

The Expression of Information Structure in Malayalam

Amanda Swenson, Athulya Aravind, Paul Marty (MIT)

Introduction

Human linguistic communication differs from the communication of other animal species by its unparalleled communicative power. This communicative power lies in two essential dimensions of human linguistic communication: it is symbolic and it is grammatical. By means of linguistic symbols, one human being can *share information* with other human beings by directing their attentional states to objects, events and situations in the outside world. By means of grammatical mechanisms (e.g., intonation, syntactic constructions), human beings can *package information* in order to optimize the form of their messages relative to the current attentional state of their interlocutors. To illustrate these two dimensions, consider the following dialogue:

- (1) A: Dileep enthu cheythu?
 Dileep what did
 ‘What did Dileep do?’
 B: Avan Anu-inte sahodari-ye nulli.
 He Anu-GEN sister-ACC pinched
 ‘He pinched Anu’s sister.’

In (1), B’s reply meets the communicative needs on the side of A: it answers A’s inquiry, and adds factual information to the common ground (CG) between both speakers (cf. Stalnaker 1974). That is, B shares with A the information that ‘Dileep pinched Anu’s sister’, which is now mutually known to be shared. But B’s reply also packages this information in a principled way. Part of this information is backgrounded, i.e. the proposition that ‘Anu has a sister’ (presupposition), while another part of it is foregrounded, i.e. the proposition that ‘Dileep pinched her’ (assertion).

In linguistics, we refer to the formal study of the way information is packaged within a sentence as Information Structure (IS). Precisely, following Chafe (1976), IS refers to the packaging of information that meets the temporary communicative needs of the interlocutors. This includes those features of linguistic constructions and productions whereby a speaker optimizes the transfer of information relative to the momentary state of the addressee’s mind and the question under discussion. The goal of this short article is to provide an overview of the possible avenues for further investigating the expression of IS in Malayalam, with a particular emphasis on the role of prosody. For these purposes, we draw our attention to two basic IS roles previously discussed in the literature: topic and focus.

The notion of ‘topic’ (T) standardly comes with a complementary notion called ‘comment’ (C). The idea behind the notion of topic, originally developed in Reinhart (1981), is that when a new piece of information is shared by two speakers, it is structurally associated with an entity, namely the entity that speakers identify as the entity which the information, the comment, is given about. Consider for instance the following pair of sentences:

- (2) a. [Dileep]_T [Anu-vine vivaham kazhichu]_C
 Dileep [Anu-ACC wedding did]
 b. [Anuvine]_T [Dileep vivaham kazhichu]_C

The sentences in (2a) and (2b) express the same proposition: ‘Dileep married Anu.’ Nonetheless, they structure this information differently. A natural way to interpret (2a) is to understand it as an answer to ‘Who did Dileep marry?’, and therefore as a comment about Dileep. By contrast, (2b) is likely to be understood as an answer to ‘Who did Anu marry?’, that is as a comment about ‘Anu.’ Following Reinhart, (2a) should thus be stored as information about ‘Dileep,’ while (2b) should be stored as information about ‘Anu.’ These intuitions are captured by the following definition:

(3) Definition of Topic (based on Reinhart 1981; Féry and Krifka 2008):

The topic constituent identifies the entity (or set of entities) under which the information expressed in the comment constituent should be stored in the CG content.

In respect to this, it's important to note that not every constituents can act as a topic. For instance, non-referring expressions such as indefinite NPs are difficult to be construed as topics. Presumably, topical constituents have to be able to refer to an entity. This connection between topicality and referentiality is sometimes referred to as the ‘definiteness’ or ‘presupposition’ effect of topics.

The notion of ‘focus’ (F), on the other hand, relates to the highlighting of new information. The association between focus and prosodic prominence has been shown to hold in a variety of typologically and genetically diverse languages, and is widely believed to be universal (Gundel and Fretheim 2004). Consider the following example:

- (4) A: Nammal-ude koottukaar-il aare War and Peace vaayich-ittunde?
 we-GEN friends-LOC who War and Peace read-PERFECT
 ‘Who among our friends has read War and Peace?’
 a. B: [DILEEP]_F War and Peace vaayicch-ittunde.
 b. B: #Dileep [WAR AND PEACE]_F vaayicch-ittunde.

