Proceedings of the 16th Joint ACL-ISO Workshop Interoperable Semantic Annotation (ISA-16), pages 88–93
Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020

Detection and Annotation of Events in Kannada

Suhan Prabhu, Ujwal Narayan, Alok Debnath, Sumukh S, Manish Shrivastava
Language Technologies Research Center
International Institute of Information Technology
Hyderabad, India
{suhan.prabhuk, ujwal.narayan, alok.debnath, sumukh.s}@research.iiit.ac.in
m.shrivastava@iiit.ac.in

Abstract

In this paper, we provide the basic guidelines towards the detection and linguistic analysis of events in Kannada. Kannada is a morphologically rich, resource poor Dravidian language spoken in southern India. As most information retrieval and extraction tasks are resource intensive, very little work has been done on Kannada NLP, with almost no efforts in discourse analysis and dataset creation for representing events or other semantic annotations in the text. In this paper, we linguistically analyze what constitutes an event in this language, the challenges faced with discourse level annotation and representation due to the rich derivational morphology of the language that allows free word order, numerous multi-word expressions, adverbial participle constructions and constraints on subject-verb relations. Therefore, this paper is one of the first attempts at a large scale discourse level annotation for Kannada, which can be used for semantic annotation and corpus development for other tasks in the language.

Keywords: Corpus Annotation, Kannada Event Analysis, Event Detection

1. Introduction

Event detection and analysis is a rapidly evolving field of Natural Language Processing (NLP) and Information Retrieval and Extraction, as it allows us to generalize temporal data in terms of actual time and relative to other occurrences and events. Providing temporal and sequential information can enrich text and its representations which can be used for multiple downstream NLP tasks such as question answering, automatic summarization and inference, in an interpretable and linguistically informed manner. However, automatic event analysis, like many discourse analysis and representation tasks, requires extensive manually annotated training data.

Kannada is a resource poor, morphologically rich Dravidian language with about 45 million speakers[1], mostly located in southern India. Work in Kannada NLP has been limited to the development of tools for syntactic and morphological analysis, almost no work has been done in semantic tasks in this language (Mallamma and Hanumanthappa, 2014), due to few experts, lack of training data and the morphological and semantic characteristics of the language. This paper is one of the first attempts to introduce a semantic analysis and enrichment task at a semantic level into Kannada, i.e. semantic level event detection and analysis.

In this paper, we aim to understand the various parts of speech, syntactic structures, and associated semantic patterns that allow the identification and representation of events in Kannada. We also present the challenges associated with identifying events in Kannada due to morphosyntactic constraints such as multi-word expressions, ubiquity of verbal, adverbial and adjectival participles, analytic verb negation, and absence of copula (Kittel, 1993). We follow a derivation from the TimeML event definition, which has been modified to adapt the zero-copula and participial constructions, so as to make it less ambiguous for annotators.

Finally, we present a dataset of 3,500 annotated sentences, along with a detailed analysis of the dataset including some basic dataset statistics. We annotate events on the Kannada Dependency Treebank (Rao et al., 2014), which consisted of approximately 4,800 event mentions. We show that our guidelines are succinct to a Kannada annotator by our high inter-annotator agreement, along with a distribution over various syntactic structures and a linguistically motivated explanation for challenges in some constructions that have been elaborated in Section 5. The corpus has been made freely available[2].

2. Related Work

In this section, we introduce some of the work done in event detection in low resource and morphologically rich languages, with a focus on TimeML event extraction, or event representation in Indian languages. TimeML was introduced by Pustejovsky et al. (2003) as a mechanism of recognizing, annotating, classifying and representing events in text for the purpose of question answering. TimeML has been used in event detection across languages such as Italian (Caselli et al., 2011), French (Bittar et al., 2011), Romanian (Forăscu and Tufiş, 2012), and Spanish (Saurı, 2010). Of course, corpora annotated with TimeML events have often been done alongside the detection of other temporal information such as time expressions, temporal links and other notions.

For languages which have syntactic structures that vary significantly from English, event detection is used as an introductory task and the definition of an event is modified to be true to the syntactic structure of the language. Examples of this include event detection in Turkish (Seker and Diri, 2010), Hindi (Goud et al., 2019), Hungarian (Subecz, 2019) and Swedish (Berglund, 2004).

