# ‘to’-marked XPs in Hindi

*Thesis submitted to Jawaharlal Nehru University*
*in partial fulfilment of the requirements*
*for the award of the degree of*

**DOCTOR OF PHILOSOPHY**

**OJASWEE BHALLA**


**CENTRE FOR LINGUISTICS**
**SCHOOL OF LANGUAGE, LITERATURE & CULTURE STUDIES**
**JAWAHARLAL NEHRU UNIVERSITY**
**NEW DELHI-110067**
**INDIA**
**2021**

Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University
New Delhi-110067 India

Dated 30th December 2021

# CERTIFICATE

This is to certify that the thesis titled **"'to' marked XPs in Hindi"** submitted by **Ms. Ojaswee Bhalla**, in partial fulfillment of the requirements for award of degree of Ph.D. of Centre for Linguistics, School of Language, Literature and Culture Studies-I, Jawaharlal Nehru University, New Delhi, has not been previously submitted in part or in full for any other degree of this university or any other university/institution.

We recommend this thesis be placed before the examiners for evaluation for the award of the degree of Ph.D.

**Prof. Ayesha Kidwai**
SUPERVISOR

Prof. Ayesha Kidwai
Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University
New Delhi-110067

**Prof. Utpal Lahiri**
CO-SUPERVISOR

**Prof. Pradeep Kumar Das**
CHAIRPERSON

Chairperson
Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University
New Delhi-110067

Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University
New Delhi-110067 India

Dated 30th December 2021

**DECLARATION BY THE CANDIDATE**

This thesis titled "**'to'-marked XPs in Hindi**" submitted by me for the award of the degree of Doctor of Philosophy, is an original work and has not been submitted so far in part or in full, for any other degree or diploma of any University or Institute.

Ojaswee Bhalla
PhD Scholar
Centre for Linguistics
SLL&CS
Jawaharlal Nehru University
New Delhi-110067

## ACKNOWLEDGEMENT

I have always liked reading the acknowledgment section of any book or thesis since it opens a window into the author's journey, which has culminated into the piece of work that is in my hands. Someone's journey could bring a smile (or maybe even give hope) to someone else who is figuring out their own path in life. I am grateful for this day that I get to write my own piece of acknowledgement.

My first and foremost thanks goes to Prof. Ayesha Kidwai, who has been the most wonderful and supportive supervisor that one could have asked. Over the last few years, she has been a teacher, a mentor, an emotional support system and a sounding board for sharing ideas about life and linguistics together. From the inception of thinking about particle *–to* to this day, her encouragement for probing the 'bigger questions' has been a constant source of motivation for me. As a human being and as a linguist, I have learned so much from you and I know that I will always be learning from you. Thank you for being there for your students during both the good and the bad days. It matters a lot.

I would like to express my gratitude towards Prof. Utpal Lahiri, my co-supervisor for this thesis. I am grateful to you for opening up the world of semantics for me and for teaching it in such an engaging and thought-provoking manner that the world started making more 'sense' because of it. Thank you for being there throughout this research journey and for taking time out to answer any question that I may have had and for just keeping the door open always.

For someone coming from a background of languages and science, choosing linguistics as the subject of specialisation was quite organic. I was lucky enough to have had a wonderful set of teachers in both JNU (New Delhi) and EFLU (Hyderabad) that introduced the field of linguistics to me. Prof. Anish Koshy and Prof. Amritavalli — thank you for making me curious about language faculty in the first place. Prof. Pradeep Kumar Das, Prof. Pramod Pandey, Prof. Vaishna Narang, Prof. Franson Manjali, Prof. Hari Madhab Ray and Prof. Pandharipande — thank you for providing me with the opportunity to explore all the sub-fields of linguistics during my post-graduation days. A special note of thanks to the wonderful mentors I got to meet

and learn from in LISSIM- IX and X — Prof. Tanmoy Bhattacharya and Prof. Paroma Sanyal. GIAN Workshops provided me with the opportunity to learn from and interact with Prof. Rajesh Bhatt and Prof. Veneeta Dayal — something that I am very thankful for.

Jawaharlal Nehru University — you are the home that I was lucky enough to find and live in for one of the most happiest periods of my life. Thank you for giving me the best place to live and learn about life in. This thesis owes a lot to you, for without your hilly roads, the unique mix of green and red in your essence and the beautiful ecosystem that you create within your walls — I wouldn't be the 'I' am today. Along with this place that gave me the freedom to think and write, I am also grateful to the UGC for their JRF scholarship that financially supported me during this period. Since I am mentioning the place, I owe a thank you to Naveen Bhaiya and Gopal Sir for keeping the system in the Centre for Linguistics office hassle-free for us students.

Now I take this space to thank my friends without whom this thesis, and my life in general, would be incomplete. Monali, you have been there for me since day one and your support through all the phases in the thesis writing process is the most wonderful source of strength that I treasured. Mayank, Goldi and Juhi, thank you for sharing the small joys and for pointing the light at the end of the tunnel during the not-so-good days. Neha, Aditi, Ishita, Archana Di and Deblina — hostel was a welcoming and a warm place because of you people. Chandni and Shreya, thank you for being great companions and wonderful friends through this journey since 2013. Mamta Di, thank you for being a rock solid support and my go-to person for whenever I needed one. You faith in me sustained me through the writing period. Shreya Di, Mathangi, Enakshi, Yangchen, Vyom, Karthick, Nupur, Anjali — this peer group was instrumental in learning how to think and work like a team and I am grateful for the companionship it offered. To Diti, thank you for being forever ready to help a friend out and for helping me access hard-to-find literature in linguistics. On that note, let me also thank the philanthropic organisations that are providing access to academic material to research scholars like us and helping us immensely in the research enquiry. A shout out to Harshita, Harnoor and Reetpal who didn't understand a word of what I was saying about linguistics but were patient enough to hear me out always. Aseema, thank you for proofreading my thesis and for helping me out in my rushed days. Shubhendra Uncle and Saachi, I am extremely grateful for running the final lap along with you both and am thankful for all the timely support you extended. There are more people who have been instrumental in bringing out this thesis. I am grateful to each one of you and I am reserving a big hug for whenever I meet you next.

Finally, I take the moment to thank the people who mean the world to me — my family. To Dama, Mumma, Papa and Didi, thank you for pushing me to give my best in whatever I choose to do (in this case, to do a PhD degree) and to be there to give me hope whenever I am lacking some. Dama, your sheer love gets me going through all the days. Mumma, you do understand me better and know just the right thing to say or do to make things alright. Papa, you are the motivation behind my pursuing a PhD and I am grateful to you for shaping my life the way you do. Didi, you are the voice of reason and the breath of fresh air that helps me out whenever things became a bit much. I am blessed to have you all in my life. This thesis is dedicated to you.

## Table of Contents

## List of Tables

## List of Figures

## List of Abbreviations

| | |
|---|---|
| 1 | first person |
| 2 | second person |
| 3 | third person |
| SG | singular |
| PL | plural |
| M | masculine |
| F | feminine |
| ACC | accusative |
| DAT | dative |
| ERG | ergative |
| GEN | genitive |
| INS | instrumental |
| OBL | oblique |
| HON | honorific |
| LOC | locative |
| SOC | sociative |
| NEG | negation |
| INF | infinitive |
| AUX | auxiliary |
| PROG | progressive |
| PFV | perfective |
| IPFV | imperfective |
| SBJV | subjunctive |
| IMP | imperative |
| PRS | present |
| PST | past |
| FUT | future |
| TOP | topic |
| TO | particle *to* |

# CHAPTER 1

# INTRODUCTION

Minimalist enquiries into Language posit a direct correspondence between a duality of semantic interpretation and a duality of syntactic operations made available by the human language faculty for the derivation of a linguistic expression (Chomsky 2008, 2013, 2019). Thus, a sentence can convey two types of meanings — a core argument-predication related meaning and a peripheral discourse-related meaning ('core' and 'peripheral' for the computational procedure) which are, consequently, a reflex of the two syntactic operations employed to build-up a sentence derivationally.1 Within this framework, discourse roles that instantiate packaging of information in a sentence (like topic or focus) belong to the peripheral discourse-related meaning conveyed by an expression. Hindi, an Indo-Aryan language, has a particle to that has been associated with such discourse-related interpretations like that of topicality (Kidwai 2000) or emphasis (Verma 1971, Lakshmi Bai 1977, Montaut 2015) in the literature. Thus, hypothetically, the semantic-pragmatic interpretation associated with this particle should correspond to a (type of) syntactic operation in the derivational mechanism.

This thesis enquires into the syntax and semantics of Hindi particle *to* vis-à-vis these theoretical assumptions and proceeds to linguistically unpack it at the syntax-semantics-pragmatics interface. It examines the *to*-marked phrases (XPs henceforth) in the language, within a generative approach as its theoretical framework.[2] This chapter introduces the particle to the reader and provides a lay of the land in terms of its empirical distribution in the language. Section 1.1 introduces this particle in terms of its description in the traditional and modern grammars of Hindi and its phrasal

[1] The two syntactic operations are External Merge (EM) and Internal Merge (IM), to be discussed in chapter 5. Argument-predication based semantic interpretation is associated with EM and other types of semantic interpretations with IM.

[2] The category label 'X' of the XP phrase is left unspecified until it becomes relevant for the discussion.

clitic status. Section 1.2 provides an overview of the existing literature on it. Section 1.3 delimits the research problem at hand and specifies the theoretical questions that emerge from it. Section 1.4 covers the empirical distribution of the particle *to* at the phrasal level and also with some specific lexical items that shed light on its occurrence restrictions. Section 1.5 gives an outline of the thesis and specifies the thesis proposal.

### 1.1 What is this particle *to*?

From a purely descriptive perspective, traditional Hindi grammars categorizes the particle *to* under two classes — as an illative conjunction marker and as a marker of emphasis. According to Kellogg, "*to* as an illative conjunction regularly introduces the apodosis of a conditional clause" whereas the second emphatic *to* "conveys a shade of emphasis that can only be found in English by a particular stress of voice" (Kellogg 1938: 490). Illustrative sentences from that work for the two distinct types of *to* particles are provided below.[3] Example (1) exhibits an illative conjunctive *to* particle (glossed as 'then') whereas example (2) shows the occurrence of an emphatic *to* particle after a noun phrase.

(1) dʒo mɛ nəhi dʒaũ to ʊeh nəhi aega
if I not will go then he not will come
'If I do not go then he will not come.'

(2) trib$^{h}$uʊənpəti dʒəgət ka kərta to mɛ hũ
lord of the three worlds earth of creator I am
'Lord of the three worlds and the creator of the earth, I am.'

Agnihotri (2007) describes the particle 'to' as an emphatic particle that has multiple meanings in the language. It can translate to 'so' as in example (3) or it can mean 'so far as x is concerned…'. This type of meaning is conveyed by the sentence in (4), that

[3] Throughout the thesis whenever a linguistic datum is cited from some other author's work, their writing system, interlinear glosses and the translation are not modified and presented as such. This author's own data is written in International Phonetic Alphabet (IPA) and glossed according to Leipzig glossing rules.

can be translated as ‘so far as food is concerned, it is hot’. These examples are from Agnihotri (2007:152).

(3) to kaun aa rahaa hai
so who coming is
‘So, who is coming?’

(4) khaanaa to garam hai
food hot is
‘Food is ready/served hot.’

Such descriptive grammars attempted to specify the usage of the particle *to* and its meaning in terms of English translational equivalents. Within the linguistics literature, this particle has been analyzed as belonging to either the class of discourse particles in functional descriptivist approaches (like Lakshmi Bai 1977 or Montaut 2015 *etc.*) or to the class of topic particles within the generative perspective (Kidwai 2000, 2004). Discourse particles are broadly those items that do not affect the propositional content of the sentence but that convey subjective information like the speaker’s propositional attitude or expectations regarding the discourse participants.[4] Topic particles are the morphological cues that signal the topical constituent in a sentence.[5]

Before delving into an overview of the linguistics literature on the particle *to*, I would like to add an empirical note at this juncture. This thesis makes a distinction between the two types of particle *to* that are externalized in Hindi: an enclitic particle *-to* (that is dependent on and marks a phrasal constituent XP) and a stand-alone, free word *to* (that is not dependent on any specific XP in the clause). A host-adjoining clitic particle has the diagnostic properties of taking phrasal scope over a set of conjoined elements and also that a pause can intervene between a (nominal) host and a particle

[4] The class of discourse particles is discussed in the introduction to chapter 4.

[5] Owing to the vast diversity of literature on the notion of ‘topic’, a full discussion about what marking an entity as topical means is deferred to chapter 2. For now, assume the ‘topic’ of a sentence to be the entity about which that sentence is framed or predicated (the ‘Aboutness Topic’ definition of Reinhart 1981). In example (i), the topical constituent is ‘the boy’ since the rest of the sentence is *about* the referent of ‘the boy’.(i) The boy ate his pizza.

cliticized with it (Bajaj 2016).[6] The particle *to* marks the conjoined nominal elements 'dog' and 'horse' (and not just 'dog') is shown in example (5). A pause, marked by the symbol '/' following the notation for intonational phrase boundaries in Ladd (1981), can intervene between a noun and particle *to* but not between the noun and the oblique affix, as shown in example (6). Based on this evidence, it is concluded that the particle *to* that marks a phrase is a clitic element.

(5) [gʰoɽ-e ɔr kʊtt-e]=ne=to
horse-OBL and dog-OBL=ERG=TO
(the horse and the dog)-to

(6) gʰoɽe-to (without pause)
* gʰoɽ/e-to (with non-licit intervening pause)
gʰoɽe/-to (with licit intervening pause)

As a starting point, I take the enclitic particle *-to* to be the morphological topic marker of Hindi (as proposed by Kidwai 2000, to be discussed in section 1.2.5).[7] Examples (7) and (8) show the enclitic particle *-to* marking a nominal constituent 'I' and a verbal constituent 'eat' respectively.

(7) mɛ=ne=to kʰana kʰa lija tʰa
I=ERG=TOP food eat take.PFV.MSG AUX.PST.3SG
'I have eaten the food.'
'About/as for me, I had eaten the food.' (*intended*)

(8) mɛ=ne kʰana kʰa=to lija tʰa
I=ERG food eat=TOP take.PFV. MSG AUX.PST.3SG
'I have eaten the food.'
'About/as for (the event of) eating, I had eaten the food.' (*intended*)

The non-clitic, free word *to* closely resembles the lexical meaning of English 'then'.[8] Example (9) illustrates the occurrence of the non-clitic *to* particle in a conditional bi-clausal construction in Hindi that consists of a protasis (the initial clause that contains

[6] Bajaj follows Sharma (2003) in analysing Hindi particle *–hii* as a phrasal clitic and not as an affix based on these diagnostic properties of clitics. I analyse the Hindi particle *–to* for these properties to ascertain its clitic status.

[7] Assuming the analysis proposed by Kidwai (2000), the enclitic particle *to* is glossed as a topic marker TOP for now until this analysis is disproven (in chapter 3).

[8] See Montaut (2015) for the differene between lexical items *to* and *tab* (lexical 'then' in Hindi) in terms of the temporal adverbial interpretation they provide.

the condition) and an apodosis (the later clause that contains the conclusion). This particle can stand alone as a question (example (10) spoken with a rising intonation) or it can be the clause-initial element in a question like (11).

(9) əgər tum aje to kʰana bənega
if you come.PFV.MSG then food make.FUT.3SG
'If you come, then food will be cooked.'

(10) to
then
'Then?'

(11) to kja hua
then what happen.PFV.MSG
'Then what happened?'

In examples (10) and (11), the clause initial *to* gives the interpretation of 'then' and it prompts the hearer to provide the event/situation that sequentially or causally follows the previously (implicitly or explicitly) mentioned situation in the discourse. In the conditional case (9), it links the two situations or events as encoded by the two clauses. No variation in the interpretation is observable beyond this for the free word *to*. Crucially, it does not mark any XP in the sentence.

## 1.2 Previous Accounts of Hindi *to*

### *1.2.1 M.K. Verma (1971)*

Working within the transformational grammar framework, Verma describes and compares the structure of the noun phrases in Hindi and English. He categorizes particle *to* be an element of the class of limiter particles. Limiters are the pre-nominal elements that "have a kind of quantitative force, are adverbial in nature" (Verma 1971:39) with the full NP functioning as their domain.[9] Limiter particles are the discontinuous elements of the class of limiters that primarily mark the scope of the

[9] He proposes that items like *sirf* 'only', *bas* 'only', *khaaskar* 'specially', *kam-se-kam* 'minimally' to be limiters in Hindi.

limiter as the full noun phrase.[10] They may additionally "emphasize the limiting implication of the limiter" (Verma 1971:86).

Within the domain of a noun phrase, the particle *to* functions as an emphatic particle that introduces an emphatic implication. In example (12) (Verma 1971:85 ex 47), the intended emphatic implication is paraphrased by the English translation of the Hindi noun phrase. Within his system, limiters can be optionally deleted but the limiter particles can remain (overtly) in the noun phrase. The sentence in example (12) is one such case with no overt limiter, where particle *to* functions as an 'emphasizor' of the noun phrase.

(12) je tin ləʈke to…
'At least these three boys (or, particularly these three boys)…'

The empirical distribution covered by this manuscript is restricted to the noun phrase only and within this description, the particle *to* is described both position-wise (it occurs after the noun) and functionally (it emphasizes and limits the full noun phrase) without delving into what the theoretical implications for the descriptive terms 'emphasis' or 'limiter' could be.