In (4), A's question communicates a particular informational expectation: A is asking B about who, out of all their friends, read *War and Peace*. While both (4a) and (4b) have the same truth-conditions, only B's answer in (4a) is felicitous in this context. In (4a), A's expectation is recognized by the focus on the subject-NP 'Dileep', which is linguistically expressed by prosodically prominent phrase accent (represented using all capital letters). By contrast, in (4b), the focus is on the object-NP '*War and Peace*', leading then to an infelicitous answer. A general definition of focus is provided in (5).

(5) Definition of Focus (based on Rooth 1992):

Focus indicates the presence of alternatives (i.e., other statements that could have been uttered instead) that are relevant for the interpretation of linguistic expressions.

It is worth pointing out that the articulation of the definitions in (3) and (5) can be refined to include further extensions of the two basic IS categories we are interested in. For instance, combining the notion of topic and focus can encompass other topichood-related notions such as contrastive topic, as illustrated in (6). A contrastive topic (CT) can be thought of as indicating (via some prosodic focus) the presence of alternative topics at the current point of the discourse context. As noted by Buring (2003), contrastive topics are expected whenever a question under discussion cannot be answered on the basis of one single topic. In (6), for instance, A's question sets up one single complex topic, namely 'Anu and Asha'; but the fact that 'Anu' and 'Asha' read different books leads B to tease these entities apart as distinct subtopics.

- (6) A: Anu-vum Asha-yum enthu vaayicchu?
 Anu-CONJ Asha-CONJ what read
 'What did Anu and Asha read?'
 B: [Anu]_{CT} War and Peace vaayicchu. [Asha]_{CT} Anna Karenina vaayicchu.
 'Anu read *War and Peace*. Asha read *Anna Karenina*.'

Similarly, the notion of contrastive focus, as exemplified in (7), can be defined as a subtype of focus in which the alternatives are given explicitly by the utterance context (e.g., tea and coffee), and where only one of the contrasted alternatives is acceptable (e.g., tea or coffee, but not both).

- (7) A: nina-kku CHAAYA aan-oo KAAPPI aan-oo vend-athe?
 You-DAT tea COP-DISJ coffee COP-DISJ want-NOMLZ
 'Do you want TEA or COFFEE?' (lit. 'Is it TEA or COFFEE that you want?')
 B: eni-kku [CHAAYA]_F aanu vend-athe.
 I-DAT tea COP want-NOMLZ
 'I want TEA.' (lit. 'It is TEA that I want.')

In what follows, we investigate the linguistic reflexes of topic and focus in the grammar of Malayalam by emphasizing the contributions of prosody (i.e., changes in intonation, pitch modifications, etc). We will use the ToBi system for the transcription of prosody throughout our discussion (cf. Beckman & Elam 1997, Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Ladd 2000). As we shall see, however, Malayalam has a wide range of linguistic devices for expressing both of these IS roles. Beyond intonation, Malayalam can further express topic and focus using syntactic mechanisms (e.g., scrambling, clefts) and morphological mechanisms (e.g., ‘as for’-phrases, focus marker *-ee + Verb-ulluu*), as well as focus-sensitive particles (e.g., *maathram*). In this article, we will provide examples of several ways prosodic and morphosyntactic structures interact to encode information structure.

Information Structure at the Syntax-Prosody Interface

First, let us turn to *scrambling*, a general term for the way in which non-canonical word orders are derived in free word order languages. Mohanan (1982) notes that, while the basic word order in Malayalam is SOV, a sentence like *kutti aanaye kandu* (‘The child saw the elephant’) can have six possible orders: SOV, OSV, SVO, OVS, VSO and VOS. The following question arises: do the different orders encode different information structures? The answer seems to be affirmative.