Much of the work done in event detection in Indian languages is based on events in social media. Rao and Devi

[1] https://www.ethnologue.com/language/kan

[2] https://drive.google.com/drive/folders/11ZXpP4mQcDcM91SKHiSNEtWi_mAkXku7

(2018) has provided a forum dedicated to social media event extraction for Indian languages. Deep learning methods have also been used for a few Indian languages such as Hindi, Tamil and Malayalam (Kuila and Sarkar, 2017). However, these events are based on the ACE definition and analysis of events, which does not consider all event predicates (Ahn, 2006), and views event analysis solely as a task in semantic prediction, without the explicit demarcation and analysis of the surrounding syntax (Ji and Grishman, 2008).

3. Kannada Grammar and Event Representation

In this section, we explore the facets of Kannada grammar that facilitate the representation of events. We begin by considering the notion of a TimeML event. According to Pustejovsky et al. (2003) and Saurí et al. (2006), TimeML defines an event as *a cover term for situations that happen or occur*, as well as *predicates in which something obtains or holds true*. Adopting Goud et al. (2019)'s definition, we consider an event mention as the textual span expected to provide complete information about an event, such as tense, aspect, modality and negation. We also consider the event nugget to be the semantically meaningful unit that expresses the event in a sentence (Mitamura et al., 2015).
Kannada is a free word order, morphologically rich language. However, by convention, verbs usually occur at the end of the sentence. Passive voice is rare. The subject often occurs in nominative case, the object in dative. There are a few primary notions of Kannada syntax which are crucial to event annotation. These include:

- **Kannada is a zero-copula language** (Schiffman, 1979). Copular constructions in Kannada occur without an overt verb (Bhat, 1981). In copular sentences, tense is represented by a modification of the predicate. These predicates are used for copular clauses. As the state is represented by a morphaffixed form of a simple nominal predicate, we do not consider these events at the moment.

1. *nanna* *hesaru* *ujwal*
   my name Ujwal
   My name is Ujwal

- **Every sentence has only one finite or conjugated verb** (Schiffman, 1979). Therefore, sentences with coordinating and subordinating verbs are modified into adjectival and adverbial participle constructions as non-finite verb forms. The tense and aspect information is morphaffixed onto the verb. These adverbial or adjectival participles provide the semantic connotation of an occurrence which describes another action or occurrence, so we annotate these participles in our event mention.

2. *Oorige* *bandidda* *Arjun,*
   town to come had Arjun
   jaatrege *hodanu*
   festival to go
   Arjun, who had come to town, went to the festival.

   Since Kannada has only one verb per sentence, relative clauses are converted into adjectival constructions, which *describe* the verb in the relative clause as a description of the subject of the main verb. Therefore, the sentence "Arjun came to town and went to the festival" can not be translated into Kannada directly. There is no possible mechanism to represent this sentence, other than the inclusion of an adverbial clause to the coordinating verb that occurs semantically prior to the main verb (i.e. is meant to take place before the main verb). This implies a general notion of sequentiality between the main verb and the adjectival construction.

- **Kannada employs tenseless negative forms** (Lindblom, 2014). Negative forms are analytically represented by a single functional negative term. While there are no semantically negative words in Kannada, a single functional negative form is morphaffixed onto the finite verb, or the non-finite adjectival or adverbial participle. Therefore, negations are considered a part of the event mention.

3. *Sumukh* *ootakke* *baralilla*
   Sumukh dinner for didn't come
   Sumukh did not come for dinner

- **Tense, aspect and modality of Kannada verbs** are represented morphologically (Shastri, 2011). Tense and aspect markers are morphaffixed onto both finite and non-finite forms. Therefore, adverbial and adjectival participles have tense, aspect and modality. Therefore, this information is inherently a part of the event mention.

4. *Ram* *tale* *tirugi* *biddanu*
   Ram head spin fell
   Ram, after getting dizzy, fell down

4. Annotation Guidelines

In this section, we provide comprehensive guidelines for the annotation of events in Kannada. Inspired by TimeML, we present these guidelines categorized by the POS of the event nugget. These parts of speech include nouns, finite verbs, non-finite verb constructions such as infinitives, as well as adjectival and adverbial participle constructions.
The TimeML definition of event was used for event annotation in Kannada, following a slight modification based on the changes adopted by Goud et al. (2019). These changes were associated with the analysis and representation of copular constructions as states. Given the morphology of Kannada and the notion of an event nugget being restricted to a lexical or supra-lexical span, we do not annotate copula as events as of now, as Kannada is a zero-copula language, and the representation of such constructions is based on predicate inflections.
In the subsections that follow, we describe the guidelines for annotating events by parts of speech and provide an example for each type. The event of that category is represented in bold in the Kannada transliteration of the sentence.