### *1.2.2 B. Lakshmi Bai (1977)*

This paper takes a functionalist-descriptivist approach that builds upon the typology of *to* particles as provided by the traditional Hindi grammars like Kellogg (1938) and Guru (1920) and thus, adopts a homonymy over polysemy analysis of this morpheme. Different lexical *to*s are argued for based on certain co-occurrence restrictions upon them (like conjunctive *to* being able to occur in 'real' questions, whereas emphatic *to*

[10] Other proposed limiter particles are the particles *hii* 'only' and *bhii* 'even/also'. They mark the scope of the limiter *sirf* in example (i) or they can occur as a numeral sub-modifier like example (ii). Data from Verma (1975:87).

(i) sırəf tin ləʈke hi aẽge
'only three boys (and no one else) will come.'
(ii) sırəf tin hi ləʈke aje
'only three boys came.'

cannot). Besides the conjunctive *to* and the emphatic *to* (as mentioned in section 1.1), a third category of request particle *to* is also posited by her to account for an imperative construction like (13). The following (example (44) from that work is produced below (Lakshmi Bai 1977: 75)):

(13) ek glas pani lana to
one glass water bring
'Please bring one glass of water.'

The motivation for positing a third category of this particle is the presumption that this type of *to* can move from a post-verbal position (as in example (13)) to a post direct-object position in (as in example (14)) without any change in interpretation; examples (13) and (14) are from Lakshmi Bai (1977: 75). This type of movement is not permissible for her conjunctive or the emphatic *to* particle.

(14) ek glas pani to lana
one glass water bring
'Please bring one glass of water.'

She assumes that the conjunctive *to* is used to connect the main clause with the conditional clause or to introduce a causal/temporal subsequent clause after a main clause. For the empirical binary distinction made in this thesis, emphatic *to* and request particle *to* can be subsumed under the enclitic *-to* particle whereas the conjunctive *to* is the free word *to*.

Lakshmi Bai's major insight is the observation that the emphatic *to* can attach with only those constituents that are 'given' in the discourse. 'Givenness' is looked at from Chafe's (1976) psychological lens and conveys what is present in the 'consciousness' of the speaker and the hearer.[11] However, the data given in the paper actually exhibits a givenness-spread onto the bigger proposition, which does not remain local on the constituent that particle *to* marks. In the case of example (15) she states that "the speaker assumes not only that the children he is addressing know that 'lion is the king of the forest' but also that this is in their consciousness at the time of his uttering the sentence in question" (Lakshmi Bai 1977: 69). This is problematic for the earlier

[11] Giveness, as a concept of Information Structure, is discussed in chapter 2.

claim that the constituents marked with particle *to* are the ones that are 'given' in the discourse.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (15) | bətʃtʃõ | tumhe | to | pəta | hi | hoga | |
| | children | you | | know | | must be | |
| | ki | ʃer ko | dʒəngəl | ka | radʒa | kehte | hɛ̃ |
| | that | lion | jungle | of | king | say | |

'Children, you may know that lion is called the king of the jungle.'

According to her, "the speaker has the freedom to select any of the candidates from a set of "given" elements as a "focus of contrast". It is precisely this focus of contrast which is marked with *to*" (Lakshmi Bai 1977: 69). Additionally, she posits *to* as an assertion marker and takes this to be the reason for its non-occurrence inside conditional or relative clause, since they are not assertions in her system. All such distributional facts are employed to characterize the discourse function of this particle in Hindi. She concludes that the Hindi *to* particle is associated with a proposition that the speaker makes an assertion about and this particle functions to bring into focus this 'given' proposition.

### *1.2.3 Bhaya Nair (1991)*

This paper focuses on the pragmatic properties of tag questions with a minor focus on the *to* particle in Bangla and its assumed functional counterpart in Malayalam – *eŋkıl*. Nair assumes that Hindi *to* particle has the same properties as Bangla *to* particle, and extends the analysis proposed for Bangla (and Malayalam) to Hindi.

The approach adopted here looks at *to* particle as an 'areal' feature of the Indian sub-continent. This *to* is an emphatic particle, which is used in those discourse contexts where a speaker makes some propositional claim and also express their attitude towards the aforementioned claim. Nair pushes for a relational link between a tag question and the *to* particle. The common factor, between these two, is argued to be that both of them "presuppose some categorical assertion in a conversational context" (Nair 1991: 207). Emphatic particle *to* is speaker-oriented and implicates the conditional certainty that a speaker has regarding the assertion being made on a

belief-continuum. However, no theoretical explanation of the cross-linguistic descriptive facts is offered.

### *1.2.4 A. Montaut (2015)*

Montaut also takes a functional typological perspective to analyse the particle *to*. A crucial difference between Montaut's proposal and Lakshmi Bai's proposal is that while the former pushes for a polysemous (rather than a homonymous) account underlying the diversity of the surface meanings of this particle, the latter had posited three separate morphemes for the three types of *to* — a conjunctive, an emphatic and a request particle. According to Montaut, Hindi has a grammatical word *to* and a discursive particle *to*. The former functions as a co-ordinator or a correlative, while the later functions either as a topic marker (when it has a restricted scope on a constituent) or as an argumentative particle (when it has a wider scope over the proposition). Such varied surface meanings and functions are given a unified derivational analysis by attributing to this particle the role of "triggering an operation which deals with alterity ('otherness'), at various levels: *to* seems to convey the speaker's judgement on a term or sequence on which it has scope, in such a way that *to*-P triggers the implicit or explicit representation of P' (P': non P or other than P)" (Montaut 2015: 12). This abstract operation is labelled as 'inter-subjectivity'.

In the psycho-pragmatic analysis adopted in this work, the importance of the interlocutor's viewpoint and the dynamic interplay of conversational discourse is highlighted. When *to* functions as a co-ordinator, it introduces a new frame for predication, which is distinct from previous discourse sequence. As a correlative (in either the conditional *agar...to...*'if-then' construction or a temporally dependent *jab..to...* 'when-then' construction), it functions as an anaphoric device with the apodosis and it "selects a given path as a choice made amongst other available paths" (Montaut 2015: 13). As a topic particle, it thematizes the term (let us assume P) it attaches to (hence, it has restricted scope here) or it changes the relation between P and the rheme or it presents a viewpoint on P that is different from a previously expressed viewpoint. The non-thematic *to* is assumed to have a wider scope — over a whole sentence. In those sentences that have no topical constituent, this particle is

used to counter an initial proposal by presenting a counter-argument or it weakens the relevance of a proposal already made. Thus, "the operation triggered by the discourse particle *to* always involves inter-subjectivity, since it negotiates with the other's viewpoint, distinct from and often conflicting with the utterer's viewpoint" (Montaut 2015: 13). Out of the two — grammatical word *to* and discourse particle *to* — it is the latter which is sensitive to the operation of inter-subjectivity.

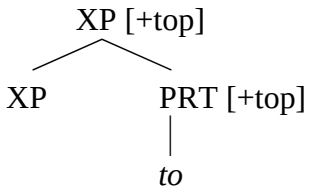
### *1.2.5 Kidwai (2000, 2004)*

One of the contributions of Kidwai (2000) is that of locating the question of topicalization as a syntactic operation in the larger issues of word order, scrambling and XP-adjunction in the universal grammar within the generative enterprise. She positions herself within the Minimalist Program framework (Chomsky 1995, 2000) and proposes an analysis of particle *to* as the morphological topic marker for Hindi. She argues for a syntactic distinction between topicalization (as a substitution operation) and scrambling (an adjunction operation), although both have a surface similarity of moving the constituent to a clause-initial position. She argues for topicalization to be a covert operation that moves the topical constituent into some position within the Complementizer Phrase (CP) domain. This position is the specifier position of a functional head 'Top' (for topic) in the left periphery of the clausal spine (as suggested by MS 1993). In her system, "the presence of topic morphology designates the Top head" (Kidwai 2000: 49) and Hindi *to* topics involve a covert topicalization operation. Thus, for her, both the topicalized topics and the *to*-marked topics occupy the same position — that of [Spec, TopP] — by the end of the derivation. This analysis works well for unmarked clause-initial sentence topics but not for the *to*-marked XPs, which may not obligatorily receive a topical interpretation and which fail to give any evidence for an obligatory syntactic movement from their base position, a matter which I discuss in detail in chapter 3 and chapter 5 respectively.

To account for the syntax of *to*, she adopts the notion of 'minor functional head' (proposed in Bayer 1996 and Rothstein 1991). The typology of functional heads operative in the syntax was expanded to include the category of minor functional

heads that subcategorize for their argument types but that do not have a theta grid or bind theta positions or project category features. Such a head modifies its complement and projects further the category of its complement, with its unique feature being percolated to the top maximal projection. In her analysis, the particle *to* is specified for a feature [+top].[12] The projected structure of a *to*-marked XP is given in (16).

(16)




To analyze the semantic import of these *to*-marked XPs, Kidwai adopts the proposal given by Miyagawa (1987) for Japanese topic particle *wa*. Entities marked with morphological topic particles can receive either a thematic reading or a contrastive reading. The *to*-marked XP in example (17) receives a thematic reading and the one in example (18) receives a contrastive reading; data from Kidwai (2000: 43 ex 39-40).

(17) ram-to kəl aega
Ram-TOP tomorrow will come
'Ram will come tomorrow.'

(18) ram-to aega, ɔr koi ae-na-ae
Ram-TOP will come else any come-not-come
'Ram will come, whether anybody else comes or not.'

These readings are a result of the manner in which a particular semantic property of the topic particle is satisfied in a sentence. This property is what Miyagawa calls the semantic property of 'set anaphoricity' of a topic particle. Set anaphoricity stipulates that the topic particle makes reference to an identifiable set of individuals/entities. This set of individuals or entities must be present in the immediate conversational context. The knowledge of this set is shared by the interlocutors of the discourse. The crucial condition is that every member of this set must be exhaustively represented in the statement containing the topic particle. The anaphoric relation is with the set as a whole and not some individual entity from the set. The thematic and contrastive

[12] [+top] is assumed to be a morpho-syntactic feature in this system.

reading difference comes about by how the set, that is set-anaphorically referred to by that topic particle, gets exhaustively represented in the conversation. A thematic reading is received when all members of the set are referentially picked out and associated with the same property by that particle. A contrastive interpretation comes about when the particle attaches to a portion of the set, which gains 'IN CONTRAST TO' relation to other members of the set — thereby exhaustively representing the full set in the sentence.

As an example of how this analysis proceeds, consider (19)-(24) in which the topic particle *-wa/-to* marks the external argument noun phrase. Japanese data is from Miyagawa (1987: 186) and Hindi data is from Kidwai (2000: 45).[13]

(19) dʒon wa hon o jonda
John TP book DO read
'As for John, he read a book.'

(20) kudʒira wa hoŋgju-dobutsu desu
whales TP mammals COP
'Whales are mammals.'

(21) ame wa futeimasu ga, juki wa futeimasen
rain TP falling but snow not-falling
'It's raining, but it isn't snowing.'

(22) dʒɔn-to kɪtab pəɽʰta hɛ
John-TOP book reads is
'As for John, he reads books.' (thematic)

(23) kutte-to wəfadar hote hɛ̃
dogs-TOP faithful be-HAB are
'Dogs are faithful.' (thematic)

(24) barɪʃ-to ho rəhɪ hɛ, pər ole nəhɪ pəɽ rəhe hɛ̃
rain-TOP be IPFV is, but hailstones not fall IPFV are
'It's raining, but there's no hail.' (contrastive)

Anaphoric nominals (like 'John' in examples (19) and (22)) or generic nominals (like 'whales' in example (20) or 'dogs' in example (23)), when marked by a topic particle,

[13] The full forms for glossing abbreviations used by Miyagawa for Japanese data are: TP - topic, DO - direct object, COP – copula.

give a thematic reading. The referent set for an anaphoric nominal is the singleton set containing that anaphor itself. The referent set for a generic nominal includes all the entities that are included in the predicate class. When a topic particle marks such nominals, the full referent set gets exhaustively represented in the sentence, thus invoking a thematic reading. In examples (21) and (24), the contextually available referent set includes the elements: {rain, hailstones}. For receiving the contrastive interpretation in (24), the particle *to* marks one element of the set ('rain') and creates a relation of 'in contrast to' with the remaining member of the set ('hailstones'). Thus, by virtue of this relation, all the members of the set are exhaustively represented in the set. Thus, in this analysis, the thematic and the contrastive semantic interpretations of a *to*-marked XP is contingent on the mechanism of fulfilment of a property of *to* as the topic particle itself.

In the early versions of the minimalist framework (Chomsky 1995, 2000), the architecture of the universal grammar involved a computational mechanism (the narrow syntax) that selected items from a mental lexicon and assembled them into an expression 'Exp'. The derived expressions are assigned a sound-meaning pair by the external performance systems —the Conceptual-Intensional (C-I) system that assigns the semantic interpretation and the Articulatory-Perceptual (A-P) system that is responsible for externalizing the expression in any language modality (oral or manual). The PF (Phonological Form) and the LF (Logical Form) are the interfaces at which the derived expressions is input to the external systems. With this as the premise, Kidwai (2004) questions whether topicalization can be subsumed under a uniform scenario of bare output conditions (Chomsky 2001) imposed on an expression by the PF and the LF interface. She argues that topicalization is a PF derivational output that interfaces with both the external performance systems (the C-I and the A-P). Crucially, she reformulated the EPP feature as the checking requirement for each sentence to have a topic and not the subject (as priorly formulated).[14] Within

[14] EPP stood for the Extended Projection Principle in the Government and Binding era (Chomsky 1981) of the generative tradition This principle was interpreted as a near-universal syntactic requirement for all finite clauses to have a subject by the end of the derivation. Later versions (Chomsky 1995 onwards) modified EPP to be a syntactic feature of a functional head that motivates the syntactic movement of a constituent to the specifier position of the head with this feature.

this proposal, for languages like Hindi, the EPP requires the topical constituent to occupy the [Spec, TP] position.[15] Her analysis works for the unmarked topical constituents but not for the *to*-marked topics since the particle *to* can attach to a variety of lexical categories and they do not necessarily move to [Spec, TP].

This section gave an overview of the various approaches that have been taken to look at particle *to* in Hindi so far. The next section specifies the research questions that this thesis is engaging with before starting an empirical investigation into the linguistic phenomenon of marking of constituents by this particle.

## 1.3 Research Questions

The goal of this thesis is to attempt to unpack the semantic-pragmatic interpretations of sentences with *to*-marked constituents, the meaning conveyed by this particle as well as the syntactic derivation of the *to*-marked XPs in the language. The question is interesting because there has been no explanatorily adequate, formal engagement with it that can account for the distributional facts and the varied range of interpretations available for this particle. At this point, two co-varying hypotheses can be formulated:

I. Is there a single lexical item *to* that enters into the syntactic computation but yields varied semantic interpretations based on the different syntactic positions it ends up at?

II. Or are there multiple lexemes of *to* in the Hindi lexicon (polysemous but homonymous, contra (I)) that each have their own fixed syntactic position but have a simpler derivation of their meaning?

An enquiry into this dichotomy thus involves a theoretical choice — between 'a simple syntax but a complex semantics' vs. 'a complex syntax with simpler semantics' analysis of particle -*to*. This thesis claims that the derivational machinery underlying enclitic particle –*to* requires a simple syntax with a complex semantic interpretative system.

At the start of the investigation, I adopt Kidwai (2000) as the null hypothesis and assume the enclitic particle *-to* to be the morphological topic marker for Hindi. This analysis presumes a one-to-one correspondence between a linguistic entity and its information structural function. This thesis questions the basic premise of such a direct mapping of a discursive notion (like the concept of 'topic') onto a morphological form (like particle *to*). The sub-questions that entail this theoretical enquiry are:

a. For the topical interpretation of *to*-marked XPs, what conception of 'topic' applies to them?

b. Do such XPs satisfy the standard diagnostic tests for topichood available in the literature?

c. If there are other (non-topical) interpretations possible, then can the assumption of the particle *to* being a grammatical reflex of a topic interpretation be maintained?

d. From a distributional point of view, what are the pragmatic conditions that constrain the marking of a constituent by particle *to*?

e. What is the semantic contribution of this particle in a sentence? How can the various semantic-pragmatic interpretations of a *to*-marked XP be derived in a compositional semantic model?

f. What is the syntax of the *to*-marked XPs, given our current understanding of the minimalist enterprise? Can it involve a feature driven movement (Chomsky 2001) or a specific type of compositional operation (Chomsky 2008) that facilitates discursive interpretations? Alternatively, can it not be accounted for within the minimalist tradition and thus, requires looking for some novel theoretical framework?

g. What is the array of features associated with this particle in the lexicon? Is there any evidence of a feature-based interaction in its computation in the narrow syntax?

h. What can the difference between *to*-marked topics, *to*-marked non-topics (if any) and the unmarked topics tell us about the architecture of grammar?

i. What evidence does Hindi provide for Chomsky's claim that "there is no direct mapping between Universal Grammar and Information Structure" (Chomsky 2008)?

Questions (a)-(b) are the subject matter of chapter 2; questions (c)-(d) are discussed in chapter 3; chapter 4 is concerned with question (e); chapter 5 attempts to propose a solution for questions (f)-(g); chapter 6 conjectures about the thesis proposal and their relation to questions (h)-(i).

To begin to answer these questions in this thesis, I take the first steps in the next section, by unpacking the distribution of enclitic particle *–to*.