When we introduced the notion of topic in (2), we saw that the constituents identified as sentence topics were in a sentence initial position. Asher and Kumari (1997) and Mathew (2014) have proposed that the sentence initial position is a default topic position. If this position is restricted to topics, then we expect constituents that serve as answers to a question to be barred from this position. Examples (8)-(9) confirm this expectation and show that the sentence initial position is indeed a topic position in Malayalam. In (8), *Rajan*, and in (9), *raajaavu*, are identified as topics by the questions. Thus, additional information elicited by the questions should be stored under them in the CG. They are both allowed in the sentence initial position, (8a) and (9a). However, the new information, i.e. *raajaavu* in (8) and *Rajan* in (9), cannot appear in the sentence initial position because this information is part of the comment that will be stored under the topic headings of *Rajan* and *raajaavu*, respectively. This supports the idea that the sentence initial position is restricted to topics.

- (8) Who is Rajan?
 a. [Rajan]_T [raajaavu]_C.
 Rajan king
 ‘Rajan is the king’
 b. #[raajaavu]_C [Rajan]_T

- (9) Who is the king?
 a. [raajaavu]_T [Rajan]_C.
 b. #[Rajan]_C [raajaavu]_T.
 (Mathew 2014, p24: 22-27)

As additional support for the sentence initial position being a topic position, Mathew points out that indefinites cannot usually occur as the initial element in a sentence. This is not unexpected, given that indefinites are often not interpreted referentially and there seems to be a link between being referential and having the ability to be a topic. In (10a), we see that an indefinite object like *randil adhikam puuchakale* cannot move to the beginning of the sentence. However, when *randil adhikam* is replaced with *aa* to make the phrase definite, the phrase is allowed to move to the sentence initial position, (10b).

- (10) a. *randil adhikam puuchakal-e Rajan Priyay-ku koduthu
 two more.than_[non-specific indefinite] cats-ACC Rajan Priya-DAT gave
 ‘*Rajan gave more than two cats to Priya.’
 b. aa puuchakal-e Rajan Priyay-ku koduthu
 ‘Rajan gave those cats to Priya.’

The facts in (8)-(10) suggest that an OSV word order is derived from an SOV word order by movement of the object to a topic position in the left periphery of the sentence. The analysis of the prosodic structure provides support for this hypothesis. Specifically, the spectrogram for an SOV sentence like *kutti aanaye kandu* shows there is a single intonational phrase (IP), Break Index 4, with two high (H) pitch accents: a nuclear H pitch accent on the subject (bold font), and a downstepped H pitch accent on the object (labeled !H). The IP ends with a L-L% boundary tone, corresponding to a falling of the fundamental frequency F0. By contrast, an OSV sentence like *aanaye kutti kandu* is made up of two intermediate phrases (iP; Break Index 3), namely [*aanaye*] and [*kutti kandu*]. While both iPs bear a H phrase accent, this accent is more prominent in the second iP than in the first iP (i.e., higher F0 maximal value), suggesting that the latter is in fact the nuclear accent of the whole IP. This observation is consistent with the idea that the Comment (i.e., *kutti kandu*), bringing about new information, is prosodically more emphasized than the Topic (i.e., *aanaye*). ToBi transcriptions for SOV and OSV sentences are given in (11) below. In (11b), the first iP consists of the object that has been moved to a topic position in the left periphery, and the second iP consists of the subject and the verb.

- (11) a. [kutti aanaye kandu]_{iP} SOV
 H* !H* L-L%

b. [[aanaye]_{iP} [kutti kandu]_{iP}]_{IP} OSV
 H* H- H* L-L%

At this point, one might ask whether we can extend a similar kind of analysis (i.e., a movement analysis with information structure implications and prosodic correlates) to other derived word orders. Tentatively, we suggest that such an extension seems plausible. Mohanan (1982) and Mathew (2014) note that placing the verb in a non-final position is marked and requires special contexts. Mohanan suggestively comments that non-final verbs are strongly contrastive. We note preliminarily that the prosodic structures of these sentences seem to reflect the work of scrambling and to correlate to some extent with Mohanan's suggestion. Instances of the SVO, OVS and VOS word orders with their ToBi transcriptions are provided in (12).