4.1. Nouns

Nouns can also be events, albeit they occur much more rarely than verbs. Nominal events are abstract nouns which pertain to a temporal phenomenon (i.e. possess a semantically inherent notion of finiteness), such as *yuddha* (war), *kshaama* (famine), *cunaavanegalu* (elections) etc.
For example:

	Bharat	***bandina***	*bagge*
	India	Strike	about

5.

mantrigallu	*yaccharike*	*needidaru*
minister	warning	give

The minister issued a warning about an all India strike.

4.2. Verbs

Verbs are often morphologically marked due to the relaxed constraints on word order, and therefore events can occur at any place in the sentence and can be identified by identifying the main verb. (Veerappan et al., 2011).

4.2.1. Finite Verbs

Finite verbs are considered events as they represent an action that alters that state of the world, possess tense and aspect information, which provides it an inherent notion of temporality.

rajeyannu	*tagedikollalu*	*ramanu*
holiday	take	Ram

6.

arji	***sallisidanu***
application	submit

Ram submitted an application to take a vacation

4.2.2. Non-finite verb forms

Kannada syntax enforces the rule of one finite verb per sentence, all other verb forms are participles (adjectival or adverbial), infinitives and subjunctives. Kannada does not have a gerund verb form.

Adjectival Participle Construction In adjectival participle constructions, the verb is converted into an adjective, so as to describe the noun participating in the main verb by its previous actions. The semantics of the adjectival participle enforces a notion of sequentiality with respect to the main verb, and also represents a notion of finiteness of the action. Furthermore, the adjectival participle is inflected with tense, aspect and modality. Therefore, these constructions are annotated as events.
For example:

7.	*avalu*	***malagalu***	*manege*	*hodalu*
	she is	to bed	home	gone

She slept after going home.

Adverbial Participle Construction Similar to adjectival constructions, the adverbial participle form is used to represent those verbs performed by or associated to the noun in dative (or accusative) case. Here, there is no direct sequentiality applied or associated with the main verb and the adverbial participle.
For example:

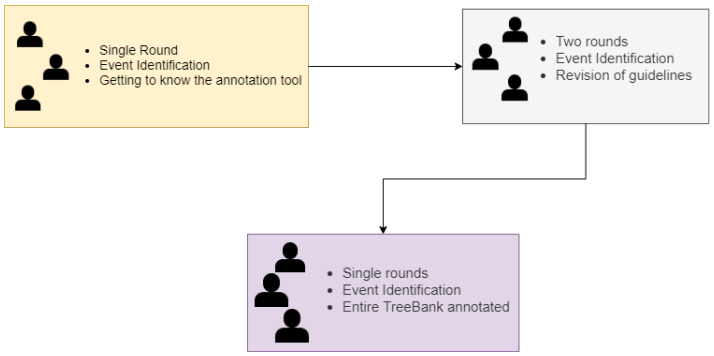

Figure 1: Annotation Procedure

8.	***oodtiruva***	*jinkeyannu*	*betegaaranu*	*hodedanu*
	running	deer	hunter	hit

The hunter shot the running deer.

Infinitives Infinitive verbs in Kannada occurs with the characteristic inflective ending of *'lu'* These infinivitve forms of the verb are also marked as events.
For example:

9.	*naanu*	*iiga*	*aata*	***aadalu***	*hoguttene*
	I'm	now	game	play	go

I'm going to play games now.

Subjunctives The subjunctive is a rare type of verb that expresses something that is desired or imagined. Subjunctives are used to indicate events that are not certain to happen, and hence subjunctive verbs are also annotated as events. Subjunctives can be morphologically inflected for tense, aspect and modality.
For example:

10.	*ninage*	*olleyadagali*	*yendu*	*naanu*	***bayasutene***
	You	be good	that	I	wish

I wish you good.

5. Dataset Preparation and Challenges

In this section, we describe the annotation procedure used to create the Kannada Event Annotated Dataset, and some of the linguistic challenges associated with this task.