## 1.4 Empirical Distribution of Particle *-to*

### *1.4.1 Distribution at the Phrasal Level*

Kidwai proposes that enclitic particle *to,* as a marker of topical entities, "can attach to any maximal projection of a lexical category (DPs, PPs, VPs) but generally cannot be 'inserted' inside that maximal projection" (Kidwai 2000: 41). Example (25) shows a *to*-marked Determiner Phrase (DP) 'my black book'; a Verbal Phrase (VP) 'read book' is marked by this particle in (26); a postpositional phrase (PP) 'to Noor' is marked by particle *–to* in example (27). This data is cited from Kidwai (2000: 42 ex 35-37).

(25) [$_{DP}$ meri kali kɪtab]-to mɪl gəji
my black book-TOP found went
'My black book was found.'

(26) ram [$_{VP}$ kɪtab pəʈ$^{h}$]-to rəha hɛ
Ram book read-TOP PROG is
'Ram is reading a book.'

(27) sita [$_{PP}$ nur-ke pas]-to gəji
Sita Noor-GEN near-TOP went
'Sita went to Noor.'

However, a closer look at the data from Hindi shows that particle *to* can mark lexical entities within these maximal projections too, besides marking other phrasal categories like Adv(erbial) Phrase and CP. Keeping the test construction same for DP, VP and PP cases, I show the various constituents that can be marked by particle *–to* in them. To tease out the *to*-induced interpretations, a potential discourse continuation is provided to pragmatically contextualize a sentence where needed.

#### *1.4.1.1 'to'-marking of a PP*

Hindi nominal morphology has been analysed in terms of three classes or 'layers' of functional category markings on the nominal (see Masica 1991, Butt and King 2004, Spencer 2005). Layer I morphology consists of the number or the oblique inflection of the noun stem depending on its declension class. Layer II morphology consists of the postpositional clitics that mark the syntactic and semantic functions of case like enclitic *–ne* (ergative), *-se* (ablative), *-ka* (genitive) etc. Layer III morphology instantiates spatial expressions in Hindi in the form of complex postpositions that involve an obligatory layer II postposition *–ka* (genitive) and a lexical term indicating spatial orientation of an entity with respect to the complement of the layer II postposition.

Example (27), repeated below as (28) for expository purpose, consists of a complex postposition in Hindi. The noun 'Noor' is a proper noun that does not inflect for layer I morphology. The layer III spatial postposition *pas* 'near' is mediated via an obligatory *–ka* (layer II postposition) marking of the noun. Kidwai claims that in (28), the particle *–to* can mark the maximal projection layer of the PP—the spatial postposition *pas* 'near'.

| | | | | |
|---|---|---|---|---|
| (28) | sita | nur=ke | pas=to | gəji |
| | Sita | Noor=GEN | near=TOP | go.PFV.3FSG |
| | 'Sita went to Noor.' | | | |

However, (29) exhibits that the particle *–to* can also mark the noun and the genitive postpositional unit *nurse=ke* within a PP projection. The overt contrastive

juxtaposition of the two PPs ‘to the nurse’ and ‘to the doctor’ facilitate the *to*-marking of the genitive marked noun in this example.

(29) sita nərs=ke=to pas gəjı pər ɖɔkʈər=ke nəhı
Sita nurse=GEN=TOP near go.PFV.3FSG but doctor=GEN not
‘Sita went to the nurse but not to the doctor.’

*1.4.1.2 ‘to’-marking of an AdvP*

Cinque (1999) proposed a fixed relative ordering of different classes of adverbs that underlie a universal hierarchy of adverbial projections in the clausal structure. The proposed fine-grained typology of adverbs can be collapsed into two broad categories based on their relative position in the syntactic spine: ‘higher’ adverbs (like speech act adverbs, evaluative adverbs, evidential adverbs, epistemic adverbs etc.) and ‘lower’ adverbs (like manner adverbs, frequentative adverbs, completive adverbs etc.). Lower adverbs are associated with the modification or the specification of the event structure and they attach low in the verbal domain. Higher adverbs are discourse-oriented or speaker-oriented and they attach in the higher CP or TP domain.

Example (30)-(35) show differentially felicitous marking of the higher and lower adverbs with particle *-to* in Hindi. A manner adverb ‘fast’, a completive adverb ‘completely’ and a frequentative adverb ‘repeatedly’ are *to*-marked in the examples (30), (31) and (32) respectively. All these are lower adverbs that attach lower in the spine as they modify the verb-denoting event. A possible discourse continuation for the sentence in each case is provided in the parenthesis in the translation, without which the sentence is interpreted to be informationally incomplete.

(30) gʰoʈa tez=to dɔʈa pər
horse fast=TOP run.PFV.3MS but
‘The horse ran fast but’ *(couldn’t win the race)*

(31) ʊs=ne ʃɛhər [puri tərɛh=se]=to bərbad kija tʰa pər
he=ERG city completely=TOP destroy do.PFV.MSG AUX.PST.3SG but
‘He had completely destroyed the city but’ *(the city rose from the ashes)*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (32) | tʊm | ʊdʰər | [baːr baːr]=to | dʒate | rehte | ho | pər |
| | you | there | repeatedly=TOP | go.IPFV.MSG | stay.IPFV.MSG | AUX.PRS.2SG | but |

'You go there repeatedly but' *(you still don't know the layout of that place)*

In contrast to the lower adverbs, particle *-to* cannot mark the higher adverbs in Hindi. Example (33) involves and episteme adverb 'probably' that conveys the speaker's epistemic state regarding the proposition. Example (34) has a speech act adverb 'honestly' and example (35) has an evaluative adverb 'unfortunately' — that conveys the speaker's evaluation of the state of affairs encoded by the sentence. All of these high adverbs cannot be marked by particle *–to* as indicated by the ungrammaticality (33)-(35).[16]

| | | | | | |
|---|---|---|---|---|---|
| (33) | * ʃajəd=to | gʰoɽa | tez | ɖɔɽa | |
| | probably=TOP | horse | fast | run.PFV.MSG | |

'Probably the horse ran fast.'

| | | | | | |
|---|---|---|---|---|---|
| (34) | * [səʈʃ mẽ]=to | mɛ=ne | us=ko | dekʰa | hɛ |
| | honestly=TOP | I=ERG | he.OBL=ACC | see.PFV.MSG | AUX.PRS.3SG |

'Honestly, I saw him.'

| | | | | | |
|---|---|---|---|---|---|
| (35) | * dʊrbʰagjavənʃ=to | us=kɪ | mɔt | ho | gəji |
| | unfortunately=TOP | he.OBL=GEN | death | happen | go.PFV.FSG |

'Unfortunately, he died.'

Thus, particle *-to* can mark adverbs in Hindi but there is restriction on its distribution *vis-à-vis* the type of adverb being marked i.e. lower adverbs can be *to*-marked but higher adverbs cannot.

### *1.4.1.3 'to'-marking of a DP*

In the domain of a DP, Kidwai (2000) had proposed that the particle *–to* can mark the maximal projection of a DP, as shown in example (25). However, as (36) and (37) exhibit, the particle *–to* can also mark the possessive determiner *meri* 'my' or the adjective *kali* 'black', respectively, within the DP projection. In both these examples, the *to*-marked lexical entity is juxtaposed with an alternative of the same lexical

[16] Following standard conventions, ungrammaticality is indicated by a star '*' prefix and pragmatic infelicity is indicated by a hash '#' prefix before a sentence.

category in the second clause. Thus, a *to*-marked determiner *meri* 'my' is contrasted with the possessive *hari-ki* 'of Hari' in (36) and a *to*-marked adjective *kali* 'black' is contrasted with the adjective *lal* 'red' in (37). Without these second contrastive clauses, the first clause is interpreted to be informationally incomplete.

(36) meri=to kali kıtab mıl gəji pər
my=TOP black book find go.PFV.FSG but
həri=kı nəhi mıli
Hari=GEN NEG find.PFV.FSG
'My black book was found but not Hari's (black book).'

(37) meri kali=to kıtab mıl gəji pər
my black=TOP book find go.PFV.FSG but
lal nəhi mıli
red NEG find.PFV.FSG
'My black book was found but not the red (book).'

It is noted that in Hindi, the enclitic particle *–to* can mark only one of the lexical entities in a DP. This is shown in (38) that has both the adjective and the noun *to*-marked and is consequently, judged ungrammatical.

(38) * meri kali=to kıtab=to mıl gəji
my black=TOP book=TOP find go.PFV.FSG
'My black book was found.'

As discussed in the preceding section on the *to*-marking of PPs, the case markers in Hindi are analysed to be the postpositional clitics (layer II) on the noun (Butt and King 2004, Spencer 2005). Within a DP, the enclitic particle *–to* can mark the case marked noun but it cannot intervene between a noun and its case marker clitic. This generalization is given in (39), where CM stands for case marker. The illicit ordering of a case clitic and the enclitic *-to* on a noun is exhibited in (40a)-(40c) for ergative case marker *–ne*, dative case marker *–ko* and the ablative case marker *–se* respectively.

(39) *[NP=TOP=CM]

(40) a. * əlı=to=ne
Ali=TOP= ERG
b. * zoja=to=ko
Zoya=TOP=DAT

c. * peɽ=to=se
tree=TOP=ABL

For a genitive DP, the particle *-to* can mark either the possessed nominal or the case marked possessor nominal but following the generalization (39), it cannot intervene between the possessor nominal and the genitive case clitic *–ka*. Example (41) shows the permissible and the non-permissible *to*-marking of a genitive case marked DP in a Hindi sentence.[17]

(41) ram=(*to)=kı=(to) saikəl=(to) pəntʃər hɛ
Ram=(*TOP)=GEN=(TOP) cycle=(TOP) puncture AUX.PRS.3SG
'Ram's cycle is punctured.'

*1.4.1.4 'to'-marking of a VP*

Hindi has a base order of the sentential constituents as given in (42).

(42) Base Order: Subject —Indirect Object (IO)—Direct Object (DO)—Verb

Within a VP domain, Kidwai (2000) had proposed that the enclitic particle *–to* can mark the maximal lexical projection (the verb), as shown in example (26). Examples (43) and (44) exhibit that this particle can (also) mark either the direct object ('the book' in (43)) or the indirect object ('Sita' in (44)) in a VP.

(43) ram=ne sita=ko kıtab=to di
Ram=ERG Sita=DAT book=TOP give.PFV.FSG
'Ram gave a book to Sita.'

(44) ram=ne sita=ko=to kıtab di
Ram=ERG Sita=DAT=TOP book give.PFV.FSG
'Ram gave a book to Sita.'

A comparison of the data in (45)-(49) puts forth the observation that the enclitic particle *–to* cannot mark the verbal element at the clause-final position. Hindi has simplex predicates that consist of one lexical verb that is marked for the aspectual information, as in (45). The lexical verb in (45) can be optionally followed by an

[17] Note that the enclitic particle *–to* can occur only once in a DP, so it cannot mark both the nominals in a genitive DP in a sentence.

auxiliary verb in, as in (46) and (47). Hindi also has complex predicate constructions that are composed of two verbal elements (a 'main' lexical verb and a semantically bleached 'light' verb) with the aspectual information encoded by the light verb (see Mohanan 1994; Butt 2003, 2010 etc.). Examples (48) and (49) involve a complex predicate construction. The common factor underlying the ungrammaticality of (45), (47) and (49) is that in all these cases the enclitic particle *–to* marks the clause final verbal element (whether it be a lexical verb as in (45) or an auxiliary verb as in (47) and (49)).[18]

(45) * ram=ne sita=ko kitab di=to
Ram=ERG Sita=DAT book give.PFV.FSG=TOP
'Ram gave a book to Sita.'

(46) ram=ne sita=ko kitab di=to tʰɪ
Ram=ERG Sita=DAT book give.PFV.FSG=TOP AUX.PST.3SG
'Ram gave a book to Sita.'

(47) * ram=ne sita=ko kitab di tʰɪ=to
Ram=ERG Sita=DAT book give.PFV.FSG AUX.PST.3SG=TOP
'Ram gave a book to Sita.'

(48) ram kɪtab dekʰ rəha=to hɛ
Ram book look IPFV.MSG=TOP AUX.PRS.3SG
'Ram is looking at the book.'

(49) * ram kɪtab dekʰ rəha hɛ=to
Ram book look IPFV.MSG AUX.PRS.3SG=TOP
'Ram is looking at the book.'

To sum up, this section has looked at the empirical distribution of the enclitic particle *–to* in Hindi at the phrasal level. It can mark elements at the maximal projection level as well as within the maximal projection i.e. an XP. This section proposes that this particle can mark entities of any 'lexical' category — nouns, adjectives, adverbs, verbs, and postpositions. The three restrictions observed on its distribution in the sentence are:

(i) It can mark only one constituent in a sentence.

[18] Chapter 5 discusses the restriction of the particle *–to* from marking the constituent at the clause final position from a theoretical perspective.

(ii) It cannot mark higher adverbs that are discourse-oriented or speaker-oriented.

(iii) It cannot mark clause-final constituents.

### *1.4.2 Distribution with some select lexical items*

#### *1.4.2.1 Distribution with '–hii' and '–bhii'*

Hindi has particles *–hii* and *–bhii* that have been analysed as phrasal clitic elements (Sharma 2003, Bajaj 2016) that can convey multiple semantic interpretations in a sentence (see Bajaj 2016 for a full discussion about *–hii*; Dayal 2019 for a recent discussion about *–bhii*). The seminal work by Rooth (1996) and the vast literature branching out from it proposes a semantic analysis that posits 'only', 'even' and 'also' to be proposition-level operators that are sensitive to the presence of relevant alternative propositions in the discourse. The proposition that is marked by these operators is called 'the prejacent'. 'Only' is interpreted as an 'exclusivity' operator i.e. the prejacent is evaluated as true as long as no alternative proposition in the context is also true. 'Also' is interpreted as an 'additivity' operator i.e., it is presupposed that some alternative proposition is true and it is asserted that the prejacent is also true. 'Even' is analysed in terms of 'scalarity' i.e. based on speaker's expectation, a likelihood scale is established and the prejacent marks the endpoint of that scale. Hindi particle *–hii* is, by now, standardly assumed to give an 'only'-type exclusivity reading and particle *–bhii* to convey both an 'also'-type additive reading and an 'even'-type scalar reading.[19] Dayal (2019) proposes that these particles are presuppositional in nature i.e., a sentence marked by these particles denotes the prejacent and adds a presupposition about their scalar, exclusive or additive meaning part.[20]

[19] Bajaj (2016) has proposed a scalar and a non-scalar reading of particle *–hii* besides an exclusivity reading. Dayal (2019) has proposed three types of readings for particle *–bhii*: 'also', 'even' and an' ever-identity' reading. I refer the reader to this literature for an exposition of these proposals.

[20] The inference relation of presupposition is discussed in chapter 3. For now, assume that presuppositions are the inferential propositions that are assumed to be true as background information for evaluating a sentence.

Example (50) shows the exclusivity reading for a sentence containing a *–hii*-marked noun. Examples (51)-(52) exhibit the additive meaning component and the scalar meaning component, respectively, added by particle *–bhii* in a sentence.

(50) ram=hɪ aja
Ram=only come.PFV.MSG
Assertion: Ram came.
*Exclusivity Presupposition*: Only Ram (and no one else) came.

(51) ram=bʰɪ aja
ram=also come.PFV.MSG
Assertion: Ram came.
*Additive Presupposition*: Someone came. Ram also came.

(52) ram=bʰɪ aja
ram=even come.PFV.MSG
Assertion: Ram came.
*Scalar Presupposition*: Ram was the least likely person to come (on a likelihood scale ranking the alternatives in the context)

Hindi enclitic particle *–to* can co-occur with these *–hi* and *–bhii* particles in the nominal domain. However, it is observed that there is an ordering restriction on the distribution of these particles. The particle *–to* cannot intervene between the noun and the particles *–hii/-bhii*, a generalization outlined in (53). If a noun is to be marked by both — the particles that function as semantic operators like *–hii/-bhii* and the enclitic particle *–to*, then these particles mark the noun first.

(53) *[NP=to=hi/bhii]

Examples (54)-(55) highlights the fact that the ordering of these particles has an effect on the grammaticality of a sentence. A tentative proposal for this particle stacking issue is discussed in chapter 5.

(54) ram=hi=to aja
* ram=to=hi aja
'Only Ram came.'

(55) ram= bʰɪ=to aja
* ram=to=bʰɪ aja
'Even/Also Ram came.'

*1.4.2.2 Distribution with 'koi'/'kuch' and 'sab'/'har'*

In logic, quantification has been described as a relation between two predicates (by Frege, the French philosopher). Intuitively, quantifiers in natural language restrict the 'quantity' of things they quantify over such that a property (the predicate) is ascribed to them (Lahiri 2017). Universal quantification (that ascribes a property to all the members of a set of entities/individuals) and existential quantification (that ascribes a property to minimally one member from the set of entities/individuals) are concepts of the first order predicate logic language. English determiners 'every' and 'some' are canonically treated to be the natural language quantifier with a universal force and the natural language quantifier with an existential force, respectively. Mahajan (2017) provides an exhaustive overview of the various types of quantified expressions in Hindi. In this section, I focus on the distribution of the Hindi enclitic particle *–to* with the existential quantifiers *koi/ kuch* 'some' and the universal quantifiers *sab* 'all' and *har* 'each'.

