ma		
8	య	ర	ల	వ	ళ		
	ya	ra	la	va	lYa		
9	శ	ష	స	హ	క్ష		
	Sa	Ra	sa	ha	kRa		

Writing Convention : Standard dialect - Telangana dialect

### Vowels:

#### A. Deletion of vowels :

- a) [VD1] Can be dropped at the middle of a word

jAwakaM - jAwkaM

nltukaM - nltkaM

loVpta - loVpata

nAgarkawa - nAgarikawa

anxarki - anxariki

- b) [VD2] Can be dropped at the end of the word

unnAyi - unnay

#### B. Addition of vowels :

- a) [VA1] The vowel - 'aM'

lakRamldevi - lachchiaMdevi

wAbElu - wAaMbElu

- b) [VA2]v-Addition of a for easy articulation

unxi - unnaxi  
unxo - unnaxo

C. Replacement of vowels :

- a) [VR1] a following y gets palatalised and becomes e - Palatalization of vowels

sAmAnyanga - sAmAnnega  
viRayam - viRReVm

- b) [VR2] A following a v gets rounded and becomes o and v is dropped

vAdu - odu  
prajAsvAmyaM - prajAsomyaM

- c) [VR3] i is changed to e , this generally occurs in verbs

undAli - undAle  
ceyAli - ceyAle

- d) [VR4] in the preposition lo o is changed to a

lo - la

- e) [VR5] Replacements of Long vowels with short vowels

vaswAru - vaswaru  
ceVwAru - ceVpwaru  
ceswAru - ceswaru

- f) [VR6] u -> i :

undaxu - undaxi ( verb+negation)

- g) [VR7] Verbs in their root form u->a ; aM->a

ceyadaM - cesudu  
winadaM - winudu  
wAgadam - wAgudu  
kattadaM - kattudu  
urakadAm - urkudu

**Consonants:**

1. Addition of consonants :

- A. [CA1] Addition of n :

ceVppAlA ? - ceVppAlna ?  
padukune - pandukune

- B. [CA2] Addition of g :

- a) [CA2a] Deixial terms : Words or phrases that require contextual information to be fully understood for example, English pronouns are deictic

a) Person Deixis:

Ayana - gAyana (his)  
glme - lme (her)  
vElylYu - gElylYu(they)

b) Spatial Deixis :

A - gA (that)

c) Temporal Deixis :

ippudu - giipudu (now)

b) [CA2b] In General

inwa - ginwa (so much ) [ quantitative pronoun ]  
axe - gaxe (same)  
Anthe - ganthe  
Anxuke - ganxuke

C. [CA3] Addition of t : [ Delete this rule ]

alAne - atlane ( as well as ) [ adverb ]  
ela - etla  
ila - gitla (like ) [ filler , adjective ]  
ilAne - gitlAne (similarly)[ adverb ,adjective]

The above rules can take place simultaneously too

2. Deletion of consonants :

- A. [CD1] A cluster of identical consonants preceding a second short vowel in a word is simplified into a single consonant. This rule operates immediately following Dropping of vowels [Rule 14 Sarma's Analysis]

oVppukoru - oVpkoru  
ekkuva - ekva

- B. [CD2] Three consonant clusters of ndr, ndl, gny, rnm are reduced by dropping the middle consonant. [Rule 15 Sarma's Analysis]

- ndr - nr : tandri - tanri ; cUdundri - cUdunri
- Ndl - nl : bandlu - banlu ; indlu - inlu
- Gny - gy : gnyanaM - gyanaM
- rnm - rm : gavarnmmeVnt - gavarment

- C. [CD3] v at the beginning of the word

vachindi - achindi

vaswunnayi - Oswunnay (+ Replacement of vowels )

vAna - Ana

vAllu - Allu

- D. [CD4] r is dropped when followed by plosive in a consonant cluster .Occurs in both nouns and verbs

prAnaM - pAnaM

brawakadaM - bawakadaM

- E. [CD5] Dropping of n (Hypercorrection)

padukovadaM - pandukovadaM

### 3. Replacement of consonants :

- A. [CR1] Voiced and Voiceless consonants :