- (12)
- a. [[kutti kandu]_{iP} [aanaye]_{iP}]_{IP} SVO
 L+H* L- H* L-L%
- b. [[aanaye kandu]_{iP} [kutti]_{iP}]_{IP} OVS
 L+H* L- H* L-L%
- c. [[kandu aanaye]_{iP} [kutti]_{iP}]_{IP} VOS
 L+H* L- H* L-L%

In these sentences, the first iP bears a bitonial L+H* nuclear accent: the pitch track shows a gradual fall from the beginning of the first prosodic word, reaching a low at the end of its first syllable, then a sharp rise into the High tone of the second syllable. This type of pitch accent has been found in many languages to be associated with a contrastive topic interpretation (cf. Constant 2014). This differs from the intonation we observed in OSV sentences such as (11b) where a H* (non-nuclear) pitch accent is anchored to the topicalized element. We said in the introduction that contrastive topics can be thought of as a subtype of topics involving focus. The appearance of a contrastive topic prosody in these examples is in line with Mohanan's comment that sentences with verbs in non-final positions tend to have a contrastive meaning. Discerning the specific IS of each word order and the movements involved will require further careful probing of the contexts which license the respective word order and additional congruency tests.

A second syntactic device with information structure implications is the *cleft*-construction. The term *cleft* refers to a particular syntactic mechanism by which languages express exhaustive focus. Clefts are typically copulative constructions with a discourse prominent portion, the 'focus', and a given portion. As first described in Madhavan (1987), clefts in Malayalam involve the copula *aanu*, an open proposition (i.e., a clause with a gap) headed by the morpheme *athu*, and the focused constituent. These constructions in Malayalam behave like clefts

cross-linguistically with respect to their interpretive effects. Clefts are associated with two presuppositions: (i) exhaustive identification/maximality, and (ii) existence. In (13), we see that the focused element, *Vinu*, is interpreted as exhaustively identifying the individual who likes Deepa's sister; consequently, the following sentence, which expresses that Suman also likes her, is infelicitous.

- (13) *Vinu* aanu Deepa-yute sahodari-ye snehikkunn-athu. #Suman-um aval-e snehikkunnu
Vinu COP Deepa-GEN sister-ACC love-NOMLZ. Suman-also she-ACC love.
 'It is *Vinu* who loves Deepa's sister. #Suman also loves her.'

The use of a cleft also leads to an inference that there exists some individual that satisfies the predicate denoted by the cleft clause. The oddity of (14) is due to the violation of this inference:

- (14) #Aarum alla Deepa-yute sahodari-ye snehikkunn-athu.
 Anyone COP.NEG Deepa-GEN sister-ACC love-NOMLZ
 '#It's no one that loves Deepa's sister.'

As with most foci, the clefted constituent must correspond to new information. This is illustrated once again with the congruency test from above. Example (15a) asks about the individual that *Vinu* likes, but in (15b), the clefted constituent, which we expect to encode new information, does not actually provide an answer to this question.

- (15) a. Aare aanu *Vinu* snehikkunn-athu?
 Who COP *Vinu* love-NOMLZ
 'Who is it that *Vinu* loves?'
 b. #*Vinu* aanu Deepa-yute sahodari-ye snehikkunn-athu
Vinu COP Deepa-GEN sister-ACC love-NOMLZ
 '#It's *Vinu* that loves Deepa's sister.'

In addition, the information denoted by the cleft-clause must already be part of the common ground. However, the cleft-clause in (15b) contains new information, namely the identification of Deepa's sister as the recipient of *Vinu*'s affection. As a result, the dialogue in (15) above is, in a sense, doubly infelicitous.

