Causative Constructions in Indo-Aryan and Dravidian Languages: Final Report

Prajneya Kumar - 2019114011
Jayant Panwar - 2019114013

April 23, 2021

Introduction

As our final course project in LTU, we try to identify and search for patterns in **causative constructions** in Hindi, English and Telugu languages.

Majority of the work carried out so far has been emphasized more on widely used English and other European languages. Indian Languages have been explored little because of the amount of information available in non-English language is relatively less. Even for English, not much research work has been conducted as far as generalizations about causative construction is concerned.

We approach this problem by first looking through a dataset of 500 sentences each for the three aforementioned languages: Hindi, English and Telugu; we annotate the data and extract sentences which we believe contain causative constructions. Next, we look for patterns in the aforementioned data we obtain and try to conclude about the environment in which causatives manifest themselves in different languages.

Literature Review

There are a lot of literature which exists for Causative Constructions, but a general pattern observed is that they are not transferable to Indian Languages. This is either because they are restricted to patterns observed in the language or due to lack of resources in IL. Some resources we reviewed which could be helpful in some way or the other include:

1. Bernard Comrie - Chapter 8: Causative Constructions: Bernard Comrie, in this chapter, explains typology of the syntax and semantics of causative constructions. Comrie explains the different types of causatives, which we will discuss further, and proposes differnt parameters that affect the existence of causatives in a sentence.
2. A Dependency Treebank for Telugu: This paper by Taraka Rama and Sowmya Vajjala does not talk about causatives in particular, but talks about annotations and development of Telugu treebank. This helped us in looking for patterns in the Telugu corpus.
3. Morphological Causative Construction in Hindi (Its Phonological Conditions and Limitation) : This paper by Hema talks about the different conditions for Morphological Causative Construction in Hindi and also discusses the limitations of morphological causative constructions in Hindi.
4. Causative Sentences in Hindi Revisited: This old journal written by Yamuna Kachru provides a very important fundamental understanding of causative constructions in Hindi and covers almost everything there is to cover about causatives in Hindi.
5. Complex Predicates in Telugu (A computational perspective): This research paper by Rahul Balusu explains in detail about the complex predicates in Telugu from a computational perspective. The paper sufficiently covers causative suffixes and morphemes in Telugu as well.

Causative Constructions

Comrie describes Causative Constructions as an expression that denotes a complex situation involving two component events:

1. the causing event, in which the causer does or initiate something and
2. the caused event, in which the causee carries out an action or undergoes some kind of change as a result of the causer's action

Type I: Lexical Causatives

In this case, idea of causation is part of the semantics of the verb itself (Wikipedia). For example, die, kill etc.

Type II: Morphological Causatives

These consist of words which have certain suffixes, prefixes or infixes attached to them to transform them into a causee/causer. For example, सुलाना , सुलवाना etc.

Type III: Periphristic Causatives

This case includes those sentences which make the use of more than one verbs or clauses to express causativity. For example, *I made him write.* The use of two verbs here, *made* and *him* denotes that this is a periphristic causative.

Midterm: Recap

In order to find causative constructions in our corpus, we opted for different strategies. Generally speaking the methods were a combination of both manual and automated scanning. We scanned the entire corpus of a given language using python code at first and also opted to do manual scanning in cases where finding the causative constructions using code was a difficult task.

Telugu

Since causative constructions in Telugu are constrained to a limited number of causative verbs and due to our non-native knowledge of Telugu, the causative constructions in Telugu corpus were scanned out using python code, while Ph.D. Scholar Hema Ala helped us in the manual annotation.

In total, we were able to extract 370 sentences containing our desired construction.

Hindi

Causative constructions in Hindi are easily identifiable owing to the morphological changes in the transitive and intransitive verbs. The suffixes such as 'वाया', 'वाना ', 'वाओगे ', etc. when attached to the verbs transform them into causative verbs. However, causative constructions are not just limited to attaching suffixes to the main verb. For example, in some sentences, 'करवाना ', 'करा दिया', and other variations of 'कर ' denote presence of causative construction as well.

Finding these causative constructions in our corpus of Hindi sentences was partly a manual task and partly an automated one. Suffixes like 'वाना ', 'वाता ', 'वायेगा ', etc. were searched for using code written in python meanwhile the other cases of causative constructions (like 'करवा दिया', 'करवाकर ', *etc.*) were found out manually.