Depending on the animacy feature of the noun, Hindi forms an existentially quantified nominal expression using either *koi* ('some' that is lexically specified as [+animate]) or *kuch* ('some' that is lexically specified as [-animate]). Enclitic particle *–to* can mark the quantified NP (example (56)) or the quantifier itself (example (57)) in an existentially quantified nominal constituent. Marking of the quantifier by the particle *–to* in (57) adds an 'at least' type semantic interpretation. This sentence can be interpreted to convey that the work done by the subject is at the minimal endpoint of a pragmatic scale that ranks the quantity of work that could be done by the subject.[21]

(56) ʊs=ne [kʊtʃʰ kam]=to kija
he=ERG some work=TOP do.PFV.MSG
'He did some work.'

(57) ʊs=ne kʊtʃʰ=to kam kija
he=ERG some=TOP work do.PFV.MSG
'He did (at least) some work.'

[21] An analysis for such 'at least' type interpretation for a 'to'-marked entity are discussed in chapter 3.

The universal quantifier *sab* ‘all’ has a cumulative meaning and the universal quantifier *har* ‘each’ has a distributive meaning. Examples (58) and (59) exhibit that a *sab*-quantified NP can be marked by particle *–to* (in (58)) but marking only the quantifier *sab* with particle *–to* degrades the pragmatic felicity of the sentence (in (59)).

(58) səb bətʃtʃe =to gʰər gəje
all children=TOP home go.PFV.MPL
‘All children went home.’

(59) ?# səb=to bətʃtʃe gʰər gəje
all=TOP children home go.PFV.MPL
‘All children went home.’

A further degradation into ungrammaticality is observed in the case of example (60) where the quantifier *har* ‘each’ is marked by the particle *–to* in the universally quantified nominal expression.

(60) * hər=to bətʃtʃa gʰər gəja
each=TOP child home go.PFV.MSG
‘Each child went home.’

The difference in the degree of acceptability of examples (59)-(60) indicate that the pragmatic-semantic import of the particle *–to* is not compatible with a universal quantifier that has a distributive reading but can be marginally associated with an quantifier with a cumulative reading. This raises questions for the semantics of particle *–to* and is discussed in chapter 4.

#### *1.4.2.3 Distribution with ‘nahi’ and ‘mat’*

Assuming standard semantics terminology, negation is interpreted to be a function that takes as argument a proposition and returns the proposition with its truth value reversed. Hindi has a clausal negation operator lexicalized as *nahi* ‘not’ and a prohibitive negation marker *mat*.

Examples (61)-(62) exhibit that the enclitic particle *–to* cannot mark elements that encode semantic negation in Hindi. These negative elements do not have a lexical

meaning that they denote. Thus, as a tentative generalization, it can be assumed that the particle *–to* can mark only those lexical items that have a lexical meaning or semantic content.

(61) * ram kɪtab pəɽʰ nəhɪ=to rəha tʰa
ram book read NEG=TOP IPFV.MSG AUX.3SG
'Ram was not reading the book.'

(62) * tʊm kɪtab mət=to pəɽʰo
you book NEG=TOP read.IMP
'You do not read that book.'

To sum up this sub-section, the distribution of the enclitic particle *–to* with the following lexical items— semantic operator type elements: *–hii* 'only' and *–bhii* 'also/even'; nominal quantificational elements: *–sab* 'all', *–har* 'each', *–koi/–kuch* 'some' and the negation markers *nahi* 'not' and *mat* was observed. This particle can co-occur with *–hii* and *–bhii* but it cannot intervene between the noun and these semantic operators. The particle *–to* can mark universally and existentially quantified NPs. It can also directly mark the existential quantifier *koi/kuch* (that evokes an 'at least' type interpretation) and to a partial degree of acceptability mark the cumulative universal quantifier *sab* 'all'. It cannot mark the distributive universal quantifier *har* 'each'. Neither can it mark the negation markers *nahi* or *mat* that convey a negation operation but do not encode some lexical semantic content. Thus, the distribution of this particle is constrained by the idiosyncratic semantic restrictions of this particle.

This section has attempted to cover the empirical distribution of the enclitic particle *–to* within the phrasal domain and with some specific operator-type lexical items. It was observed that this particle can mark nouns, adjectives, determiners, verbs, adverbs and postpositions. All of these are lexical categories that encode some semantic content. This particle cannot mark negation particles that do not encode a lexical meaning *per se*. This particle cannot mark clause final elements whether they be a lexical verb or an auxiliary verb. This particle cannot mark discourse-oriented or speaker-oriented higher adverbs. There is an ordering restriction when this particle and some other phrasal clitic element with a semantic (case markers or additivity, exclusivity and scalarity operators) or a syntactic function (case markers) have to

mark a noun. These semantic or syntactic phrasal clitics mark the noun first and then enclitic particle *–to* marks this unit. A constituent that is marked by this particle may receive an 'at least' type interpretation or may require contrastive discourse relation with some relevant alternative in the discourse context. A distributive quantifier is not compatible to be marked by particle *–to* but a cumulative quantifier is partially compatible with the interpretative import of this particle. All of these distributional facts are assumed to be indicative of the linguistic packing of the particle *–to*.

An analysis of this particle as a topic particle cannot account for the distributional facts observed above. If the analysis that the particle *–to* marks topical constituents, then it cannot predict the marking of adjectives, determiners, quantifiers or the postpositions. Neither can it propose a non theory-internal explanation for the ordering sequence *vis-à-vis* the case markers or the semantic operators. It cannot explain the derivation of an 'at least' type interpretation (by *to*-marking of *koi/kuch*) or the contrastive interpretation (by *to*-marking of determiners, adjectives etc.) yielded by the marking of constituents that are not sentence topics *per se*.

This thesis proposes an analysis of the particle *–to* in terms of the notion of discourse salience.[22] Salience, assumed to be the degree of the relative prominence of a unit of information at a given point in time (Chiarcos *et al* 2011), can be imposed on an entity by a morphological cue, a prosodic cue or a syntactic cue (Heusinger and Schumacher 2019). This thesis proposes that the enclitic particle *–to* is a morphological prominence lending cue in Hindi. The distributional facts observed in this section can be explained by analyzing the meaning of discourse salience conveyed by the particle *-to* and its interaction with the syntactic configuration and the pragmatic considerations.

## 1.5 The Roadmap Ahead

This thesis provides evidence to claim that the discursive interpretations like topicality or contrast are not directly manifested via a syntactic operation or a

[22] Chapter 4 discusses the concept of discourse salience in section 4.2.3.

morphological form in the language but they are a product of the interaction of certain syntactic, semantic and pragmatic factors. The enclitic particle *–to* in Hindi is a discourse salience marker and it imposes salience on the lexical category it marks. The interpretative effects of topicality and contrast can be derived from its semantic core of imposed salience and the syntactic position as well as contextual pragmatic information of the sentence. This particle has no role to play in the syntactic derivation of a sentence but affects its interpretation (and subsequent convergence) at the interface with the external performance systems.

Chapter 1 has introduced the object of enquiry — the enclitic particle *–to* in Hindi — in terms of its description in the Hindi grammars as well as reviewing the available theoretical proposals for it in the linguistics literature. It has specified the theoretical questions that this thesis deals with. The goal of unpacking this particle at the syntax-semantic-pragmatics interface began by exploring the empirical distribution of this particle in the language.

Assuming a topic marker analysis (Kidwai 2000) as the null hypothesis, Chapter 2 questions the basic premise of directly mapping the particle *–to* to a discursive role like topic. To this effect, section 2.1 discusses the concept of Information Structure, gives an inventory of various IS notions like the dichotomies of topic-comment, focus-background and given-new information. The various frameworks that formalize the IS notion of topic are discussed in section 2.2. The tests proposed in literature that diagnose the topical status of a constituent as well as some formal models that are topic-oriented are discussed in this section. Section 2.3 diagnoses the status of topics in Hindi using the diagnostic tests and concludes that topic interpretation is not contingent on marking by this particle in Hindi. It also exhibits that this particle behaves differently from the canonical topic particles like Japanese *–wa* and Korean *–nun*. Assuming language design to be economical and non-redundant, it questions the motivation for selecting this particle from the lexicon if topic interpretation is not contingent on its presence.

To further strengthen the claim of dissociating topicality from this marker, Chapter 3 claims that a non-topic interpretation can be accorded to a *–to*-marked XP in Hindi. The methodology followed here categorized lexical entities into a topical type versus a non-topical type based on the different theoretical frameworks that model topics. Section 3.1 exhibits that this particle can mark non topic-type entities like adjectives, quantifiers etc. Section 3.2 diagnoses sentence initial position to be biased for a topic interpretation and then proceeds to show that the particle *–to* can mark topic type entity (an NP) that does not receive a topic interpretation (because it is in-situ and not in a topic bias position). Based on these factors, it is claimed that the enclitic particle *–to* is not a topic marker since it marks non-topical entities and not all entities it mark receive sentence topic interpretation. The discourse interpretation of a *to*-marked object NP is that of a contrastive relation with a contextually available set of alternatives. This pure contrast interpretation is facilitated via an uncertainty implicature and a scalar implicature as discusses in section 3.2.4. A derivation of these implicatures is proposed in section 3.3 that analyses them to be conversational implicatures.

Chapter 4 proposes that the meaning of particle *–to* cannot be accounted for in terms of truth conditional semantics since it does not affect the truth conditional meaning of a sentence (section 4.1). Adopting a multidimensional model of meaning, the proposed meaning of this particle is of a use conditional type. Section 4.2 discusses the concept of discourse salience and proposes that the meaning conveyed by the particle *–to* is to impose discourse salience or speaker salience on the entity it marks. Marking of topics and contrast are the main motivations for imposing salience on a constituent (Mulkern 2007). This meaning has been formalized as a procedural meaning (section 4.3.1), as an expressive meaning (section 4.3.2) and as a use-conditional meaning (Section 4.3.3). Adopting a Hybrid Semantics framework, the meaning of a sentence containing a *–to*-marked constituent is proposed to contain a tuple of a truth conditional content (modeled as a set of world where the sentence is true) and a use conditional content (modeled as a set of contexts where the sentence is felicitous). Section 4.4 situates this particle within the typology of use conditional items (Gutzmann 2013) and analyses it to be functional, expletive non-shunting UCI.

Section 4.4.2 exhibits that this particle displays all the characteristic properties that identify a UCI.

Chapter 5 deals with the syntax of the particle *–to*. The Interactional Spine Hypothesis framework (Wiltschko 2021) is discussed in section 5.2. Section 5.3 lists the advantages of adopting this theoretical framework over others and proposes an analysis of syntactic incorporation of this particle within a clausal spine that projects an interactional domain above a propositional domain. The grounding layer in the interactional domain syntactically manifest an interlocutor's ground or mental state. It is proposed that the particle *–to* associates with a $\text{Ground}_{\text{speaker}}$ functional head to assert that the speaker holds a propositional attitude specifically relative to the salient marked entity in their ground. A syntactic derivation of this proposal is discussed in section 5.3.2. Section 5.4.1 attempts a featural decomposition of this particle and claims that this particle is bundle of a phonological feature and a semantically relevant interface feature [UCI]. It has no syntactic feature that is relevant for agreement or movement operations. Section 5.4.2 proposes an analysis for the merge of this particle with a lexical root within the Problems of Projection framework (Chomsky 2013, 2015). Section 5.4.3 proves that this particle does not interfere with syntactic agreement operation or with movement operation within the narrow syntax. A tentative proposal about the nature of interface conditions that seem to be operative for licensing of the discourse salience interpretation of this particle is floated.

Chapter 6 traces the development of the argument of this thesis through the chapters 1 to 6 in section 6.1. Section 6.2 conjectures about the theoretical implication of the thesis proposal for the architecture of grammar in light of the minimalist goals of 'genuine explanation'. It suggests a tentative tripartite modeling of the Conceptual Intensional interface such that distinct modules license distinct semantic interpretations to a sentence.

# CHAPTER 2

# INFORMATION STRUCTURE AND PARTICLE *-to*

The preceding chapter introduced the Hindi enclitic particle *-to* and discussed its distribution with various phrasal categories, context-sensitive particles and other lexical items in the language. Kidwai's analysis of *-to* as a topic marker (2000) has been taken as the starting point for this thesis. What it means to be a 'topic' (as an information-structural notion) was only briefly mentioned, without delving into any expansive discussion of it or its status in the linguistic theory. This concern brings us directly to the *topic* of this chapter, moving from the broader question of 'what is information structure' (section 2.1) to the narrower question of 'what is a topic' (section 2.2) since topics are unanimously considered to be a part of the information structuring of a sentence. After setting the base for enquiry, the status of *to*-marked XPs as a mechanism of topic-marking for the information structuring in Hindi is analyzed in section 2.3 in this chapter.

## 2.1 Information Structure

### *2.1.1 Introduction*

Linguistic theories have looked at both *what* a speaker says (the truth-conditional content of a proposition) and *how* they say it (the structuring or encoding of a proposition). Two (or more) sentences can have exactly the same propositional content. For example, the pair in (1a) and (1b) are interpreted to be 'true' in exactly the same situations.[1]

[1] In formal semantics, examples (1a) and (1b) would both hold true in all those possible worlds where the speaker belongs to the set of entities that like Linguistics in those worlds. This can be formally represented in predicate logic as: (where M is the model; w - world variable; g - assignment function)

[| (1a) |]$^{M,w,g}$=1 iff for that world w, {speaker} ϵ {x| x likes Linguistics}

[| (1b) |]$^{M,w,g}$=1 iff for that world w, {speaker} ϵ {x| x likes Linguistics}

Thus, the truth-conditions for the examples (1a) and (1b) are the same.

(1) a. I like Linguistics.
b. Linguistics, I like.

However, these sentences have different structures and their structuring is sensitive to the discourse in which a sentence has to be anchored. To exhibit this, consider the question-answer pair (the smallest expository unit possible for a discourse) in (2) and (3). If a discourse context contains a preceding question like example (2 Q) below, then only (1a), repeated here as (2 A.1), is considered as a felicitous response to the question in this context.[2] Uttering example (1b), repeated here as (2 A.2), would lead to infelicity, indicated by the '#' mark before the sentence.

(2) Q: What do you like?
(2) A.1: I like Linguistics.
A.2: # Linguistics, I like.

If the discourse is modified to contain a preceding question like example (3Q), then this infelicity is removed. In such a context, uttering example (1b), repeated here as (3A), is a felicitous response. Examples (1)-(3) thus exhibit that structuring of a sentence is sensitive to the discourse in which it is uttered.

(3) Q: Who likes Linguistics (between you and your sister)?
A: Linguistics, I like.

Thus, Language is a higher-order system that not only provides a means for communicating a thought (between a speaker and an addressee) but also provides the means to package the information contained in a proposition according to the contextual requirements of the discourse participants.[3] This structuring of a sentence, according to the requirements of the interlocutors, is called the Information Structure

[2] A note about labelling of examples in this thesis: In those examples that contain a question-answer pair, the label contains the example number followed by Q for question and A for answer. In case of multiple answers to the same question, the answers are labelled as A.1, A.2, A.3 and so on.

[3] A sentence is 'informational' in the sense of information defined as a reduction of uncertainty. A sentence effects a change in the knowledge repository of participants in conversation by reducing the uncertainty (that existed in a person's knowledge repository before the utterance move was made) about something in some manner. According to Vallduví (1992: 13), "the information carried by two sentences with equal propositional content is different when the reduction of uncertainty they bring along to the hearer's knowledge-store is different".

(IS) of the sentence. As can be inferred from this formulation, IS is an interface phenomenon that deals with both — the linguistics-specific modules of sentence derivation and processing (like syntax, semantics, morphology, prosody and pragmatics) and "the extra-linguistics aspects such as interlocutor's psychological perception of the world" (Féry and Ishihara 2016). Over the years, multiple approaches have been taken to study IS and its grammatical encoding in a language. These are discussed in sub-section 2.1.2. The common denominator underlying all these approaches to IS is that it is presumed that IS interacts with both grammar and discourse.

According to one view, the basic requirements of IS are: anchoring of the linguistic message in the discourse and guaranteeing 'informativeness' in it (Strawson 1964). Two IS notions are posited that correspond to each of these IS requirements: *topic* (for anchoring of sentence and maintaining discourse coherence) and *focus* (for adding new information). A 'notion' of IS is understood to be "the formal and communicative aspects of language for the expression of information structural roles" (Féry and Ishihara 2016: 2). However, since IS interfaces with non-linguistic domains too (like an interlocutor's 'psychological perception of the world'), the description of IS notions needs to be expanded to include these extra-linguistic mental states as well as the concepts of location or temporality that are crucial for anchoring of a sentence in discourse.

The literature on IS is vast with a huge diversity of theories as well as definitions of such notions that are related to information structuring of a sentence. No uniform treatment of IS or its grammatical encoding across languages is available. However, there does exist consensus amongst IS linguists that topic (of a sentence/discourse), focus (in a sentence) and 'contrast' (if contextually available) are to be considered the basic IS notions according to which a speaker structures their utterance.[4] These notions are discussed in sub-section 2.1.3. Cross-linguistic and language-specific

[4] Contrast as a linguistically relevant notion has been added to this basic IS list via recent proposals like Molnár (2002, 2006), Repp (2010), Neeleman and Vermeulen (2012). Focus has received the maximum attention on its discourse-semantic properties and manifestations in language (cf. Molnár and Winkler 2006).

studies have shown that these IS-related notions have an observable reflex in a language at the levels of prosody (intonation), syntax (linear order), morphology (topic or focus marker) and interpretation (quantifier scope).