Prosodically, a sentence like (16a) has a single IP which is made up of three main pieces. The ToBi transcription of (16a) is provided in (16b). Firstly, it involves the alignment of a nuclear pitch accent with the prominent syllable of the clefted constituent *Vinu*. Here, *Vinu* bears a high pitch accent (H*), which is realized by a local maximum in the fundamental frequency (F0) of

the speech signal. Secondly, the copula *aanu* bears a prominent bitonal L*+H pitch accent: a Low target is associated with the first syllable of *aanu*, and then followed by a rise to a peak, corresponding to the +H portion of the pitch accent. Thirdly, the sentence in (16a) has a Low phrase tone (L-) associated with the stretch of material between *Vinu aanu* and the end of the IP. We notice that, despite the presence of some F0 peaks on *Deepa* and *sahodari*, probably reflecting stress assignment for these prosodic words, the melody for this part of the utterance looks essentially deaccented. This behavior might be an instance of the phenomenon of post-nuclear deaccenting, sometimes used to derive the asymmetry between pre-focal and post-focal material. While non focus-marked material preceding focus can bear (prominent) pitch accents, cf. (11b), post-focal material cannot, except in contexts with multiple foci. Hence, post-focal material often appears to show a steady declining F0 that can cover many words. At the phonetic level, this phenomenon corresponds fairly directly to pitch range compression: pitch range is expanded on focus to reach a local maximum F0, but compressed after it. In the present case, the cleft component of the sentence, i.e. *Vinu aanu*, is marked by phrasal stress prominence, which is realized by greater intensity and duration, and followed by post-focal pitch range compression. This prosodic pattern is commonly found in constructions that encode exhaustive focus as we will see in more detail in the sections to come.

- (16) a. *Vinu aanu Deepa-yute sahodari-ye sneehikkunn-athu.*
 Vinu COP Deepa-GEN sister-ACC love-NOMLZ.
 ‘It is Vinu who loves Deepa’s sister.’
 b. [*Vinu aanu Deepa-yute sahodari-ye sneehikkunn-athu*]_{IP}
 H* L*+H (!H*) (!H*) L-L%

Morphological Markers of IS and their Prosody

In addition to syntactic mechanisms, Malayalam can also encode IS roles with morphological mechanisms. First, we will focus on those mechanisms that encode topic notions. Asher and Kumar (1997) recognize the forms *aanu engil*, (17a), and *ennu veccaal*, (17b), as expressing the notion of topic. An additional form, *kaaryathil*, (17c), can also be used.

- (17) a. *Raaman aan-engil bhayangkara matiyan aanu*
 Raman COP-COND extremely lazy-MASC COP
 ‘As for Raman, he is very lazy.’
 b. *kutti-kk-aan-ennu veccaal nalla pani undu*
 child-DAT-COP-COMP COND good fever be
 ‘As for the child, (he) has a high fever.’
 (Asher & Kumari 1997, p184: 897, 901)

- c. [Asha-yude kaariyath-il aval-ude vesham bhangiyullathalla ennu] Unni wicharikkunnu.
 Asha-GEN matter-LOC she-GEN dress ugly COMP Unni thinks
 ‘As for Asha, Unni thinks her dresses are ugly.’

As topic marking constructions, the markers in (17) obey the same types of congruency tests show in (8)-(9). For example, (17a) could be the answer to the question “What type of person is Raman?” but not “Who is a lazy person?”. Reinhart (1981) identifies ‘as for’ left-dislocation in English as a test for determining the aboutness topic of a sentence. The constructions in (17) serve this function in Malayalam: (17a) is about *Raman*, (17b) is about the *kutti* and (17c) is about *Asha*. Additionally, as Mathew (2014) points out, *aanengil* exhibits the expected definiteness effects, (18). We see in (18a) that a non-specific indefinite cannot be marked as a topic by *aanengil* where as in (18b) we see that a definite can be marked as a topic

- (18) a. *randil adhikam puuchakal-e aan-engil Rajan Priyay-ku koduthu.
 two more.than cats-ACC COP-CONDT Rajan Priya-DAT gave
 ‘As for more than two cats, Rajan gave (them) to Priya.’
- b. aa puuchakal-e aan-engil Rajan Priyay-ku koduthu.
 those cats-ACC COP-CONDT Rajan Priya-DAT gave
 ‘As for those cats, Rajan gave (it) to Priya.’