In total, we were able to extract 97 sentences containing our desired construction.

English

In order to look for causative constructions in English, we listed certain verbs that occur in causative constructions with greater than a chance frequency. Once the sentences containing the verbs were filtered, we manually checked each one for a false positive, i.e., the sentences in which the verb did not take a causative form but was filtered solely due to its presence in the sentence. This process was repeated until we had scanned the entire English corpus for all the possible cases of causative constructions.

Here, we made a list of verbs which would help us classify causative sentences. These list of verbs included all three forms of: "let", "allow", "permit", "make", "made", "force", "require", "get", "got", "help".

In total, we were able to extract 1436 sentences containing our desired construction.

Analysis of Causatives

Individual Realisations and Environments

Hindi

In Hindi sentences, causatives are used in a very simplistic manner. They are expressed in verbs' suffixes. For example, the suffixes "वाना", "वाया", "वाई", "करवा", "वाओगे", "करा दिया", "वाता", "वायेगा", "वाकर" when attached to verbs, they display causativeness.

These are mostly morphological causatives, the examples of which in the corpus are:

1. इस राज्यव्यापी हस्ताक्षर अभियान के माध्यम से प्रदेश भर में 10 लाख हस्ताक्षर करवाकर राज्यपाल को सौंपे जाएंगे।
2. संघ के प्रदेशाध्यक्ष बृज मोहन गुप्ता ने सरदार पटेल व इंदिरा के बलिदान से विद्यार्थियों को अवगत कराया।

We saw the presence of periphristic causatives as well. These could be seen with the presence of words such as मजबूर. For example,

1. उन्होंने कहा कि पाकिस्तान की भारत के प्रति हीन भावना ही है, जो बार-बार उसे गलत मार्ग पर चलने को मजबूर कर रही है।

Lexical causatives are not easily realised in the language.

In general, a base verb root can be converted into a causative by adding the ा or वा suffix. We can define the causative levels as follows:

1. Base Verb - खा
2. First Causal - खिला
3. Second Causal - खिलवा

In each step of causative derivation, the valency of the verb increases.

English

In English sentences, causatives are expressed by certain verbs like "let", "allow", "permit", "force", "require", "get", "got", "help". For example,

1. Mahanta said that you should not allow inferiority to enter your mind and devote your mind to God.
2. The Katha will be completed and the final oblation to the Havan will be made on Saturday.

There is also a special case where causatives can be expressed using the verb "has", "had" or "make", "made", combined with another verb, if the sentence has the following format:

has –NOUN PHRASE– <another_verb>

For example:

1. Ram had the robot clean the room.

Clearly, it must be noted that even though both the above cases are periphristic, the valency increases in the second case.

These are periphristic causatives.

For each of the following verbs above, the following generalizations can be made:

1. **TYPE I: HAVE, MAKE, LET** - SUBJECT + HAVE + OBJECT + BASE VERB
   Ram had the robot clean the room.
2. **TYPE II: GET, HELP** - SUBJECT + GET + OBJECT + INFINITIVE VERB
   Ram got the robot to clean the room.
3. **TYPE III: GET** - SUBJECT + GET + OBJECT + PAST PARTICIPLE VERB
   Ram got his room cleaned.

English also shows lexical causatives with certain verbs. For example, the causative for *die* becomes *kill*.

Telugu

In Telugu sentences, causatives are expressed either by words such as:

చేయిస్తాను
చేయించాను
చేయిస్తానా ?
చేయించానా?
చేయించను
చేయించలేదు
చేయించనా?
చేయించలేదా?
చేయిస్తాను
చేయించలేదు?
చేయించను?
చేయించాను?
చేయించాలని
చేయించారు
చేయించుకున్నాడు

etc.

which, simply translated, are words such as will do, make do, did etc.

or by verbs that contain the infix *"inchu", "inchi", "inchadu", "inchindi", "inchaaru"*.

The above observations show that there are mostly morphological causatives only in Telugu, manifesting them in verbs containing the infixes stated above. For example,

1. **timindiyaku chndina aiduguru kriktarlanu aisoleshanku pampincharu**
   Five cricketers who belonged to TeamIndia were sent for isolation
2. **tana dabbulanu rohit sharma chllinchadani tlipadu**
   He said that Rohit Sharma had his money squandered

Generalizations

Generalizing a common type of causative constructions for all the three languages discussed till now is something that is not possible. This is due to the fact that these languages showcase different forms of causative constructions. For example, English uses periphristic causatives and causative verbs for introducing causative construction in a sentence whereas Telugu uses morphological causatives by making use of infixes. Hindi acts as a sort of middle ground as it introduces causative construction through various ways that comprise morphological causatives (through suffixes attached to the verb) and periphristic causatives (by making use of certain verbs).

