

Language and Society (CL2.203)

Monsoon 2021, IIIT Hyderabad

Project Interim Report

Abhinav S Menon

Data Collection

As of now, I plan to collect data in casual conversation. Most of my friends and relatives of my generation living in Bangalore speak in the same way, and I have been able to collect some data from their day-to-day speech itself.

If the data proves insufficient, it may necessitate collection through interviews. However, this method will not be optimal, as some of the syntactic divergences in Indian English are in the syntax of interrogative statements.

It is also possible to collect data from online sources. Public Instagram reels are a viable option. Some content creators use the specific dialect of Indian English under study. If this approach does not have privacy-related concerns, it can be taken up.

Statistics and Analysis

24 sentences have been collected. 8 of these are questions, 3 imperatives, and the remaining 13 are declarative statements.

There is no stratification based on gender. All speakers are in the age group 18-30; no further divisions can be made based on age.

Furthermore, the participants are largely of South Indian origin. As all the syntactic features to be considered are common to all these languages (and some common to Hindi as well), this also is not a factor.

Therefore, the differences among the speakers due to environment and social factors will *not* be considered at all. Only the divergences of Indian English from Standard British English, and the corresponding features of the Indian languages (as mentioned in the statement) will be studied.

A preliminary analysis of the data reveals the following.

Statements

Interrogatives

We note four types of questions common in the dataset. One of these is present in ordinary British English as well.

1. **Rising intonation.** The tone of the sentence rises towards the end, indicating that it is a question. This method is common in both English and Indian languages.
2. **Simple substitution.** In open questions, the part of a statement that is to be questioned is simply replaced with the corresponding *wh*-word. Unlike British English, *wh*-fronting does not occur. This occurs in all Indian languages under consideration, including Hindi. Alternatively, the *wh*-word may be fronted, but inversion may not occur. As inversion is not a feature of the Indian languages, this cannot be compared with them.
3. **Universal pronominal tag.** This is a method of tagging to form yes-no questions, using only the third-person neutral pronoun (*i.e.* forming the universal tag *is it?* or *isn't it?*). This feature, interestingly, has no *perfect* parallel in the Indian languages, although all of them do feature universal tags not involving any pronouns. This will be further investigated in the final project.
4. **Tagging by schwa.** A simple schwa sound with rising intonation (usually written as *uh?*) is commonly used as another tag for yes-no questions. This has a clear parallel in all South Indian languages and other dialects of Indian English spoken in South India.
5. **Tagging by “no”.** This is a form of tagging yes-no questions, not specific to Indian English, but common in the speech of the native speakers of many European languages. The universal tags not involving pronouns, as mentioned in point 3, are possible candidates for this feature. This, too, will be further investigated.

Imperatives

One feature is common to imperatives in Indian English, as distinct from British English – the pronoun can be mentioned in second-person imperatives. In Standard British English, this is not allowed.

This feature, too, is present in all five Indian languages under consideration.

Declaratives

No special features with regard to the overall structure of declaratives can be seen in the dataset.

Noun Phrases

It is common for the *definite* article to be omitted from noun phrases. However, an omission of the indefinite article is not found in the dataset.

Some of the languages we are considering do have an indefinite article, while none of them have a definite article. This could be the reason for the divide. This will be further studied.

Furthermore, in place of the indefinite article, the numeral *one* is sometimes used. This is a feature of the languages in our set which do not have indefinite articles.

Other

Tenses

One notable feature is the conflation of the future tense and the simple present, which is common to some other South Indian languages. Frequently, the future tense is used to describe a habitual action.

Only

The word *only* is often used as a marker of emphasis or specification. All the Indian languages we consider have a particle with this function that also carries out the standard function of the English word *only*.

Conjunctions

There are two conjunctions of special interest. The first is *but*, which can be used at the end of a sentence in Indian English. In such constructions, it is equivalent to *though* in standard English.

Another is a conjunction only in Indian English – the word *means*. It is used with the same function as *if...then*, *i.e.*, as a conditional marker; it comes between the condition and the consequence. Some South Indian languages have a word which is used to indicate both the typical meaning of the English word *means* and the *if...then* construction.

Proposed Algorithm

As has been done in this report, I plan to proceed in the following way for the rest of the project:

1. Categorise the data according to syntactic constructions and morphological features used.
2. Classify the features broadly and study each class separately.
3. For each feature, identify the Indian languages which have the same or a similar feature.