The same facts hold for *veccaal* and *kaaryathil*. Furthermore, we see that in Malayalam, just as in English ‘as for’ constructions, the markers in (17) require some contrastive topic context. To see this, imagine the following scenario¹. A professor is teaching a class on nanosyntax. On the first day of class, the professor walks into the room and the first thing out of his/her mouth is one of the sentences in (19). In this scenario, all of the sentences in (19) are odd things to say since there is nothing else in the CG with which we can contrast *nanosyntax*.

- (19) a. nanosyntax aan-engil, mikka alkarum ath-ine-kuriccu kettitte illa.
 nanosyntax COP-CONDT most people it-ACC-about heard NEG
 ‘As for nanosyntax, most people have never heard of it.’
- b. nanosyntax aan-ennu veccaal, mikka alkarum ath-ine-kuriccu kettitte illa.
 nanosyntax COP-COMP CONDT most people it-ACC-about heard NEG
 ‘As for nanosyntax, most people have never heard of it.’

¹ Thanks to Noah Constant (p.c.) for suggesting this scenario.

- c. nanosyntax-inte kaaryath-il, mikka alkarum ath-ine-kuriccu kettitte illa.
 nanosyntax-GEN matter-LOC most people it-ACC-about heard NEG
 ‘As for nanosyntax, most people have never even heard of it.’

The prosodic analysis further supports a contrastive topic interpretation of these sentences. Work on English has identified contrastive topics as being associated with the L+H* pitch accent (cf. Constant 2014). The ToBi transcriptions of the sentences in (17) are given in (20). In (20a) and (20c), we observe an L+H* pitch accent on the NP marked by the topic marker, followed by a prosodic break (Break Index 3) which is indicative of a distinct intermediate intonational phrase. This L+H* pitch accent is also observed in (20b); however, it is anchored this time to (part of) the topic marker *veccaal*. We also notice that no major prosodic break was found in this sentence. The reason for these differences between (20a, c) and (20b) remains an open question.

- (20) a. [[Raaman aan-engil]_{IP} [bhayangkara matiyan aanu]_{IP}]_{IP}
 L+H* L* H- H* L-L%
- b. [kutti-kk-aan-ennu veccaal naala pani undu]_{IP}
 H* L+H* L+!H* L-L%
- c. [[Ashayude kaariyathil]_{IP} [avalude vesham]_{IP} [bhangiyullathalla ennu Unni wicharikkunnu]_{IP}]_{IP}
 L+H* L- H* L- H* !H* L-L%

Based on these data, as well as the data we saw for non-verb final sentences in (12), it looks like contrastive topics in Malayalam are marked with the same type of prosodic cues as contrastive topics in English. It also supports that the constructions in (17) encode additional contrastive meaning beyond that of topics in OSV sentences (recall that the topicalized object here forms an intermediate intonational phrase with a H* pitch accent and a H- final phrase accent). Gaining a better understanding of the function of each of these morphological topic marking constructions and how a single morphological marker can encode multiple types of topics is an interesting avenue for future work.

Malayalam has two morphological strategies for marking focus: *maathram*, *-ee +Verb-ulluu*. Examples of these strategies being used to focus the subject can be found in (21a-b). We see that both of these strategies encode exhaustive focus in (21c-d): *Deepa* is interpreted as exhaustively identifying the individual who likes *Vinu*; consequently, the following sentence, which expresses that *Asha* also likes him, is infelicitous.

- (21) a. Vinu maathram Deepa-yute sahodari-ye sneehikkunnu.
 Vinu only Deepa-GEN sister-ACC love
 ‘Only Vinu loves Deepa’s sister.’

- b. Vin-ee Deepa-yute sahodari-ye sneehikkunn-ull-uu
 Vinu-ee Deepa-GEN sister-ACC love-be-uu
 ‘~Only Vinu loves Deepa’s sister.’
- c. #Deepa maathram vinu-vine sneehikkunnu. Asha-yum sneehikkunnu.
 Deepa only Vinu-ACC loves. Asha-CONJ loves
 ‘Only Deepa loves Vinu. And so does Asha.’
- d. #Deepa-yee vinu-vine sneehikkunn-ull-uu. Asha-yum sneehikkunnu.
 Deepa-ee Vinu-ACC love-be-uu Asha-CONJ love
 ‘~Only Deepa loves Vinu. And so does Asha.’