In a vein similar to Krifka (2008), this thesis avoids the conflation of terminology evidenced in most literature on IS, and maintains a conceptual distinction between an IS constituent, an IS referent and an IS interpretation — all subsumed often under a single term. For example, a topic constituent, a topic referent and a topic interpretation are all distinct. A topical constituent is an expression in the sentence that is marked for a topical interpretation. A topical referent or entity is the real-world denotation of a sentence-based topical constituent.

A major theoretical issue that IS has faced since its beginning is its incorporation into linguistic theory; i.e., depending on the architecture of grammar posited, different theoretical frameworks have either integrated discourse (and IS) within the grammar or located it at the periphery. The former includes multi-level models of grammar that have a parallel correspondence architecture where computations for phonological, syntactic, semantic and (some) pragmatic representations proceed in parallel. Such grammars incorporate IS in the architecture better since IS is "integrated as a level on a par with argument structure, semantics, etc." (Erteschik-Shir 2007:55) in them. Seminal works in these multi-layered theories are: Lexical Functional Grammar (Vallduvi and Engdahl 1995), Head-driven Phrase Structure Grammar (Pollard and Sag 1987; Sag and Ginzburg 2000), Combinatorial Category Grammar (Steedman 1996, 2000b), Role and Reference Grammar (Van Valin 1993b, 1999b; Van Valin and LaPolla 1997), Functional Syntax (Kuno 1987, Kuno and Takami 1993) and Functional Grammar (Dik 1997a). The latter category included much of the early generative enterprise, which had not given much importance to information structuring of a sentence (until Rizzi 1997) and had kept IS at the boundary of syntax, semantics and phonology part of computation.

Generative grammar has looked at language primarily as an instrument of thought. Chomsky (2000:75) assumes that "language is not properly regarded as a system of

communication. It is a system of expressing thought".[5] Within a generative framework, the focus is on accounting for human mental Faculty of Language (FL) that has a basic computational block (narrow syntax) that takes input from the Lexicon and whose output, a Phonological Form (PF) and a Logical Form (LF), interfaces with two language-external performance systems: the Articulatory-Phonetic (AP) system and the Conceptual-Intensional (CI) system respectively. Linguistic properties are encoded as features of lexical items in the Lexicon. Spell-Out is a point in derivation where once a unit of computation is completed (a 'phase'[6]), its phonological features are removed and shipped off to the PF component, before the derivation proceeds towards the LF component. This model of grammar is represented in figure 2.1.

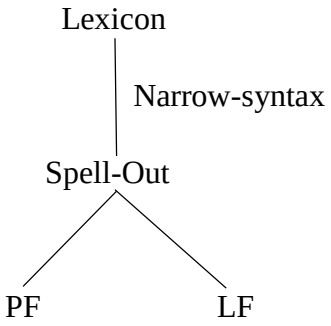



Figure 2.1. Inverted Y-model of Grammar

Incorporation of IS in this inverted Y-model of grammar poses some serious theoretical challenges. Early minimalist frameworks (Chomsky 1995, 2000) assumed that all syntactic operations are feature driven. So for IS notions to be operationalizable within this formal system, they have to be formalised as syntactic IS features. A featural analysis of IS notions and their typological status (as formal or substantive features; syntactic, morphological, or phonological features) is however a

[5] Contra other approaches, like Discourse Grammar, that analyse language as an instrument of communication and focus on accounting for context-dependent aspect of meaning (cf. Winkler 2005).

[6] In Minimalism, phases are the smallest propositional units that are isolatable and independent (for example, they can stand alone as a fragment answer). CP and v*P are the two phases in a sentence according to Chomsky (2001).

matter of debate with different positions taken by different linguists.[7] IS notions have an effect at both PF (linear order, prosody) and LF (discursive interpretations, scopal relations with quantifiers etc.). Since there is no common interface between PF and LF outputs in this model, IS features have to be posited as being present in narrow syntax to be available for both the PF and LF outputs. Chomsky, in his *Minimalist Program* (1995), posited a principle of 'Inclusiveness' that an optimal FL and an efficient system of sentence derivation follows. According to this, "any structure formed by the computation…is constituted of elements already present in the lexical items selected for the numeration; no new objects are added in the course of computation apart from rearrangements of lexical properties" (Chomsky 1995:228). Therefore, if IS features are not to violate 'inclusiveness', they must be included in the feature geometry that comprises a lexical item. This is again problematic since an element is not topical or focal in the lexicon but rather it 'becomes' a topic or focus during the course of derivation. Thus, incorporating IS in a generative model of grammar is challenging. Some proposed models of IS that deal with these challenges are discussed in sub-section 2.1.4 of this chapter.

### *2.1.2. IS: Different Perspectives*

The research area of IS has attracted attention since the end of 19th century but the term 'Information Structure' was coined only in 1967 by Halliday.[8] Over the years, multiple approaches have been adopted to analyse IS as a linguistic phenomenon and it has been given a different label in each framework, like 'Information Structure' by Halliday (1967), 'Packaging' by Chafe (1976), 'Information Packaging' by Vallduví (1990), 'Topic-Focus Structure' by Erteschik-Shir (1997) etc. This section highlights some of the different perspectives that have been adopted to define the notion of IS itself in the literature.

[7] Erteschik-Shir (2006a) posits an optional assignment of [topic] or [focus] features as lexical features to an item in the lexicon. Another approach is taken by Miyagawa (2010, 2017), who creates a new syntactic feature category of discourse features (parallel to phi features) that are valued as topic or focus during the course of derivation.

[8] von der Gabelentz (1869) used the term 'psychological subject' to denote an object that a speaker is thinking about, and the term 'psychological predicate' to refer to what the speaker is thinking about that object.

Mathesius, the founder of Prague Linguistics Circle, has been called 'the father of modern IS' (Lambrecht 1994). He reformulated the psycholinguistic terms of von der Gabelentz into IS notions, namely that of a 'theme' — what a sentence is about — corresponding to Gabelentz's 'psychological subject'; and a 'rheme' — what is being said about the theme —corresponding to his 'psychological predicate'. Example (4) is an English sentence that exhibits this theme/rheme bifurcation of a sentence in a language.

(4) [The Earth]THEME [revolves around the Sun]RHEME

In the above sentence, the constituent *the Earth* is the theme since the sentence is about it. Rheme, which corresponds to what is said about the theme, is the predicate *revolves around the Sun*. As can be observed in (4), a subject-predicate type of sentence construction directly maps onto a theme-rheme articulation of a sentence. This binary distinction has been translated into the terms 'topic' for theme and 'focus' for rheme by some linguists (see Féry and Ishihara 2016).[9]

The Prague School tradition was characterized by a pragmatic and functional flavour of linguistic analysis. This tradition was taken up by Firbas (1964, 1966) and later on by Sgall, Hajičová and Panevová (1986). Firbas worked within a Functional Sentence Perspective framework and developed a theory of dynamic communication. He coined a term — Communicative Dynamism (CD) — a concept that is gradient in nature and provides the degree to which a sentential expression pushes the communication forward. Information Structure is assumed to have a direct relation with the linear order of constituents in a sentence, which is accounted for in terms of the degree of CD of the constituent. Elements that are contextually known do not contribute to the communication and have the least CD. Elements that add new information in the discourse have, in contrast, the highest CD on a scale. This theory posits that elements with a lesser CD value linearly precede those with a higher CD value in a sentential structure. Theme, then, is "the sentence element (or elements)

[9] However, this direct mapping of terminology is problematic since focus can be contained inside topic but rheme cannot be contained inside a theme. This is exhibited in section 2.1.3, where topics and focus are discussed.

carrying the lowest degree(s) of C[ommunicative] D[ynamism] within the sentence" and must be the first element in a sentence. The rheme or the remnant is that important part of the sentence that "pushes the communication forward" (Firbas 1964: 272). Later on in Prague School, a second theory was developed by Sgall, Hajičová and Panevová (1986) — the Topic Focus Articulation (TFA). This theory was based on the criteria of contextual boundedness. Within a TFA framework, topics are generally context bound whereas focus is the element(s) that is not context bound.

While the Prague School linguists adopted a functional-pragmatic perspective to analyse IS, Halliday based himself within phonology to account for this phenomenon. According to this approach, Information Structure is what accounts for "the distinction of focus, presupposition and propositional attitude towards entities in the discourse conveyed by phrasal intonation" (Halliday 1967). Focus is what cannot be 'recovered' from the discourse that preceded the sentence. Prosodic cue of intonation marks the newness or oldness of a phrase.

A psycholinguistic perspective was taken by Chafe who formulated a theory of 'Packaging'. In this theory, what gains relevance is the "states of the addressee's mind" (Chafe 1976). Thus, focus shifts to the hearer's mental world and his current attentional state. 'Packaging' of information in an utterance aims at satisfaction of the temporary communication requirements of the addressee's mind. The choice of a particular form is based on either some pragmatic function or it is based on the requirement of accomodating the mental state of the addressee. For example, asking of a question creates a particular attentional state. The packaging of information in the answer has to recognize the addressee's attentional state and highlight (focus-mark by pitch accent or other strategies) only the new information that was asked for.

Information Structure has also been given a purely syntactic account by linguists like Erteschik-Shir (1996), Frascarelli and Hinterhölzl (2007), López (2007), Miyagawa (2010, 2017), Holmberg (2020) etc. To mention one such approach at this point, Erteschik-Shir (1996) describes Information Structure syntactically as the meta-structure that gets annotated over the hierarchical structure exiting the derivation at

Spell-Out. Topic and Focus are the two IS features that the sentence structure is annotated for. A sentence has to obey the constraints of all — the Phonological Form, the Logical Form and the Information Structure — to be uttered grammatically and felicitously in a given discourse.

A cognitive approach has also been taken by linguists to analyse IS. According to Zimmermann and Féry (2009: 1) "Information Structure is an intermediate cognitive domain that operates between the strictly linguistic modules of syntax, morphology and phonology and the broader cognitive domains that fix an individual's system of beliefs via the process of pragmatic reasoning, information update and inference mechanism". In this view, broader cognitive domains, whose function is to establish and regulate a person's belief and knowledge states, are brought inside the study of IS. The formal expression of IS in language has an impact on the cognitive processes of inference and reasoning. The locus of information storehouse is the larger cognitive domain in a speaker's mind. Overt marking of IS notions (like topic, focus or given versus new) on a linguistic unit facilitates the process of information update for a speaker and also helps in actualization of their belief states. This marking may also have some pragmatic and cognitive effects like increasing the salience of a discourse referent or triggering certain implicatures.

The cognitive modules-based perspective, followed by Zimmerman and Féry, finds a resonance in Krifka (2008). Krifka's theory of IS is hinged on the notion of Common Ground[10] (CG) and the dynamics of communication. Krifka (2008: 15) defines CG as "information that is mutually known to be shared and continuously modified in communication". This is relevant for IS because the main purpose of communication is "transfer of information and its optimization relative to the temporary needs of interlocutors" (Krifka 2008: 15). In informal terminology, CG consists of all the propositions (and in recent work, even discourse referents) that are assumed to be shared belief or shared knowledge between the interlocutors. Doxastic and epistemic

[10] The origin of the term 'common ground' is found in Grice (1989), where it roughly means the presumed background information that is shared amongst the participants of a conversation. For seminal work on the concept of Common Ground, see Stalnaker (1974, 2002).

modal logic are invoked to give a semantic account of 'belief' and 'knowledge' respectively. A formal semantic definition of CG is given in the textbox below.

> For any proposition Ø is "it is common ground that Ø in a group if all members accept (for the purpose of the conversation) that Ø, and all believe that all accept that Ø, and all believe that all believe that all accept that Ø" (Stalnaker 2002: 716).

Krifka assumes a Chafe-style communication requirement based 'packaging' of information in a sentence. Each utterance by the participants of a discourse 'updates' the material available in the current common ground between those participants. This material is called the CG Content. This CG content continuously undergoes some modification (addition of new information, deletion or change of previous information) in a discourse. CG Management deals with the way CG content is acted upon in an interaction between participants so as to be relevant to the current discourse. The CG that is input to a discourse consists of 'presuppositions' and this varies from the resultant CG at the end of a discourse because the latter has assertions or 'proffered content' added to it. Within this framework, Information Structure is functionally related to both the CG content and the CG management functions. He posits that "associate those aspects of IS that have truth-conditional impact with CG content, and those which relate to the pragmatic use of expressions with CG management" (Krifka 2008:246).

While there are several viewpoints about the formal expression of IS in language, it itself has been presumed to be universal for all natural languages (Zimmermann and Féry 2009). The source of cross-linguistic variation comes from the mechanism of expressing an IS notion in the grammar of a particular language. The next sub-section takes stock of these IS notions. Later, section 2.2 discusses the one IS notion that is relevant for probing *to*-marked XPs in Hindi in detail - *topic*.

### *2.1.3 An Inventory of IS Notions*

Within IS, a linguistic message is structured at different dimensions according to the communicative intentions of the speaker as well as the requirements imposed by linguistic and extralinguistic contexts. Sentences can be decomposed at various levels to see the different criteria of information structure orientation. These levels provide the following binary articulations of a sentence:

- Topic-Comment
- Focus- Background
- Given/Old- New

A *topic-comment* structure is intuitively similar to the theme-rheme distinction that was mentioned in section 2.1.1. The concept of *topic* does not have a unanimously accepted definition but the most accepted one is from Reinhart (1981) — topics are what the sentence is 'about'. The rest of the sentence conveys information about it and is called a *comment* for that topic. Example (5) exhibits a sentence decomposition into a topic-comment sub-structure based on the criteria of what the sentence is about and what information is conveyed about it.[11]

(5) Context: A group of classmates is discussing their school principal. Specifically, his attributes are being listed. One student mentions:

[He]TOPIC [is also rich]COMMENT

At the second level, a sentence can be decomposed into a *focus* part and a *background* part. Roughly, focus is the information towards which the speaker wants to draw the hearer's attention and marks it as more prominent. The non-focus parts of a sentence are the background for that focus. Example (6) shows this bifurcation and a focus-marked constituent:

(6) Context: A teacher is counselling her students after they received their report cards. She is telling students where they need to increase their effort. She says to one student:

[11] Since most expressions of a sentence have the potential to be the topic/focus of a sentence, expository sentence for any IS notion requires contextual grounding to demarcate which expression gets informationally structured. So, an overt context is specified before introducing a sentence that is IS-marked.

[You should work harder in]BACKGROUND [Physics]FOCUS

At the third level, a distinction is made between contextually *old* (already mentioned either implicitly or explicitly and hence, *given* in the discourse) or contextually *new* information (no prior mention in the discourse) in a sentence. Example (7) shows this decomposition of a sentence into the given/old-new frame of IS:

(7) Context: It is lunch time and two office employees are discussing food options in their canteen. Speaker A asks speaker B regarding what he wants to eat, to which Speaker B responds:

[I want to eat]OLD [Subway from outside]NEW

In the context above, Speaker A's enquiry is responsible for making speaker B's 'I want to eat' as given in the discourse because of its implicit mention in A's question. The new part that gets added to the common ground is 'Subway from outside'.

IS-notions often overlap on a constituent i.e., an expression in a sentence can simultaneously convey more than one information structural role. It is a common tendency in natural language for topics to be given and for focus to be the new part of the sentence or for focus to correspond with the comment part and for topic to correspond with the background part. This is exhibited via the sentence pair in example (8).[12] The example data in (8) and (9) are taken from Hinterwimmer (2011: 1876), diagrammatic representations in figures 2.2 and 2.3 are mine.[13]

(8) a. Tell me something about John.
b. John married BERTHA.

[12] A conversationally grounded set, of either two coherent sentences (like example (8)) or a congruent question-answer pair, is the smallest unit of discourse that provides sufficient contextual information without prior mention to assess IS packaging of an assertion.

[13] Note about a stylistic convention followed in the original text and repeated here: the constituent getting the main prosodic stress (the object NP in (8b)) is capitalized as an intonational cue.

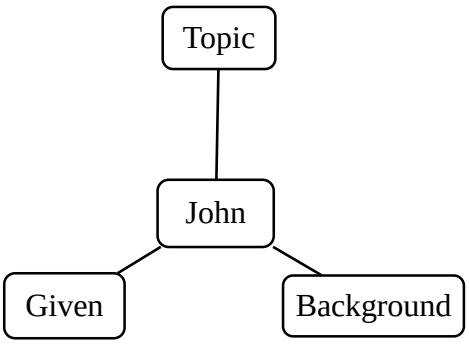



Figure 2.2: Co-incidence of IS notions on the subject *John*

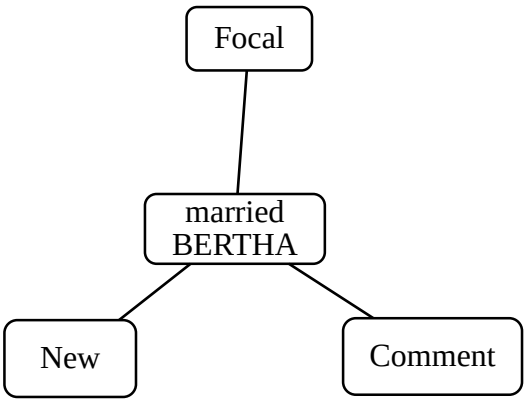



Figure 2.3: Co-incidence of IS notions on the predicate *married BERTHA*

‘John’, having prior mention in example (8a), is consequently given for (8b) and also the topic about which that sentence provides information. It is also the background in this example since it does not draw the hearer’s attention. In contrast, the predicate ‘married Bertha’ is the new information that also forms the comment part for the sentence topic. It is more prominent (stress on ‘Bertha’) and draws the hearer’s attention and hence, focal for example (8b).