- a) [CR1a] When two words enter into Sandhi, the initial **voiceless stop** of the second word becomes **voiced** mostly when the last consonant of the first word is a voiced one or when the first word ends in a vowel. Sometimes the vowel may drop as in the case of the rule above . [ Rule - 8 ]

- k - g

pedda + kAleV -> peDDagAleV

ippudu + kUda -> ippudugUda

- ch - j

pulini + caMpina -> pulinijampina

tayAru + cesi -> tayArujesi

- w - x

battalu + wemmannadu -> battalxemmannadu

appulu + teswU -> appuldeswU

- p - b

Inwaku + pUrvaM -> inwakbUrvam

nllwoVni + pandwadi -> nllwoVnbandwadi

- b) [CR1b] A voiceless stop preceding a voiced stop becomes voiced and similarly a voiced stop preceding a voiceless stop becomes voiceless [ Assimilation ]

- kg - gg : pATaRAaku + gaxelexu -> pATaRAaggaxelexu

- wx - xx : warvAwa + xAka -> warvAxxAka

- pb - bb : kadupu + bAWa -> kadubbAWa

- xw - ww : ruxxu +wAru -> ruwwAru

c) [CR1c] Not only in sandhi , these changes apply within a single word. This rule hasn't been stated by sarma . The examples are as follows :

- k - g : kAxu - gAdu ; kadA - gadA
- c - j : cAlu - jAlu ; cepinchukuntam - jepinchukuntam
- p - b : peVttAru - beVttAru

d) [CR1d] The final voiceless consonant of a word rarely gets voiced when preceded by a voiced consonant [ Rule-9 Sarma's Analysis ]

nAku + ledu -> nAgledu

Manaku + jEwAlu -> managjEwAlu

B. [CR2] p->v :

a) [CR2a] Sometimes the initial /p/ of a word becomes /v/ when the word follows another word .[Rule 10 Sarma's Analysis]

nIllu + pettadam -> nIlluvettadam

lekka + peVtandi -> lekkaveVtandi

b) [CR2b] Might occur in individual words too .

C. [CR3] Assimilation :

a) [CR3a] Any consonant preceding /y/ Gets assimilated

- cy -> cc : sacyaM - saccaM
- jy -> jj : ujjyogaM - ujjogaM
- wy -> yy : muwyaM - muwweM

b) [CR3b] An r following k or b gets Assimilated with it [ Rule 27 ]

IEbreVri - IEbbeVri

sakramanga - sakkamanga

c) [CR3c] A d preceding j assimilates with it [ Rule 31 ]

idi + jeVsinru - ijjeVsinru

Adi + janpoVyindi - ajjanpoVyindi

d) [CR3d] A d and r preceding a l or n gets assimilated with it [ Rule 26 + Rule 29 ]

vAdini -> vAdni -> vAnni

Vaxxidulu - vaxxidlu -> vaxxillu

Urlu - Ullu

vyAkarnaM - vyAkannaM

e) [CR3e] An s preceding w or d assimilates with it [ Rule 28 ]

ceswAru - cewwAru  
iswa - iwwa  
ceswa - jewwa

f) [CR3f] x ->w :

arwaMkaad - arxaMkaad

D. [CR4] Dissimilation

a) [CR4a] Retroflexization :ll->dl . Followed by addition of n

illu - indlu  
kallu - kandlu  
yella - yendla

E. [CR5] Aspiration :

a) [CR5a] s->R

lesi - leRi  
veVsi - eVRi  
muriRipotunnadu - murisipotunnadu

F. [CR6] c with s at the beginning of the sentence word

cudandi - suduri , cuduri - suduri  
Cakkaga - sakkaga

G. [CR7] r with l

evaru - evalu

H. [CR8] v->t : when previous syllables ends with eV/e :

tinevallaki - tinetollaki  
Kotladevalla -kotladetolla  
tirigeVvAdu - tirigeVtodu  
undevAdu - undetodu