5.1. Annotation Procedure

In order to identify events in our corpus, we iteratively annotated the dataset in four rounds. There were three annotators, undergraduate and graduate students annotating this corpus. Each annotator is a native speaker of Kannada and between the ages 20 and 25. We used the BRAT tool for annotating event spans. (Stenetorp et al., 2012).
For the first three rounds of annotation, the sentences were shuffled and each annotator was asked to annotate half the dataset for events, based on a version of the guidelines provided. After each round of annotation, the inter-annotator agreement was calculated based on the *complete overlap*. Two annotations are said to be in complete overlap if both annotations have the same span of text annotated as an event. Partial overlap is not considered for the inter-annotator agreement score.

The first round was dedicated to the annotators getting used to the tool in general, while the second and third rounds was intended to modify the guidelines for event annotation. In the final round of annotation, the dataset was shuffled and the entire Kannada Dependency Treebank was annotated for events by each annotator. As before, annotations were compared and the inter-annotator agreement was computed for *complete overlap*. The final annotations in the dataset presented with this paper consist of that annotation which was agreed upon by the annotators after computing the scores by majority agreement.

5.2. Challenges in Event Annotation

We describe here, in detail, some of the major linguistic challenges faced during annotating events in Kannada. We consider this description to be crucial due to the properties shared by Kannada and other Dravidian languages. We believe that a thorough analysis of these challenges will make the annotation of events in other Dravidian languages easier.

Copular constructions As mentioned in Section 3., Kannada is a zero-copula language. This means that the tense information for existential terms is carried by the predicate of the sentence. However, the semantic equivalent of the verb "to be" is not represented lexically in sentences. This raises the question that for a sentence such as:

11. *karu* *kempu* *bannadallide*
car red colour
The car is red in color.

is "having colour" the event trigger? In this paper, we do not consider this an event nugget, because neither is the event trigger a lexical marker, nor is it consistent with the definition of event nugget for the other syntactic categories of events. Note that in the past tense, the standard copular construction is represented as an adjectival phrase with tense information. Therefore:

12. *karu* *kempu* *baNaDaagittu*
car red coloured
The car was red in color.

Explicator Compound Verbs An explicator compound verb (ECV) is a sequence of two verbs, where the main verb of a sentence is followed by a delexicalized verb in the construction (Abbi and Gopalakrishnan, 1991). South Asian languages show an affinity to ECV constructions for semantic emphasis as well as verbalization of nominal predicates (Kachru and Pandharipande, 1980). An example of such a construction includes:

13. *Ramanu* *ravananannu* ***kondu*** ***hakidanu***
Ram Ravana kill laid
Ram decimated Ravana.

In this sentence, we see that the compound verb *kondu hakidanu* is a single semantic unit which represents the event of "killing". The second word *hakinadu* (literally, "laid to rest") loses its lexical meaning in order to act as an emphasis marker for the main verb *kondu* (literally, "to kill").

Conjunct Verbs Conjunct verbs in Kannada are a combination of a noun or adjective with a verb, such that the verb is delexicalized and serves as a verbalizer to provide a single semantic notion of an occurence or action. The verbalizer is often inflected with the tense, aspect and modality, while the noun contains the semantics of the event. Therefore, while the verb is included in the event span, the noun accounts for the event nugget. For example:

14. *naanu* *raamige* ***bharavase*** ***niDiddene***
I Ram to trust request
I promised Ram.

We see here that the phrase *bharavase niDiddene* is an idiomatic phrase (literally "to request trust") which forms a noun and verb compound which has a single semantic connotation, i.e. "to promise". Therefore it is a compound verb.

6. Generalizing Dataset Development

In this section, we present the procedure and analysis of the development of guidelines for event detection from a holistic, language independent perspective. While the paper thus far depicts the detection and annotation of events, the development of guidelines for event representation, and the associated challenges specifically for Kannada, we would like to emphasize that events are a real-world constructs which are being given a representation in a given language. Therefore, there are some task-specific but language agnostic steps that can be taken for annotating and representing events in morphologically rich languages in general.

As presented above, an understanding of the linguistic philosophy behind event representation is critical to understand how a language's syntax allows for the linking of participants to an action or occurrence. Furthermore, event analysis depends on the morphological or syntactic components associated with marking time, duration, telicity, durativity as well as case relations (Pustejovsky, 1991).

Therefore, the first step in the generalized understanding of events in a morphologically rich language is the isolation of inflections that provide tense, aspect and modality information. The heterogeneity of the markers provides the various possible inflections (and irregular constructions if any) in which an event can take place. While in most language tense and aspect are fairly rigorous (Giorgi and Pianesi, 1997), the modality of verbs and verbal predicates need to be analyzed on a granular level.