However, this overlap of IS notions on an expression is not always obtained. A minimal change in context from (8a) to (9a) is sufficient to change the co-incidence relations obtained in (8b) (that were represented via figures 1 and 2).

(9) a. Tell me something about John. Who did he marry?
b. John married BERTHA.

In the pair in example (9) above, the first sentence establishes the aboutness topic for its subsequent response. Thus, 'John' is the topic for (9b) while the remaining 'married Bertha' becomes the comment part for this topic. The given part is not just 'John' but 'John married' (since this was mentioned in (9a)). This is also the background part and the only focused constituent that requires hearer's attention is 'Bertha' in (9b).

Besides the overlaps of IS notions mentioned in example (9), topics can also introduce novel information and given items can be focused too. In example (10) below, the topical constituent 'a friend of mine' is introduced in the discourse and adds new information to it. This topic is not given or mentioned prior in the discourse, also it is morphologically indicated by use of the English indefinite article 'a'.

(10) [A friend of mine]TOPIC got married this year.

Within a system that adopts the above mentioned criteria for identifying the topic and the focus of a sentence, a constituent that is topical may also be focal in the same context. In example (11), the question specifies the referent set (denotations of {you, your brother}) from which the topic element of the answer utterance has to be selected. Example (11 A) is about 'I', so it is topical. It is also the highlighted part that seeks its hearer's attention. Thus, it is the focal constituent too for this sentence.

(11) Q. Who drank the wine, you or your brother?
A. [[I]FOCUS]TOPIC drank the wine.

Thus far, certain IS notions were introduced informally, based on the various sentence dimensions that incorporate them on a binary distinction basis. These notions were described as an information structural role that a constituent expresses. Languages have a tendency to allow co-occurrence of multiple information structural roles on a single constituent. This section has also highlighted the fact that any theoretical enquiry into IS or the linguistic reflex of an IS notion requires a battery of diagnostic tests (that are sensitive to all contextual variables in the discourse) to verify the status of an IS notion role that is involved in any IS-related linguistic phenomenon.

Information structure theories have evolved extensively over the years (see Jacobs 1983, Rooth 1985, 1992, Rochemont 1986, Erteschik-Shir 1997, Vallduví 1992 to name a few). Formal models have been proposed to deal with the IS-related linguistic phenomenon. Some such IS models are presented in sub-section 2.1.4. Many linguists have had a fundamental impact on our current understanding of each IS notion. Apart from topic, which is discussed in section 2.2, I only cite the relevant works here that have analysed other IS notions.[14] For a formal treatment of focus see Rooth (1985, 1992), Selkirk (1984) etc; for givenness see Schwarzschild (1999), Rochemont (2014).

To expand upon one approach that has provided formal definitions of the basic notions of IS (and is widely accepted), I discuss Krifka's framework of IS (introduced in section 2.1.2) next.

*Krifka (2008)*

Krifka assumes a communicative model of language where communication entails an exchange of information between the speaker and the hearer. Information is not stored as an unstructured set of propositions but is rather stored according to a sorting/organizational algorithm. Information is stored in Common Ground (CG) and this CG is continuously modified during the course of communication by each utterance move. In his theory, CG has two dimensions — CG content and CG management. This distinction is employed to "associate those aspects of IS that have truth-conditional impact with CG content, and those which relate to the pragmatic use of expressions with CG management" (Krifka 2008: 246).

To define focus, Krifka adopts an alternatives-based approach proposed by Rooth (1985, 1992). Rooth's theory of *Alternatives Semantics* (1985, 1992) claims a direct correspondence between focus-marking of a constituent and the availability of

[14] Topic-marking is the only grammatical reflex of IS notions (from those mentioned in section 2.1.3) that directly bears upon the research question of this thesis.

alternatives to it.[15] "Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions" (Krifka 2008: 247). He builds upon this claim to define a focus property F that a constituent in a language can have. He also defines givenness formally as a feature of an expression integral to his theory of CG. An entity type, CG function related definition of topic constituent was also proposed by him. These formal definitions of basic IS notions are cited below in the textboxes.

> "A property F of an expression α is a Focus property iff F signals
> (a) that alternatives of (parts of) the expression α or
> (b) alternatives of the denotation of (parts of) α are relevant for the interpretation of α." (Krifka 2008: 248)

> "A feature X of an expression α is a Givenness feature iff X indicates whether the denotation of α is present in the CG or not, and/or indicates the degree to which it is present in the immediate CG." (Krifka 2008: 262)

> "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." (Krifka 2008: 265)

The distinction of CG content and CG management, proposed by Krifka's theory, is particularly relevant for analysing the IS-notion of focus and its grammatical encoding in language. Focus has both a pragmatic use (exhibited via examples (6), (8) and (9) above) and a semantic use. The former dimension falls within CG management and the latter in CG content. Szabolsci (1981) has argued for truth conditional relevance of focus for Hungarian that has a designated focus position. Languages also employ focus-sensitive operators, like *only* and *even*, which associate with a constituent in a sentence (cf. Jackendoff 1972, Rooth 1985, Krifka 2008). Example (12) exhibits a sentence whose truth value depends on the constituent (the direct object in example (12a) and the indirect object in example(12b)) that gets associated with a focus particle *only*. This example is from Molnár *et al* (2019: 6).

[15] The set of alternatives (called F-Alternatives) are formalised as a set of propositions that contain a variable in the place of the focused constituent. According to Büring (2016), this focusing is an operation of assertion that 'asserts the exclusion' of alternatives. The focused proposition is the only true assertion from the set of alternatives.

(12)a. John only showed Mary the [PICtures]$_{FOCUS}$.

b. John only showed [MAry]$_{FOCUS}$ the pictures.

Krifka (and many others) propose a typology of focus based on the pragmatic function it performs in a sentence. These include information focus, contrastive focus, exhaustive focus, scalar focus, emphatic focus and identification focus. Another typology based on the criteria of the nature of constituent getting focus is: broad versus narrow focus, verum focus, sublexical focus.

### *2.1.4 Some Formal Models of IS*

The issue of integration of Information Structure in linguistic theory still remains as an open question. Differing approaches have been adopted by a formal system of grammar like Generative Grammar as opposed to Discourse Grammar (refer to section 2.1.1). In the former, the syntax-semantics interface is responsible for yielding the information structural interpretations along with the scopal relations. In the latter, information structure interpretation is relegated to (Discourse-) pragmatics — a component that is either separate from rest of the derivational blocks or is assumed to interface with the LF component. Since this thesis is based within the generative framework, I restrict my attention to formal grammar-based approaches to IS. In this section, I provide a brief outline of some of the formal models of IS proposed in the literature, focusing initially only on models that theorise both topic and focus features together.

*F-Structure: Erteschik-Shir (1997)*

*F(ocus)-structure* is a model for IS introduced in Erteschik-Shir (1997). It uses the metaphor of a file card updating system. The discourse referents in a context set are analogous to a set of file cards, which are the potential topics for that discourse. “If the attention of the hearer is drawn to (the referent of) X, then the hearer (metaphorically) selects the card for X and puts it in a place of prominence, namely on top of his stack of file cards” (Erteschik-Shir 1997). In this system, *topic* and *focus*

interact and both act as triggering instructions that manipulate this stack. A set of 'f-structure rules' are posited according to which this stack of file cards is manipulated.

These f-structure rules include:

(i) Topic tells the hearer to find an existing card on top of his file

(ii) Focus tells the hearer to make a new card (for indefinite) or find an existing card (for definite) and put them on the top of the file

(iii) An update operation makes the hearer modify the topic card to enter focus on it.

A topic-focus pair marked in the annotation of a sentence is called its F-structure. This model stipulates that F-structure boundaries must be phonologically visible since they can be mapped onto intonational phrases.

*Σ-structure: Zubizaretta (1998)*

Within the generative enterprise, Principles and Parameters model (Chomsky 1981) and the Minimalist Program (Chomsky 1995, 2000) both pose problems for an IS architecture (refer to section 2.1.1). Two major theoretical concerns were raised for direction incorporation of IS within the architecture of grammar. Firstly, since grammatical encoding of IS affects both PF and LF representations, IS features must have access to both. But there is no direct interface between LF and PF outputs in the generative model of derivation. Secondly, the Inclusiveness Condition (that no new object be added in the course of derivation) gets violated unless topic and focus features are introduced lexically.[16]

Zubizaretta (1998) weakens the inclusiveness condition by introducing Σ-structure — a (post-syntax) point in the derivation upto which a single phrase marker has been derived. In this Σ-structure two features, [F](=focus) and [prosodic prominence], can

[16] Erteschik-Shir (2006a) overcomes the Inclusiveness problem by making topic and focus a matter of introduction by lexical selection. Each selection of a lexical item licenses an optional assignment of a [top] or [foc] feature, which "may percolate to the maximal projection of the lexical head they are assigned to".

be introduced. A post-syntactic but pre-phonological rule of P-movement is posited for this stage in derivation. P-movement or prosodically motivated movement can move non-focused material out of focused positions. Since P-movement has an impact on LF, Σ-structure feeds into the λ-structure or LF in this model. The output of LF is then sent to PF and to Assertion Structure.[17] The structure of the grammar proposed by her model (Zubizaretta 1998: 31) is represented in figure 2.4.

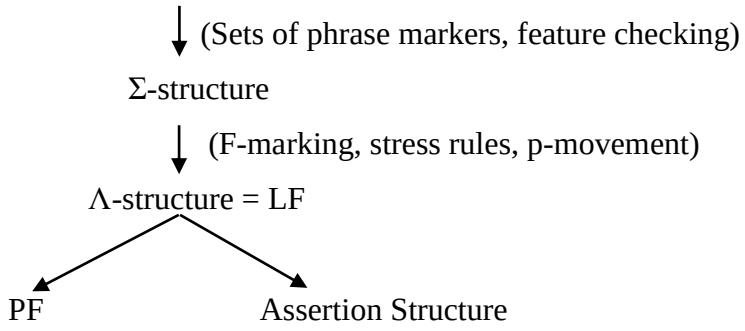



Figure 2.4. Zubizaretta's Model of Grammar

In her model, LF provides input to the PF. This is problematic as output of LF operations are not meant to be pronounced. For a detailed review of issues generated by this analysis, see Winkler and Göbble (2002).

*The Cartographic Framework: Rizzi (1997)*

Rizzi (1997) developed a cartographic approach and decomposed the single functional head C(omplementizer) in the left periphery of a clause into multiple functional heads. He posited functional heads of Force, Top(ic), Foc(us) and Fin(iteness) that project phrases of their own. Each functional projection maps onto a unique function or interpretation. ForceP specifies the clause type — whether interrogative, declarative or imperative; TopP hosts the topic of the clause; FocP hosts the focus or the wh-phrase of the clause; FinP marks the finiteness of the clause. There is a fixed order in which these functional heads map themselves in the left periphery of the clause. In this model, a clause can have only one focus phrase but it can have more

[17] Assertion structures are associated with implicit or explicit context questions (Erteschik-Shir 2007:59)

than one topic phrase. Figure 2.5 exhibits the split-up C domain in the clause structure in this model.

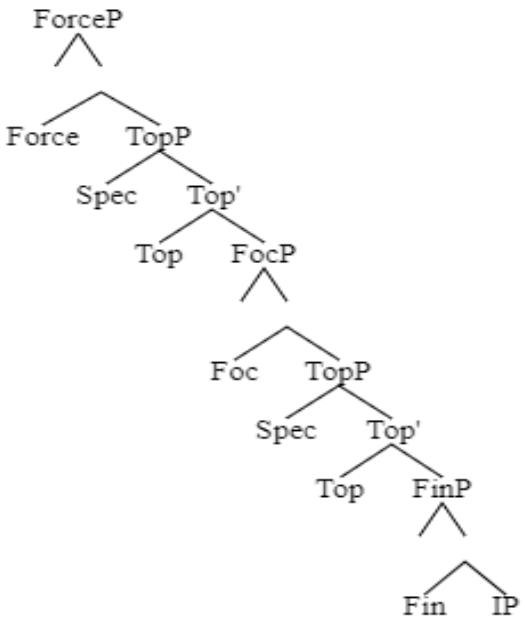



Figure 2.5: Rizzi's cartography of the left periphery

These spec-head configurations and the stipulated criterial realtionships are the mechanisms through which Rizzi derives information structural interpretations of Topic and Focus. A Topic Criterion is posited that requires that for any XP to get topic interpretation, it must be in a spec-head configuration with a Top functional head. Checking of [Top] feature requires this local configuration. IS related movement is a type of A-bar movement (not related to argument structure properties) and the permissible landing sites for a topic constituent moving out of IP are the two [Spec, TopP] positions in the clausal structure given in the figure 5 above.

To conclude, this section has presented an overview of a few models adopted by linguists to integrate IS and IS notions within the architecture of grammar in various frameworks. It is clear that a descriptively adequate model of IS, that further aims for explanatory adequacy, must integrate an account of grammatically coded syntactic, morphological or prosodic properties of a sentence with its discourse-semantic dimension. However, at this stage, I am not making a theoretical choice amongst the various frameworks.

Coming back to the issue of *to*-marked XPs in Hindi, this enclitic discourse particle has been analysed as an overt topic marker for the information structuring of this language. The next section presents the multiple definitions that are proposed for topics (section 2.2.1) and provides a brief literature review of IS studies focussed on *topic*-related linguistic phenomenon (section 2.2.2). This is followed by section 2.2.3 that specifies the battery of tests used by linguists to identify the topical constituent of a sentence. Section 2.2.4 outlines certain formal models that are oriented towards the IS notion of topic.

## 2.2 Topic

### *2.2.1 What is a Topic?*

Topics have been characterized as a concept relevant for organizing existing information and for storing new information. For example, Kuno (1972) analyses topics as 'sorting keys' that are used for filing and accessing information. This IS notion of topic is linguistically realized in two flavors: a Discourse Topic and a Sentence Topic (see van Dijk 1977, Reinhart 1981). Discourse topics have a broad scope (over the full discourse) as compared to the narrower scope of a sentence topic that is local to a sentence. Within a sentence, the discourse topic may overlap with a sentence topic but this is optional since a sentence can possibly have different discourse and sentence topics. In a context, violation of a sentence topic is more difficult to repair as compared to discourse topic violation.[18] Discourse topics have also been called as Question Under Discussion (QUD) topics (von Fintel 1994). Since most linguistic theorizing takes a sentence to be the object of its study, the basis of enquiry for the grammatical reflex of IS notion topic in language has also focused on sentence topics.

According to McNally (1998), sentence topics have either been treated as an entity or entity-type discourse referent (cf. Reinhart 1981; Portner and Yabushita 1994) or as questions that can be modeled as presupposed salient set of alternatives (cf. von Fintel

[18] This is provided as evidence by Reinhart (1981) for claiming different "psychological realities" behind a sentence topic and a discourse topic.

1994, Büring 1997, Roberts 1995). Out of all such formal analyses of topics, the most accepted one is the Aboutness (sentence) Topic posited by Reinhart (1981). She analyses 'topic of' as a relation of 'BEING ABOUT' that holds between two entities. Thus, 'X is a topic of Y' (where X and Y are two entities) means that 'X is about Y'. The relation 'topic of' is not a unique relation since multiple entities can be a 'topic of' something. Also, this is assumed to be a pragmatic relation (and not a semantic relation) which is relative to a discourse. This identification of sentence topics using the property of 'aboutness' has faced serious criticism in the linguistic literature (see Krifka 2008, Büring 2016).[19]

Many non-aboutness based definitions of *topics* have been proposed in the literature. For Chafe (1976:51), sentence topic is "the frame within which the sentence holds". Jacobs (2001) maintains a distinction between aboutness topics and frame topics. He proposes a formal definition of a 'frame' X for a structure — (X Y). X is "a domain of possible reality to which proposition expressed by Y is restricted" (Jacobs 2001:656). Whereas Chafe's domain of frames was restricted to a spatial, a temporal and an individual type; Jacob expanded this domain of frames to include conditionals as possible frame-setters.[20] Jacob's inclusion of conditionals as a 'frame' contrasts with Krifka's conception of a 'frame'. According to Krifka (2008), a frame includes "the general type of information that can be given about an individual". This can include a frame-setting adverb but must not contain 'specific' information like conditionals. Example (13) exhibits a frame topic ('financially') in a sentence in English.

(13) [Financially]$_{\text{FRAME}}$, he is content.

[19] Krifka states that the Reinhart-style definition of topic "[t]his presupposes that information in human communication and memory is organized in a certain way so that it can be said to be 'about' something. This does not follow from a general definition of information. For example, relational databases or sets of possible worlds, both models for information, do not presuppose any relation of aboutness" (Krifka 2008: 40).

[20] This is conceptually similar to Haiman (1978) who treats conditionals as topics. I suggest that the non clitic, free word *to* introduced in chapter 1 can be interpreted to be marking a frame of predication or a topic in the conditional or temporal bi-clausal constructions.

Another non-aboutness based definition of topic is attributed to von Fintel (1994), Roberts (2012). von Fintel proposed a concept of QUD that refers to a set of propositions in a discourse context. These propositions are potential discourse topics in that context. A sentence topic is claimed to be the most salient discourse topic in this approach. This claim has been falsified by Roberts (2011) who provides evidence for the same pair of sentences to have different sentence topic and discourse topic.