I. [CR9] v->y :

ivaala - iyaala  
kattetivi - kattetiyi  
ivvavA - iyavA ?  
Ilrojuki / evalatiki - yalatiki

weliyaxu - welvaxu ( Hypercorrection)

J. [CR10] d->t [ Devoicing ] :

Doesn't depend on place . when the previous syllable ends with a nasal  
then the onset of the following syllable is devoiced .

Lexanakunda - lexanakunta  
gAkunda - gAkunta  
undatledu - untaledu

Hypercorrection : Some times the reverse may also happen

K. [CR11] t->x  
enti - enxi

Hypercorrection : Some times the reverse may also happen

L. [CR12] va/ta/ca ->swa ( Negative verbs in particular )

M. [CR13]Palatalization : Any consonant preceding /y/

- wy - cy :  
sawyaM - sacyaM  
muwyaM - mucyaM
- xy - jy  
uxyogaM - ujjyogaM

4. [ CDu1] Duplication of consonants :

According to Nalimela Bhaskar in his book Telangana Padakosam , “

Duplication of consonants “

masi - massi  
gasagasAlu - gassAlu  
dayagala - dayagalla  
lekapoyina - lekkunna

## Syllables :

1. [SD] Deletion of Syllables :

A. [ SD1] Dropping of syllable( with C as d) in words [ Spatial diexial terms ]

Adigo - ago  
Akkada - Ada  
Ikkada - Ida  
eVkkada - eDa/yAda

B. [SD2] The second syllable of the shape CV is dropped after a short vowel  
when the C is either /x/ or /r/

moVxalu - moVlu  
UrikeV - UkeV

- C. [SD3] The second syllable of the shape CV following a long vowel is dropped

bAgAneV - bAneV  
lekapotheV - lepotheV

- D. [SD4] The second syllable of the shape C1 V C2 of which C1 and C2 are the same or similar., is reduced to C2

eVnanwa - eVnwa  
samanjasanga - sanjasanga

- E. [SD5] If the fourth syllable is /va/ (This seems very specific to particular word ) or a repeated third syllable, it is dropped

seVnagavaraM - seVnagaraM  
pariswiwiga - pariswiga

- F. [SD6] The third syllable CV where C is /k g y v h/ is dropped

rUpAyalu - rUpAlu  
ganaka - gana  
caxuvukoVni - caxukoVni

- G. [SD7] In case of negative verbs last syllable is dropped

ceyaledu - ceyAle  
rAledu - rAle  
brawakaledu - bawakale

- H. [SD8] Dropping of last syllable

annaya - anna  
unnAnu - unnA

## 2 . [SR] Replacement of syllables :

- A. [SR1] Verbs in past tense

- a) [SR1a] The second last but one syllable's long vowel gets replaced as  
M - Masculine

Third person - Plural - M/NM : in/i/shortening of vowel

cesAru - jesinru  
kurcopeVttAru - kursopeVttinru  
pampincAru - pampinru  
pettAru - pettinru  
poyAru - poyinru



vowels)

cesAru - jesiru

pettAru - Petirr ( + duplication of consonants + dropping of

Third person - Singular - Masculine : in

pAdAdu - pAdindu

ayyAdu - ayyindu

brawikAdu - bawikindu

Second person - Singular/Plural - M/NM: ina

chesAvA - jeRinavA ( +c->j ; s->R)

tAgAvu - tAginav (+Deletion of vowels )

First Person - Singular - Maculine/Non-Masculine : in

pettAnu - pettina

*First Person - Plural - Masculine/Non-Masculine : in*

*chesAmu - cesinAm*

- b) [SR1b] In general for any verb irrespective of number/Gender/Person : Root Word + e ( with last syllable being dropped )

pAkinxi - pAke

perigAyi - perige

B. [SR2] Verbs in Present tense :

- a) [SR2a] Third-person Plural and 1st person plural:

Containing r as onset in last syllable , previous syllable [nnA]is dropped this might be accompanied with drop of vowels[Generally last vowel] and sometimes reduplication of consonants[rr] too . There might be partial dropping of the syllable too [n is retained ]

veswunnAru - eVswur

gunjuwunnAru - Gunjuwur

Nimpukutunnaru - nimpukuturu

vastunnArA ? - osturA ?

unnAru - unru

aiwunnAru - aiwunru

anukuntunnAru - anukuturru

kaduwnnarA - kaduwunrA

ippiswunnAdanta - ippiswundanta

- b) [SR2b] Second person - singular -Non-Masculine :

tunxi/wunxi/wunnaxi - wAnxi / waxi [ w->t sometimes ] (syllable change)

untunxi - untaxi

avuwunxi - Aiwaxi

Paduwunxi - padtaxi

Ceswawunxi - jeswataxi

Untunxi - untaxi

*Second person - singular - Masculine :*

Same rules as above but xi will be replaced by du

C. [ SR3 ] Replacement of honour affixes :

Honor affixes cudu + andi => cudu + ri/dri/ru

cudandi - cuduri

padakandi - padakuri

Undandi - unduru

pondi - pondri /ponri

randi - randri

Some more points :

1. The prepositions/vibhakthis of ancient telugu are still used intact in telangana dialect which are almost never used in standard dialect. Take an example of "Chaturthi vibhakti" - koVrakun/KE

**References :**

- 1.Nalimela bhaskar - Telugu Padakosam
2. Thesis of Pavan Chakravarthy
3. Sarma's rules on Telangana Dialect
- 4.Tilak's Blog posts
- 5.Subramanyam Marripoddi Blog posts
- 6.Teenmar news

**Comments:**

1. What are our contributions?
  - a. Annotated dataset? (Are we providing this? I think not, since we did not transcribe each speech word by word)  
Ans : So now given these rules , we annotate speeches( 3 speeches are already transcribed) . All the observations may not apply in KCR in speech

. We identify them (Some elimination methods) and display them in the paper.

- b. Annotation guidelines? (Will the guidelines be useful for other domains, where code switching happens? How are these guidelines different from earlier work on comparison of Telangana vs standard dialect? Will these guidelines provide some help when working on other languages? What are the language specific guidelines, and what are the generic guidelines which can be extended to other languages of the Dravidian family?)

And :

- Domains where these rules can be useful : dialect identification in speech or text.
- There hasn't been any paper differentiating Telangana dialect and standard dialect at word level in text data . And these observation are not from single source ( Check doc Telangana reference material ) They are from different books and also include observation which we are making ( for example Teenmar news)
- Myself not being familiar with other languages and dialects of other languages can' really comment on how these can be used in other Dravidian languages

2. What is our main aim? (If you could describe this whole project in two to three lines, what would it be?)

- I have added another doc named Work-Flow into the folder

3. How is our work going to be useful? (Does it have any insights about how such comparison can be done in different languages?)

- At the end of the day we are also looking at the sociolinguistic perspective of political speeches and how they change with respect to the audience and topic which is addressed . But here we are looking only at one person speeches . Can be extended to other political leader in one language and see whether there are similar observation happen in other languages too

4. What is the final closing statement? (What have we found finally?)

- More code-switching ( in terms of Telangana Dialect ) in speeches with mass crowds and more code-mixing ( English Phrases and sentences ) when it comes to refined audiences like Government employees,MLA . We would also like into take into account the topic of the speech but unclear on how the final closing statement would be wrt to this

5. How are code switching and code mixing influencing politics? (First hypothesis. Next, supporting stats.)

And.Didn't get the question . But that's what we wanna show through statistical analysis and Pragmatic analysis of the speech .