Morante and Sporleder (2012) presents a thorough study into annotation and corpus linguistics into the role of modality and negation as extra-propositional aspects. Indeed, in event annotation, both negation and modality play a role in the complete description of an event. However, languages vary in their representation of modal verbs and negative polarity, and therefore, development of guidelines for these event features becomes a language specific problem. Indeed, while the guidelines developed and challenges faced in Section 4. are to be seen, if development in event detection takes place in Telugu, Tamil and other Dravidian languages, it should be noted that modality and polarity are represented differently in each of these languages. Therefore, it is one of the major challenges in event representation.

Data Type	Total Number
Tokens	37020
Sentences	3583

Table 1: Corpus Statistics

Event type	Total Number of Events
Single Word Events	3114
Multi Word Events	1686

Table 2: Event Type Statistics

7. Corpus Statistics

In this section, we explain the corpus, and some basic statistics associated with it, including the dataset size, the number of events, and their category-wise distribution. Finally, we consider the computation of the inter-annotator agreement and show that the dataset is in fact quite reliable, as it shows a high Fliess' Kappa Score of 0.91 in the final round of annotation.

We annotated the Kannada Dependency Treebank (Rao et al., 2014)[3] for event mentions. The Dependency Treebank corpus consists of 37,020 tokens, from distinct domains such as tourism, general and conversational.

As presented in Table 1 the dataset is divided into 3,583 sentences. The dataset has been annotated with 4,800 events, out of which 3,114 events consist of a single word in the event span, while 1,686 events have a multiword event span as seen in Table 2. There are 686 sentences which do not have any events as they are entirely copular in nature. This implies that sentences with multiple events are not uncommon in general.

In all the rounds of annotation mentioned in section 5., inter-annotator agreement was computed using the Fleiss' Kappa metric for multiple annotators (Fleiss, 1971), where the categories for annotation are 1 for complete match and 0 otherwise. Fleiss Kappa score is computed as follows:

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e} \tag{1}$$

where $P - P_e$ is the actual degree of agreement achieved and $1 - P_e$ is the degree of agreement above chance. Given N tokens to be annotated and n annotators, with k categories to annotate the data. We first calculate the proportion of annotations in the j category as:

$$p_j = \frac{1}{Nn} \sum_{i=1}^{N} n_{ij} \ , \quad 1 = \sum_{j=1}^{k} p_j \tag{2}$$

We then calculate P_i, the degree of agreement with the ith annotator as:

$$P_i = \frac{1}{n(n-1)} \sum_{j=1}^{k} n_{ij}(n_{ij} - 1) \tag{3}$$

$$= \frac{1}{n(n-1)} \left[\left(\sum_{j=1}^{k} n_{ij}^2 \right) - n \right] \tag{4}$$

Finally we calculate $\bar{P}$ and $\bar{P}_e$ as:

$$\bar{P} = \frac{1}{N} \sum_{i=1}^{N} P_i \tag{5}$$

$$\bar{P}_e = \sum_{j=1}^{k} p_j^2 \tag{6}$$

The inter-annotator agreement in the final round of annotations is 0.91, after four rounds of annotation. We noted that the inter-annotator agreement increased across the stages.

8. Conclusion

In this paper, we present a comprehensive set of guidelines for the annotation of events in Kannada, based on TimeML guidelines of event detection. Using these guidelines, we annotate the Kannada Dependency Treebank with event nuggets. This dataset is the first attempt to perform semantic annotation tasks in Kannada, which is a low-resource language. We introduce some basic features of Kannada grammar associated with representing event information, as well as some of the challenges in detecting events in this language.

In the future, this dataset can be expanded both in size and in annotations to include other facets of the TimeML annotation schema, and provide further insight into the automated detection of events and other temporal information. Given the nature of the guidelines, challenges and the description of the corpus annotation procedure, we hope that the development of event annotated corpora for other Dravidian languages becomes easier.

9. Bibliographical References

Abbi, A. and Gopalakrishnan, D. (1991). Semantics of explicator compound verbs in south asian languages. *Language Sciences*, 13(2):161–180.

Ahn, D. (2006). The stages of event extraction. In *Proceedings of the Workshop on Annotating and Reasoning about Time and Events*, pages 1–8.

Berglund, A. (2004). Extracting temporal information and ordering events for swedish. *Master's thesis report*.

Bhat, V. (1981). The copula in kannada. *Papers in Linguistics.(Festschrift for RC Hiremath), Sharat Prakashan, Mysore*.