However, not having a common form of causative construction across all the three languages does not mean that there is nothing common about causative constructions in the three languages that cannot be generalized. One of the most important aspects of causative constructions, the environment, is a feature that can be generalized for all the three languages. We just need to keep one thing in mind when discussing the environments that English is a SVO word order language whereas Hindi and Telugu are SOV word order languages.

The generalized environment for causative construction requires that there be a **subject, object, action verb, <causative component>**. Subject obviously identifies the causer who causes some action and Object denotes the causee that actually does the action. The action verb denotes the action done by causee. The most important component of our generalized environment is the *causative component.* This can be anything from a specific verb to a suffix, depending on the language under consideration. For example, in English it could take the form of a specific causative verb or lexical item in case of periphristic causative construction whereas in Telugu causative component might be an infix attached to the action verb. In Hindi, the causative component might be a suffix attached to the action verb or a certain lexical item or verb that signals periphristic causative construction in the sentence.

Finally, the order of the components are taken care of according to the language in question, i.e., SOV if Hindi or Telugu and SVO if English. Therefore, a causative construction in English, Telugu, and Hindi will always have the four major components: **subject (causer), object (causee), action verb (caused action), <causative component>**

We can also see the same environment rule working perfectly in some other languages as well. For example, in Songhai:
Jamaal nga-ndi tasu di
Jamaal eat <causative component> rice the
Jamaal got someone to eat the rice.

This sentence in Songhai has all the four components of our generalized rule. Similarly, we can take another example but this time from Kannada:
Avanu nanage bisketannu tinnisidanu
He I biscuit eat <causative component>
He fed me a biscuit.

In Kannada, the word order is SOV so the causative component is attached to the action verb in the sentence at the end. Nevertheless causative construction in Kannada also happens to fall in line with our generalized causative rule by incorporating all the four components in its construction.

Therefore, our generalized environment rule for causative constructions that we were able to derive after thorough analysis and research of hundreds of sentences in English, Telugu, and Hindi, works for other languages as well like we saw in the examples of Songhai and Kannada.

Drawbacks and Challenges

Greenberg's Universals

Greenberg's Universals focus mostly on Tyopology, Syntax and Morphology. Even after extensive research on Greenberg and causatives, we could not find a relationship that allows us to give a universal generalization about causatives.

The parallel corpora of English and Hindi did not ensure that if the sentence is of causative construction in English corpus it will be of causative construction in Hindi as well. We came across many such challenging translated cases where the construction became different. Therefore, in order to solve this we opted for a thorough individual investigation into both the English and Hindi corpus without worrying whether it's translation is of causative construction in the corresponding language or not.

For the Telugu corpus, it was kind of Ms. Hema to help us in finding causative constructions in the Telugu corpora as neither of the team members

have linguistic knowledge regarding Telugu.

As we can see, the number of Telugu and Hindi sentences we could obtain which contain causative construction is relatively less than those we could find in English. As far as Telugu is concerned, this was because we were not able to find a large enough dataset for Telugu language, and thus had to borrow a corpus from another team with a different construction. Our corpus consisted of 900 lines, and thus we could find only a limited number of sentences there. As far as Hindi is concerned, we had a corupus 520 good complete sentences obtained from the IIT Bombay DevSet Corpus (consisting of 520 sentences), which yielded around 50% of the sentences we have obtained. To obtain even more sentences, we took the training set of the IIT Bombay Hindi Training corpus, whose majority of sentences were very short.

Conclusion & Future Work

We were able to successfully scan and analyze our corpus with respect to causative constructions. Furthermore, we tried to find implicational rules that try to correlate the structure of the sentences to the presence of the causative constructions. We looked at patterns occurring in individual languages(Hindi, English and Telugu) and spcified environments in which causatives occurred in each of the three languages. Furthermore, we tried to come up with a generalization that can be applied to all three languages; and tested how they might work with other languages as well. We looked at Greenberg's universals to see how they relate to the presence of causatives, which however, did not give us good results.

*****