This section has highlighted the first difficulty that any research of IS-related linguistic phenomenon faces — that of defining the notion *topic*. This has been stated as "there is very little consensus among linguists on any (…) specific definition. Multiple properties contributing to topichood have been described, but none of these properties seem either necessary or sufficient to classify something as a topic" (van Bergen and de Hoop 2009: 173). The next section provides an overview of how the notion of *topic* and its grammatical manifestation has been viewed in some IS studies.

### *2.2.2 Overview of 'Topic' in the IS Literature*

Historically, the notion of *topic*, first developed as *theme* by Prague School (refer to section 2.1.2), was taken to be a contextual notion that determined the word order of a sentence. Topic (or theme) was the contextually 'known' element with the lowest C[ommunicative] D[ynamism] value and hence, preceded all other elements of a sentence. Linguists working within the same framework also exhibited that this contextual notion is fundamental for maintaining sentence cohesion in a discourse. Daneš (1974) proposed three types of thematic progressions that natural languages employ to maintain cohesion of successive sentences in a discourse. The first type is 'topic chaining' that involves keeping the theme constant across sentences.[21] This progression is exhibited in example (14). In the given sequence of sentences, the topic of second sentence 'he' is co-referential with topic of first sentence 'Ravi'. This constancy of topical referent across a discourse in (14) helps in maintaining its cohesion.

[21] Erteschik-Shir (2007) and not Daneš (1974) provides these labels for various types of thematic progressions. Daneš only elaborated on the mechanism of thematic progression.

(14) [Ravi]THEME likes to cook new dishes. [He]THEME has worked hard to improve his culinary skills.

The second type of thematic progression is 'focus chaining' that requires the rheme of a preceding sentence to become the theme of the subsequent sentence. This type of progression is shown in example (15); this example is taken from Erteschik-Shir (2007: 3).[22] The antecedent of the theme of the second sentence 'she' is the focussed constituent of first sentence 'girl'. Thus, focus of one sentence is 'chained' to become the topic of the subsequent sentence.

(15) There's a [girl]RHEME in the class who the teacher likes. [She]THEME answered all the questions the teacher asked.

The third type of thematic progression involves deriving topics from a 'hypertheme'— a set of referents that are available in the discourse. Either the members of the set denoted by hypertheme could be explicitly mentioned in the previous discourse or their availability could be inferred by the lexical meaning of a hypernym that is already specified in the discourse. The topic of a subsequent sentence is selected from the referent members of this hypertheme set. This progression type is exhibited in examples (16) and (17). A pair of sentences in examples (16) and (17) show the derivation of the same theme 'physics' in a second sentence, selected from the hypertheme set available in the first sentence. The members of this set are explicitly specified in example (16) but are inferred in example (17), since hypernym 'core sciences' has hyponym members {physics, chemistry, biology} under it.

(16) Let me tell you about [physics, chemistry and biology]HYPERTHEME. [Physics]THEME is an intuitive subject while biology is the easiest of them all.

(17) Let me tell you about [core sciences]HYPERTHEME. [Physics]THEME is an intuitive subject while biology is the easiest of them all.

In cross-linguistic IS studies, most linguists have assumed a one-to-one correspondence between an IS notion and the linguistic device used to encode that IS

[22] Theme/rheme annotation on a constituent in examples(14)-(15) is done by me and not available in the original text of this data.

notion in a language.[23] Thus, the concept of topic is often described in literature using the grammatical strategy (syntactic, morphological or prosodic) employed by a language to anchor it in discourse (for example: Danish (Erteschik-Shir 1997, 2007); Catalan (Vallduví 1992) etc.). One such syntactic strategy is that of topicalization (the others include — phonological de-accenting or specific intonational contour or morphological marking by particles, to be discussed later).

Topicalization is a type of syntactic movement that is motivated by the requirement of overtly marking a topic in a language. It dislocates a constituent to the left periphery of the clausal spine. The constituent moves from a clause-internal position and the landing site of such a movement operation is the specifier position of a functional Top(ic) head in the CP domain (cf. Rizzi (1997) amongst others). Within the binary distinction of A v/s A-bar movement in syntax (Mahajan 1990), topicalization is standardly assumed to be an A-bar movement (that is not motivated for any theta or case-related criteria). Languages vary in terms of their degree of preference for employing this mechanism to mark a topical constituent. For example, while Hungarian strongly prefers this movement, English rarely employs it and Danish keeps it optional. The phenomenon under discussion is exhibited via a minimal pair for English in example (18). The derivation of the sentence in (18b) involves topicalizing the direct object *him* of (18a) and moving it to the left periphery.

(18) a. I met him yesterday.
b. Him, I met yesterday.

The topical status of 'him' can be verified by checking the felicity of (18b) as an answer to a corresponding question such that the question specifies which expression will get a topical interpretation in its response. This is one of the diagnostic tests for checking the topichood of a constituent. Such a congruence in question-answer pairs is shown in examples (19) and (20). In the conversational frame of example (19), only the referents introduced in speaker A's question are eligible candidates for the (aboutness) topic in its response. Speaker C's assertion is a felicitous response since its topic is 'I' (already mentioned in discourse). Speaker B's response, which varies

[23] Notable exceptions are Ellen Prince, Gregory Ward and Ilkyu Kim.

from Speaker C's response only one factor —moving 'him' to clause initial position, is infelicitous in this context.

(19) Speaker A: Whom did you meet yesterday?
Speaker B: # Him, I met yesterday.
Speaker C: I met him yesterday.

In contrast, example (20) shows a felicitous response with a topicalized 'him'. The difference is that the question here explicitly specifies the set of aboutness topics {you, my brother} available for its response, which included the referent of topicalized 'him'. Thus, this question-answer pair is congruent and it exhibits that topicalization operates only on potential topical constituents.

(20) Speaker A: I heard that you met my brother and sister. When did you meet my brother?
Speaker B: Him, I met yesterday.

Presuming topics to be necessarily discourse-old or given, Erteschik-Shir (1997, 2007) analyses the Danish topicalization pattern as a diagnostic test to exhibit what class of nominal expressions qualify as topics in a language. This topical set includes pronouns (example (21a)), definite NPs (example (21b)), generics (example (21c)) and a contrastive NP (example (21d)). The data for Danish is from Erteschik-Shir (2007:8), the 'intended' version of English translations are mine.

(21) a. hende mødte jeg i går.
her met I yesterday
'I met her yesterday.'
'Her, I met yesterday.' (intended)

(21) b. pigen mødte jeg i går.
the girl met I yesterday
'I met the girl yesterday.'
'The girl, I met yesterday.' (intended)

(21) c. blomster ser man om foråret.
flowers sees one in the spring
'One sees flowers in the spring.'
'Flowers, one sees in the spring.' (intended)

(21) d. pigen mødte jeg i går. drengen mødte
the girl met I yesterday, the boy met
jeg først i dag.
I only today
'I met the girl yesterday, I met the boy only today.'
'The girl, I met yesterday. The boy, I met only today.' (intended)

Topicalizing a non-specific, indefinite NP leads to ungrammaticality (example (22a)), because the referent of such an NP is not *given* in the discourse and hence, cannot be made topical. However, a specific indefinite NP (an indefinite modified by a relative clause; thus receiving some contextual specification) can topicalize in Danish. This is evidenced in example (22b).

(22) a. * en pige mødte jeg i går
a girl met I yesterday
'I met a girl yesterday.'
'A girl, I met yesterday.' (intended)

(22) b. en pige som jeg mødte i går
a girl that I met yesterday
gav jeg en god bog
gave I a good book
'I gave a good book to a girl that I met yesterday.'
'A girl that I met yesterday, I gave a good book (to).' (intended)

Like Danish, Catalan also grammatically encodes the topic role of a constituent using syntactic strategies. Vallduví (1992) posits two types of old-information topics for Catalan — namely, *links* and *tails*. Both these topic types involve syntactic movement from their base position in a sentence. Links are the left-detached topical constituents and tails are the right-detached topical constituents permitted by this language.[24] The proposed structural configuration for such left or right detached XP topics are:[25]

(i) Link structure: $[_{IP}\ XP_1\ [_{IP}\ldots cl_1\ldots t_1\ldots]]$

(ii) Tail structure: $[_{IP}\ [_{IP}\ldots cl_1\ldots t_1\ldots]\ XP_1]$

[24] He considers left- or right-detachment as a left- or right-adjunction to IP (Vallduvi 1995: 127). A similar adjunction-to –IP analysis for topicalization was proposed for English (Baltin 1982) and Italian (Rochemont 1989).

[25] Abbreviations used: t-trace; cl-clitic (Catalan leaves a co-referential clitic pronoun in the main clause which is co-indexed with the detached element)

Vallduví compares Links to 'address pointers' in a file system in which new information is listed under the address specified by the Link. Since Link is a command to 'go to' an existing address, they only appear when there is a change of 'address' (Erteschik-Shir 2007:11). Thus, a sentence topic of a previous sentence cannot have its topical status continued in the next sentence (referred to as a continued topic) and be topicalized as a Link for that sentence. Only a switch/shifted topic, which is a 'new' topic as compared to a previous sentence topic, can become a Link.[26] Example (23) exhibits a Link construction. The shifted topic *aixo* is the left-detached Link in this sentence (data from Vallduví 1992).[27] It moves from its base position leaving behind a co-indexed trace *t*.

(23) Context: [after mentioning something nasty that the hearer had done to the speaker a long time ago]

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Aixo$_1$ | ho$_1$ | tine | clavat | $t_1$ | al | fons | del | COR |
| this | obj | have | pstppl-stick | | at.the | depth | of.the | heart |

Lit: 'This I have it stuck deep in my heart.'
'This I won't forget how it hurt.'

In contrast to Links, the constituents that are right-detached and thus become tails, must have an antecedent in the previous sentence, either as a topic (so it becomes a continued topic) or a focus (a type of 'focus chaining' - mentioned in the beginning of this section). Tails cannot be inferred from a hypertheme mentioned in the preceding sentence; they must have an overt antecedent. Example (24) exhibits a Tail construction from Villalba (1998). The question in example (24) contains the antecedent *els llibres* which gets right-detached in the response utterance in (24 A). This Tail element is italicized in the example.

| | | | | | |
|---|---|---|---|---|---|
| (24) Q. | on | va | posar | els | llibres? |
| | where | PAST-3 | put | the | books |

'Where did (s)he put the books?'

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (24) A. | em | sembla | que | els | va | posar | al | despatx, | *els llibres* |
| | to-me | seems | that | them-MASC | PAST-3 | put | in-the | study | the books |

'It seems to me that (s)he put the books in the study.'

[26] Aissen (1992) claims different syntactic positions for continued and shift topics. Switch topics are CP-external whereas continued topics are CP-internal. His analysis is based on data from Mayan languages.

[27] The link element is underlined. Gloss: obj-object; pstppl – past participle.

Example (25) exhibits a Link ‘the boss’ and a tail ‘broccoli’ in the same sentence in Catalan (data from Vallduví 1995:133). The focused constituent (verb *odia*) remains inside the IP and is marked with capitals in the data, followed by its English equivalent.

(25) L’amo$_1$ [l$_2$ ‘ODIA t$_2$ t$_1$], el broquil$_2$
The boss [HATES] broccoli.

The Danish and Catalan topicalization data has been used by linguists to corroborate the semantic-pragmatic properties associated with *topics* by Strawson (1964). According to him, topics have three core features:

(i) They are what a statement is *about.*

(ii) They are used to invoke “knowledge in the possession of an audience” and

(iii)“The statement is assessed as putative information about its topic” (Strawson 1964: 97-8).

Links provide evidence for the first proposed feature (aboutness) because they have been conceptualized as ‘addresses’ where new information gets entered. This new information has to be ‘about’ the referent of that address to be entered there. The distributional facts of XPs that topicalize in Danish provide evidence for topics to be ‘old/given’ information. This argument correlates to the second feature of topics proposed by Strawson since only discourse given/old constituents exist in the knowledge possessed by the interlocutors and is capable of being invoked.

The third feature establishes topic to be the pivot on which truth value of the sentence is calculated. A sentence topic with no referent in the context leads to a truth value gap. The reason for this is that such a sentence whose denotation is not complete cannot be evaluated for either a true or a false value. Such a truth value gap is translated as ‘undefined’ (represented as value ‘#’ instead of ‘1’ for true and ‘0’ for false) in the formal semantic language. Example (26) exhibits a situation where the topical constituent (‘the King of France’) in response to a question like example (26) cannot be evaluated since in our present world knowledge no such entity exists. The semantic truth value for example (26 A) is provided in (27).

(26) Q. How tall is the King of France?
A. The King of France is six feet tall.

(27) $[|\ (23)A\ |]^{M,w,g} = \#$ (where w=the possible world that we live in)

This correlation between the truth value of a sentence and the sentence topic has also been developed by Reinhart (1981), who employs a formal notion of 'context set' (adopted from by Stalnaker (1978)). She proposes that "the context set of a given discourse at a given point is the set of propositions which we accept to be true at this point". In this theory, a new assertion, if true, adds a proposition to the context set. Sentence topics have the role of classifying these new propositions in the context set. As all sentences must have a truth value, consequently all sentences must have topics. However, this dependence of truth value assignment on sentence topics is not accepted by all. Vallduví questions topics (Links) being the locus of truth value assignment as there are even linkless sentences in Catalan (see example (22 A) above). Since all sentences have a semantic truth value in a given universe of discourse, then sentence topic can't be the universal locus for truth value assessment. This argument gets further support from theories like Kuroda (1972), Ladusaw (1994), Lambrecht (2000) etc. which establish a category of IS constructions that do not involve topics. They posit that IS has three types of sentence constructions:

Type 1: Thetic sentence: A sentence where the whole sentence is in focus and there is no overt topic constituent. Example (28) illustrates a thetic construction as a response to question.

(28) Q. What happened?
A. [She cleaned her room!]$_{FOCUS}$

Type 2: Categorical sentence: They have focus on the predicate and do have an overt topic. Example (29) exhibits the same sentence (thetic example (28 A)) packaged differently as a categorical construction in (29 A) in response to a question (29 Q).

(29) Q. What did she do?
A. [She]$_{TOPIC}$ [cleaned her room]$_{FOCUS}$

Type 3: Argument Focus sentence: They have focus on an argument (usually contrastive focus). This type has also been analysed as categorical sentence with a contrastive topic subject. Example (30) exhibits argument focus on 'my uncle.'

(30) Q. I heard that your aunt fell from the balcony?
A. [My uncle] $_{\text{FOCUS}}$ fell from the balcony.

Gundel (1974) rejects the claim that thetic sentences have no topic. She argues for their topic being "the particular situation (time and place) about which it is asserted". This expands the criterial property of topics to not only denote an entity-type constituent, whose denotation is a referent in the discourse, but also to include a frame-type definition of topic, that does not have a real world referent associated with it. Erteschik-Shir (1997) coined the term "stage topic" for such implicit topics which refer to the spatio-temporal setting of a sentence. Example (31) exhibits an implicit stage topic in an English sentence.

(31) There's a stranger outside the door.

This spatio-temporal frame of the sentence becomes an overt stage topic for the sentence in example (32).

(32) Outside the door, there's a stranger.

This claim of the existence of stage topics in natural language is supported by evidence from Bantu languages. They topicalize locative and temporal adverbials just like argument topicalization (Bresnan and Kanerva 1989).

Besides syntax, morphology is also employed by languages to grammatically encode topicality. Evidence is adduced from languages like Japanese and Korean that resort to morphological markers to signal the information structural status of their constituents. Kuno (1972) proposed an analysis for Japanese particles *wa* and *ga*. He claims that *wa* is the topic particle and *ga* is the focus particle for this language. Example (33) and (34) exhibit a *wa*-marked topic constituent, data from Kuno (1973) and Kuno (1972: 271) respectively.

(33) John-wa nihon-ni ikitakagatteru
John-TOP Japan-LOC want-to-go
'John wants to go to Japan.'

(34) John wa watakusi no tomoati desu
John I 's friend is
'John is my friend.'

In his analysis "[wa] marks either the theme or the contrasted element of the sentence. The theme must be either anaphoric (previously mentioned) or generic, while there is no such constraint for the contrasted element" (Kuno 1972: 270). A semantic analysis for this particle *wa* was proposed by Miyagawa (1987), based on a property of set-anaphoricity inherent to such particles. His analysis is already presented in section 1.2.5 in chapter 1. Data from Miyagawa (1987: 186) exhibiting thematic topics and contrastive topics in Japanese, presented in section 1.2.5 as examples (19)-(21), is repeated here as example (35)-(37). A thematic topic interpretation is yielded by examples (35) and (36) while a contrastive topic interpretation is yielded in example (37).

(35) dʒon wa hon o jonda
John TP book DO read
'As for John, he read a book.'

(36) kudʒira wa hoŋgju-dobutsu desu
whales TP mammals COP
'Whales are mammals.'

(37) ame wa futeimasu ga, juki wa futeimasen
rain TP falling but snow not-falling
'It's raining, but it isn't snowing.'

Similarly, Korean particle *(n)un* has been analysed as a morphological topic marker for the language (see Bak (1977, 1984), Choe (1995) etc.). Examples (38) and (39) exhibit a *nun*-marked sentence topic; Korean data from Bak (1984) and Choe (1995) respectively.