Bittar, A., Amsili, P., Denis, P., and Danlos, L. (2011). French timebank: an iso-timeml annotated reference corpus. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 130–134. Association for Computational Linguistics.

[3] https://tdil-dc.in/index.php?option=com_download&task=showresourceDetails&toolid=1979&lang=en

Caselli, T., Lenzi, V. B., Sprugnoli, R., Pianta, E., and Prodanof, I. (2011). Annotating events, temporal expressions and relations in italian: the it-timeml experience for the ita-timebank. In *Proceedings of the 5th Linguistic Annotation Workshop*, pages 143–151. Association for Computational Linguistics.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378.

Forăscu, C. and Tufiş, D. (2012). Romanian timebank: An annotated parallel corpus for temporal information. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 3762–3766.

Giorgi, A. and Pianesi, F. (1997). *Tense and aspect: From semantics to morphosyntax*. Oxford University Press on Demand.

Goud, J. S., Goel, P., Debnath, A., Prabhu, S., and Shrivastava, M. (2019). A semantico-syntactic approach to event-mention detection and extraction in hindi. In *Workshop on Interoperable Semantic Annotation (ISA-15)*, page 63.

Ji, H. and Grishman, R. (2008). Refining event extraction through cross-document inference. In *Proceedings of ACL-08: Hlt*, pages 254–262.

Kachru, Y. and Pandharipande, R. (1980). Towards a typology of compound verbs in south asian languages. *Studies in the Linguistic Sciences*, 10(1):113–124.

Kittel, F. (1993). *A Grammar of the Kannada Language: Comprising the Three Dialects of the Language (ancient, Medieval and Modern)*. Asian Educational Services.

Kuila, A. and Sarkar, S. (2017). An event extraction system via neural networks. In *FIRE (Working Notes)*, pages 136–139.

Lindblom, C. (2014). Negation in dravidian languages: A descriptive typological study on verbal and non-verbal negation in simple declarative sentences.

Mallamma, V. R. and Hanumanthappa, M. (2014). Semantical and syntactical analysis of nlp. *International Journal of Computer Science and Information Technologies*, 5(3):3236–3238.

Mitamura, T., Yamakawa, Y., Holm, S., Song, Z., Bies, A., Kulick, S., and Strassel, S. (2015). Event nugget annotation: Processes and issues. In *Proceedings of the The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation*, pages 66–76.

Morante, R. and Sporleder, C. (2012). Modality and negation: An introduction to the special issue. *Computational linguistics*, 38(2):223–260.

Pustejovsky, J., Castano, J. M., Ingria, R., Sauri, R., Gaizauskas, R. J., Setzer, A., Katz, G., and Radev, D. R. (2003). Timeml: Robust specification of event and temporal expressions in text. *New directions in question answering*, 3:28–34.

Pustejovsky, J. (1991). The syntax of event structure. *cognition*, 41(1-3):47–81.

Rao, P. R. and Devi, S. L. (2018). Eventxtract-il: Event extraction from newswires and social media text in indian languages@ fire 2018-an overview. In *FIRE (Working Notes)*, pages 282–290.

Rao, A. B., Murali Krishna, S., and Nayak, A. (2014). Developing a dependency treebank for kannada. *International Journal of Engineering Sciences and Research*, 15:5–7.

Saurí, R., Littman, J., Knippen, B., Gaizauskas, R., Setzer, A., and Pustejovsky, J. (2006). Timeml annotation guidelines. *Version*, 1(1):31.

Saurı, R. (2010). Annotating temporal relations in catalan and spanish timeml annotation guidelines. Technical report, Technical Report BM 2010-04, Barcelona Media.

Schiffman, H. (1979). A reference grammar of spoken kannada.

Seker, S. E. and Diri, B. (2010). Timeml and turkish temporal logic. In *IC-AI*, volume 10, pages 881–887.

Shastri, G. (2011). Kannada morphological analyser and generator using trie. *IJCSNS*, 11(1):112.

Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., and Tsujii, J. (2012). Brat: a web-based tool for nlp-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 102–107. Association for Computational Linguistics.

Subecz, Z. (2019). Event detection and classification in hungarian natural texts. *European Scientific Journal July*.

Veerappan, R., Antony, P., Saravanan, S., and Soman, K. (2011). A rule based kannada morphological analyzer and generator using finite state transducer. *International Journal of Computer Applications*, 27(10):45–52.