While Mathew (2014) translates both markers as ‘only’ in her thesis, she suggests that *maathram* best corresponds to the semantics of English ‘only’ and that *-ee V+ulluu* serves a slightly different function. She suggests that its function is that of contrastive focus, as defined in Rizzi (2013): “Contrastive focus introduces new information that contrasts with some natural expectation imputed to the interlocutor” (p56). However, given a sentence like (21b), it’s not clear to us what the “natural expectation” is here, as there are no overt alternatives given in the sentence as there were in the ‘TEA or COFFEE’ case in (7). However, the intuition that *maathram* and *-ee V+ulluu* have distinct semantics seems to be on the right track. Evidence in favor of this analysis comes from the fact that both markers can simultaneously and non-redundantly be used to focus a single constituent in a given sentence. In (22) both *maathram* and *-ee V+ulluu* target the subject, *Vinu*. If they encoded the same semantics, we would expect the sentence to have a redundant feel (e.g. *Only only Vinu likes Deepa’s sister*), yet it does not.

- (22) Vinu maathram-ee Deepa-yute sahodari-ye sneehikkunn-ull-uu.
 Vinu only-ee Deepa-GEN sister-ACC like-be-uu
 ‘Only Vinu likes Deepa’s sister.’

Prosodically, the sentences in (21a) and (21b) show strong similarities. Their ToBi transcriptions are given in (23). They both have prosodic prominence on the subject (i.e., a nuclear H pitch accent associated with vowel lengthening) followed by pitch compression that drops to a L-L% final phrase boundary tone. This prosodic pattern is similar to the one induced by focus sensitive particles such as *only* (cf. association with focus). Note that, in (23a), *maathram* also bears a bitonal L*+H accent (similar to the one we found on the copula *aanu* in (16b)).

- (23) a. [Vinu maathram Deepayute sahodariye sneehikkunnu]_{IP}
 H* L*+H (!H*) (!H*) L-L%

feliculously follow (24a). As such, it patterns with *maathram* and not the cleft, suggesting that exhaustivity is not presupposed, but asserted in the *-ee V+ulluu* construction.

We see in (25a) that clefts can non-redundantly be used with *maathram*. Given that the semantics of clefts differs subtly from that of *maathram*, specifically with respect to what is presupposed versus asserted, it's not surprising that they both can be used to focus a single constituent in a non-redundant matter. However, (25b) reveals that *-ee V+ulluu* cannot be used in a cleft-sentence. This is despite the fact that in (24b) and (24d), we saw that the two constructions pattern together in meaning. The ungrammaticality of (25b) thus leaves us with a puzzle for further investigation.

- (25) a. Vinu maathram aanu Deepa-yute sahodari-ye sneehikkunn-athu.
Vinu only COP Deepa-GEN sister-ACC love-NOMLZ
'It is only Vinu who loves Deepa's sister.'
- b. *Vinu-ee aan-ull-uu Deepa-yute sahodari-ye sneehikkunn-athu.
Vinu-ee COP-be-uu Deepa-GEN sister-ACC love-NOMLZ
'~It is only Vinu who loves Deepa's sister.'

Turning now to the prosody, we would like to point out that, based on the data we have analyzed so far, it appears that clefts (16a), *maathram*, and *-ee V+ulluu* sentences have strong similarities in their overall prosodic structures. We have suggested along these lines that these similarities might tell us something about the prosodic correlates of exhaustive focus. Nonetheless, we have also seen that there exist noticeable differences regarding the specifics of the IS for these sentences (cf. what is asserted vs. presupposed for the different focus strategies). At this point, we have no evidence that these fine-grained differences are reflected in the prosody. This (temporary) conclusion is consistent with the general agreement that there is not one-to-one mapping between IS specific phonological and syntactic categories (cf. Féry 2007).

Conclusion

In this article we have surveyed the ways in which the various interfaces interact to package information. We have seen how syntactic notions like clefts and focus markers like *mathram* and *-ee V+ulluu* interact with each other and with prosody to express focus in similar but different ways. We have also seen that the notion of topic is encoded via syntactic processes such as scrambling, topic markers and different prosodic patterns. There are many open questions for future research. A few examples include: figuring out the precise compositional semantics for the *-ee V+ulluu* construction and the interaction of different focus strategies, gaining a better understanding of the movements involved in different types of scrambling and

their information structure implications, and conducting more, in depth research on the contributions of prosody to information structure.

Acknowledgements

We would like to thank all those who recorded sentences for us: Dr. K.S. Chandraseker, Josefreeda Edwin, Susan George, Sudha Jacob, Juhi Jan James, Lopin K.L., Nithin M., Shijith S, Sonu S., Aiswaria G. Shajan and T.S. Varghese. We also want to thank everyone in the Phonetics Lab at the University of Kerala for all of their help along the way. This work has been supported by a MISTI/MIT-India internship grant, and a Ken Hale grant for linguistic fieldwork. All errors are our own, as ever.

References

- Asher, R.E. & T.C. Kumari. 1997. *Malayalam*. Routledge: London and New York.
- Beckman, M. & G. Elam. 1997. Guidelines for ToBi labelling. *The OSU Research Foundation*, 3.
- Beckman, M. & J. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook*, 3(1), 5-70.
- Buring, D. 2003. On D-Trees, Beans, and B-Accents. *Linguistics and Philosophy*, 26(5), 511-545.
- Chafe, W.L. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. *Symposium on Subject and Topic*. University of California. 25-55.
- Constant, N. 2014. Contrastive Topic: Meanings and Realizations. Ph.D. Dissertation. University of Massachusetts Amherst.
- Féry, C. 2007. The Fallacy of Invariant Phonological Correlates of Information Structural Notions. In *ISIS: Working Papers of the SFB 632 (6)*. Ed. by C. Féry, G. Fanselow, M. Krifka.
- Féry, C. & M. Krifka. 2008. Information Structure. Notational Distinctions, Ways of Expression. In *Unity and Diversity of Languages*. Ed. by P. van Sterkenburg. Amsterdam: John Benjamins. 123-136.
- Gundel, J.K. & T. Fretheim. 2004. Topic and Focus. In *Handbook of Pragmatic Theory*. Ed. by L. Horn and G. Ward. Blackwell. 174-196.
- Ladd, D. 2000. *Bruce, Pierrehumbert and the Elements of Intonational Phonology*. Dordrecht: Kluwer.
- Madhavan, P. 1987. Clefts and Pseudo-Clefts in English and Malayalam: A Study in Comparative Syntax. Ph.D. Dissertation. CIEFL: Hyderabad.
- Mathew, R. 2014. The Syntactic Effects of Head Movement: Wh and Verb Movement in Malayalam. Ph.D. Dissertation. University of Tromsø.
- Mohanam, K.P. 1982. Grammatical Relations and Clause Structure in Malayalam. In *The Mental*

- Representations of Grammatical Relations*. Ed. by J. Bresnan. MIT Press: Cambridge, Massachusetts.
- Pierrehumbert, J. 1980. The Phonology and Phonetics of English Intonation. Ph.D. Dissertation. Massachusetts Institute of Technology.
- Reinhart, T. 1981. Pragmatics and Linguistics: An Analysis of Sentence Topics in Pragmatics and Philosophy. *Philosophica anc Studia Philosophica Gandesia Gent*, 27(1), 53-94.
- Rizzi, L. 2013. Topic, Focus, and the Cartography of the Left Periphery. In *The Bloomsbury Companion to Syntax*. Ed. by Luraghi and Parodi. Bloomsbury: London/New York.
- Rooth, M. 1992. A Theory of Focus Interpretation. *Natural Language Semantics* 1, 75-116.
- Stalnaker, R. 1974. Pragmatic Presuppositions. In *Semantics and Philosophy: Essays*. Ed. by M. K. Munitz and P. K. Unger. New York University Press: New York.