(38) chelswu-nun yenghi-lul salanghanta
Chelswu-TOP Yenghi-OBJ love
'Chelswu is in love with Yenghi.'

(39) Chelswu-nun CA- N- TA
Chelswu-TOP sleep PRES DEC
'Chelswu sleeps.'
'Speaking about Chelswu, he sleeps.'

This particle is also analyzed as a contrastive topic marker when intonational prominence is attached to it, as in example (40 A) from Büring (2016: 69).[28]

(40) Q. Who did what?
A. [Joe-nun]$_{CT}$ ca -ko SUE- nun nol- assta
Joe CT sleep and Sue CT play PAST
'Joe slept and Sue played.'

A third strategy employed by languages to mark topics is by manipulation of the phonological-prosody associated with a sentence. In cases where the binary division of sentence into topic-comment structure co-incides with a given versus new sentence constituent bifurcation, the topic can be prosodically unmarked i.e., instantiated with a 'low or middle pitch intonation contours' (Molnár 1998). Example (41), from É. Kiss (1987), exhibits a prosodically unmarked topic subject 'the committee' and a nuclear accent marked predicate in the response (41 A) to a question (41 Q). [29]

(41) Q. hogy döntött a bizottság a tervezetek ügyében?
how decided the committee the plans concerning
'How did the committee decide concerning the plans?'

(41) A. a bizottság 'elfogadata a javaslatot
the committee accepted the proposal-ACC
'The committee accepted the proposal.'

Topics can be prosodically marked too. In cases where topical constituents exhibit such a prosodic pattern, they trigger additional semantic and pragmatic effects. Cross-linguistic studies have identified a specific prosodic pattern— called as a "B-accent", a "fall-rise", an "I-Kontur"— that a topical constituent usually gets marked with in a language. This accent involves a bitonal L*+H or a tritonal L*+H L (low-high (low)) pattern.[30] Jackendoff (1972) proposed the label of "B-accent" to refer to a rise-fall-rise contour that mark contrastive topics. This "B-accent" is juxtaposed to an "A-

[28] This is contrary to Hetland (2007), who argues that *nun* does not require intonational prominence to mark a CT. The function of *nun* particle is to mark out alternatives required for a contrast.

[29] The nuclear accent is signaled with an apostrophe ' mark preceding the predicate in the example.

[30] Using TBA convention, L-stands for low tone, H- high tone, L*- rising intonation

accent" – a plain fall intonation that marks the sentence focus (Bolinger 1965, Jackendoff 1972). Example (42) exhibits a contrastive topic constituent 'Fred' marked with a B-accent and an A-accent accompanying the focal constituent 'beans'.

(42) Speaker A: Well, what about Fred, what did he eat?
Speaker B: [Fred]$_{\text{B-Accent}}$ ate the [beans]$_{\text{A-Accent}}$.

In Büring's seminal analysis (2016), contrastive topics (CTs) are marked by a B-accent too and this L*+H tonal contour has a specific semantic effect— scope inversion. Example (43), from Büring (2016), involves two scope taking quantifiers – a negation 'not' and a universal quantifier 'all'. In the surface scope in the unmarked case, universal quantifier out-scopes the negation. When the universal quantifier is marked as CT by associating a B-accent with it, a sentence interpretation with inverse scope order is obtained. In this scope inversion reading, negation out-scopes the universal quantifier.

(43)

| | | | | |
|---|---|---|---|---|
| ALLE$_{\text{CT}}$ | politiker | sind | NICHT | korrupt |
| all | politicians | are | not | corrupt |

'Not all politicians are corrupt.'

Büring (2016: 250) extends the analysis proposed in Alternative Semantics, for deriving focus-induced alternatives, to derive the alternatives that become relevant for CT marking of a constituent. Example (44) exhibits the formalisation proposed for 'F-alternatives' when a constituent 'me' gets focus-marked in a sentence.

(44) She wants to kick [me]$_{\text{FOC}}$ out.
F-Alternatives: the set of propositions like 'she wants to kick x out,' for some individual x

Example (45) exhibits a minimal pair with example (44), the difference being that the subject 'she' is CT-marked in this case. The set of CT-alternatives that gets invoked by this sentence is formalized below.

(45) [She]$_{\text{CT}}$ wants to kick [me]$_{\text{FOC}}$ out.
CT-Alternatives: the set of question meanings like 'Who does y want to kick out?' for some individual y

Büring (1997) posits a notion of "Residual Topic" that is created in contexts of CT-marking of a constituent. Residual topic is not a topic per se but a topic-related inference. According to this, CT-marking of a constituent implies but does not exclude the other relevant alternatives to that constituent.

Molnár (1998) provides evidence that weakens this claim of direct correspondence between prosodic marking of a topic constituent and the specific discourse-semantic effects attained. They exhibit that data, like example (46), where an element that cannot be topical (negation particle) gets CT-marked by a fall-rise or "I-Kontur" (Jacobs 1997). Thus, this prosodic marking yields discourse-semantic effects but it is not obligatorily encoding topicality.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (46) | ich | habe | $\text{NICHT}_{CT}$ | getrunken, | weil | ich | $\text{TRAURIG}_{F}$ | bin |
| | I | have | not | drunk | because | I | sad | am |

'I didn't drink because I am sad.'

Thus, languages can use prosody to grammatically encode IS notion. A focus constituent always gets prosodic prominence or nuclear stress in a sentence, but prosodic marking of topicality is relatively less clear. Topics can remain unaccented a shown in example (41A) but they can also carry a specific prosodic pattern that has certain discourse-semantic effects (as in example (43)).

This section has exhibited the various syntactic, morphological and prosodic strategies employed by various languages cross-linguistically to grammatically encode the notion of topicality. Topics not only have a grammar-internal effect but also a grammar-external effect. Topics tell how a speaker's mental knowledge storehouse is accessed and modified during the course of a communicative interaction between natural language speakers. Topics have both a CG content and a CG management dimension. However, no uniform treatment of *topic* has been possible in IS studies, owing to the vast diversity of definitions and linguistic phenomena postulated for giving an account of this IS category.

### *2.2.3 Diagnostic Tests for Topichood*

The methodological tool of diagnostic tests becomes crucial for investigating a linguistic phenomenon linked to a concept as difficult to define as *topics*. Fortunately, a battery of tests has been proposed to identify the topical constituent in a sentence in IS literature. Gundel (1974, 1985) proposed the following tests: *What about X* test, *Speaking of X* test, *As for X* test. Reinhart (1981) has proposed *Say about X that S* test, *About/of X* test, *Left-dislocation* test. Each of these tests checks the topichood status of a constituent occupying the position X. As one strategy for testing, a question-answer pair is constructed such that the question employs a template provided by the test type and the answer is assessed for keeping its topic fixed and thus, maintaining discourse cohesion. A second strategy involves using these test templates as linkers for continuing the discourse and introducing a proposition to be added to it. The felicity status of the sentence under evaluation indicates the result of these diagnostic tests.

Examples (47) and (48) exhibit a *What about X* test performed on English data to diagnose the topical argument of the sentence. This data is from Roberts (2011).[31] Example (47 Q) expects an answer whose 'aboutness' topic is the subject argument 'Mary'. Sentence in example (47 A) fulfils this requirement and makes 'Mary' the constituent about which this sentence is structured. Thus, (47 A) is a felicitous response to (47 Q).

(47) Q. What about Mary? What did she give to Harry?
A. *Mary* gave [a shirt]$_{\text{Rheme}}$ to Harry.

Example (48 Q) imposes the topical status on the indirect object 'Harry'. A response to this question shout be structured about the referent 'Harry'. Felicity of (48 A) is interpreted as a positive result for this diagnostic test.

(48) Q. What about Harry? What did Mary give to him?
A. To Harry Mary gave [a shirt]$_{\text{Rheme.}}$

[31] The stylistic convention followed in the original data (Roberts 2011) is maintained in this example. Topics are italicized and other IS notions like rheme are subscripted with the constituent that has that discourse role.

These tests presume an 'aboutness' property ascribed to sentence topics. This assumption is made explicit in the names of *What about X* test and *About/of X* test. By fixing a constituent as 'about' which a (preceding) question is asked, these tests directly identify the aboutness sentence topics in their responses. However, these aboutness topic tests are not identical. The *About X* test implies topic chaining i.e., topic of the sentence under observation should be a continued topic. The *What about X* test implies a shift in topical status between discourse referents. This switch/shifted topic is also implied for the *Speaking of X* test. For this test, the discourse referent should be relatively salient to become a shifted topic, a condition not required for *What about X* test that can pick out even non-salient entities in the discourse (Roberts 2011). The *As for X* test implies a contrast between the topical element X and other salient alternatives (implicit or explicit) in the discourse. Thus, this test is used to identify a contrastive topic in a sentence. The *What about X* test may also identify contrastive topics of a sentence (if contrastive topics are given a Büring-style disjunctive questions interpretation).

From the above observations, Roberts (2011) has claimed that each test has a different pragmatic condition under which it can be felicitously used.[32] In the case of the aboutness-based diagnostic tests specified above, they might yield non-identical results within the same discourse context. The pragmatic requirements of one test may be fulfilled in one context, while the pragmatic requirements of the other may not. Such a situation is exhibited in example (49), data from Ward (1985:73).

(49) Context: (a report in Philadelphia Inquirer, p. 8-C, 9/1/83)
Then Tom Cruise went to work for Francis Ford Coppola, on this spring's semi-successful film version of "The Outsiders". *Coppola*, he found to be "just like one of the guys. And he totally trusted me. He let me go anywhere I wanted to go with the character…"

a. About Coppola, he said that he found him to be…
b. #What about Coppola? He found him to be…
c. #As for Coppola, he found him to be…
d. #Speaking of Coppola, he found him to be…

[32] A strong critique of these tests can be seen in Büring (2016: 81) who states that "none of the topic tests seem sufficient or necessary to identify topics, nor is it clear that they actually test for the same thing".

The context in example (49) is a news report about a certain actor and a certain movie director. The context ends with the actor talking about that director (probably to a reporter- not relevant for us). A continuation of this discourse, using various topic tests, is analysed in examples (49a)-(49d). In this situation, example (49a) is pragmatically felicitous and fits into the discourse-context. This is so because (49a) requires 'Coppola' to be a continued topic from the preceding sentence. This requirement is satisfied in the current context. Discourse continuations in examples (49b) and (49d) is infelicitous because their test frames imply a shift in topical status. However, the preceding sentence was also about the same topical referent ('he'~Coppola). Thus, no shifting of topicality happens and using test templates that imply these topic types leads to infelicity. The test template of example (49c) implies the presence of an alternative entity in the preceding discourse from which 'Coppola' is contrasted and made topical. However, no such alternative in available in the discourse and Coppola is not a contrastive topic. This explains the infelicity of example (49c). Thus, the result of these tests is not just sensitive to which constituent gets fixed with a topical interpretation but also to the larger discourse-pragmatic considerations of the context in which the sentence is anchored.

In conclusion, this section has provided the concepts and diagnostic machinery requisite for a formal investigation in *topic*-related linguistic phenomenon in any language.

### *2.2.4. Some Formal Models of Topics*

This section discusses two theoretical frameworks that are topic-oriented (as compared to IS-oriented models introduced in section 2.1.4) and that propose analyses that maintain a hierarchically distinct typology of various type of topics within the clausal spine. A third framework that is not strictly IS oriented but that provides an alternative proposal to incorporate IS effects (including aboutness topic interpretation) within the functional spine is introduced at the end of this section.

*Frascarelli and Hinterhözl (2007)*

Working within the cartographic framework, they posit a typology of three types of topics that occupy distinct structural positions within the C-domain. These are:

(i) Familiar Topic (F-topic): this is what the sentence is about but at the present stage of discourse, it is backgrounded. The topic constituent moves to the C-domain from a TP-internal position.

(ii) Contrastive Topic (C-topic): This is a topical constituent that gets contrasted against a set of discourse-given entities. This constituent moves to a position higher than F-topic in the C-domain.

(iii) Shift Topic (S-topic): This constituent introduces a new topic or re-introduces an old topic at the current stage of discourse. Frascarelli (2007) has also called this topic type as Aboutness-Shift Topic. This topic constituent is either moved from inside TP or it is externally merged in the C-domain. The position of S-topics is higher than C-topics.

Evidence is provided from Italian and German to posit this hierarchy of different topic types. Each topic is phonologically marked by a distinct intonation curve. For Italian, S-topics are marked by a low-high tone; C-topics by a high tone; and F-topics by a low, flat tone. A clause can have co-occurrence of more than one topic type but the relative order of topics has to remain the same. This hierarchical ordering is shown in figure 2.6 below.

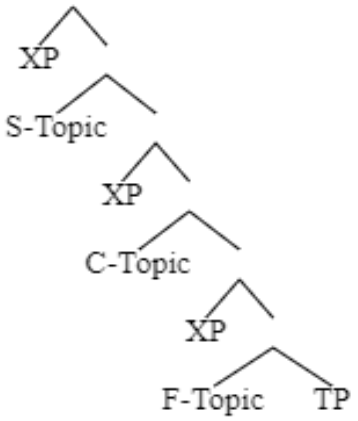



Figure 2.6. Topics in the Italian Left Periphery

In their model, it is possible to have several F-topics in the left periphery of the clause but only one S-topic or C-topic. A root restriction applies on S-topic and C-topic because of which they can occur only in the root clause and not in an embedded clause. This restriction does not apply to F-topics, which can freely occur in both root and non-root clauses.

*Holmberg (2020)*[33]

Holmberg adopts the topic hierarchy posited by Frascarelli and Hinterhözl (2007) and the δ-feature (discourse feature) proposed by Miyagawa (2017) to provide a syntactic analysis of topic-marking of constituents in North Hail Arabic (NHA). He posits topic (and focus) to be narrow syntactic categories that are realised as syntactic features (δ-features, like φ-features) on topic functional heads and the constituents (DP or PP) that get marked as topical during the course of derivation.[34] The topic head has a δ-feature valued as either F-topic, C-topic or S-topic. A constituent is lexically specified for an unvalued δ-feature. It can either get valued by an Agree mechanism (between a topic functional head and the lexical item) or it can get valued by default rules. Topic particles are overt instantiations of the abstract topic functional heads. In their theory, these particles are externally merged in C-domain in the hierarchical order of Frascarelli and Hinterhözl (2007). Languages can have two types of topic particles – agreeing v/s non-agreeing topic particles. Evidence is provided for NHA to be a language that has topic particles that either agree with the subject or the object argument in the main clause.

The agreeing type of topic head has a set of valued δ-feature but unvalued φ-feature. It probes for a matching XP that has valued features but unvalued δ-feature. It enters into an Agree relation with the goal which results in valuation of [u-δ] feature of goal XP and [u-φ] feature of the topic head. This goal (subject or object NP) moves into the specifier position of this topic head and gets valued there as [δ: F-topic/C-topic/S-

[33] This section is based on a seminar presentation by the author hosted online by Chinese University of Hong Kong on 29.09.2020. The presentation was based on a joint work between Anders Holmberg and Murdhy Alshamari on a Saudi Arabian 'dialect' of Arabic.

[34] Contra Chomsky, Gallego & Ott (2019) who hold the view that topic and focus are not features inherent in lexical items.

topic]. The non-agreeing type of topic head has a valued δ-feature but no unvalued φ-feature. The XP that has a [u-δ] feature moves into the specifier position of the topic head to get its discourse feature valued.

Some of the default rules posited by this framework are: an XP in predicate phrase gets it [u-δ] feature assigned a value of [information-focus] by default; in the TP domain it is assigned [F-top] by default and a [u-δ] not probed by but still in the scope of a C-top head, gets assigned [F-top] by default.

Thus, in his model, a constituent becomes a particular type of Topic via the mechanism of feature valuation during the course of syntactic derivation.

*Wiltschko (2020)*[35]

An alternate model for incorporating Information Structure in the architecture of grammar is proposed by Wiltschko (2014) in her Interactional Spine Hypothesis (ISH). According to ISH, human language faculty uses the same formal architecture to account for the two components of Universal Grammar — *p(ropositional)* language and *i(nteractional)* language.[36] P-language is the dimension of language involved in constructing truth-conditional propositions. I-language is the dimension of language involved in situating an utterance move made by a speaker in a communicational setting by managing the dynamics of common ground as well as regulating turn-taking sequence amongst the interlocutors. This layer involves discourse markers and intonations that do not contribute truth-conditional meaning to a sentence.

A universal spine has an i-language component built above a p-language component. The functions performed by the p-language part of the universal spine are: linking, anchoring, point-of-view and classification of a proposition. After this p-structure is

[35] This section is based on a seminar presentation by the author hosted (online) by Chinese University of Hong Kong on 15.09.2020. The ideas discussed here were later incorporated in a book by the author in 2021, to be further elaborated upon in chapter 5.

[36] A syntactic configuration that involves a functional head selecting a contextually-restricted argument as a complement, with which it enters into a feature-valuation style Agree system.

derived, 'Unit of Language' (UoL) has a grounding domain (situates the utterance in the common ground) and a responding domain (specifies whether the current utterance falls in the response set of its preceding utterance). Grounding and responding functions fall in the i-language component of the universal spine. In this model, IS is spread across both i-language and p-language. In addition, the interpretations of information structural notions (that have no semantic effect, like (aboutness) sentence topic) are obtained by manipulating the i-language domain configurations of a sentence. The clausal spine in this framework is represented in figure 2.7.[37]

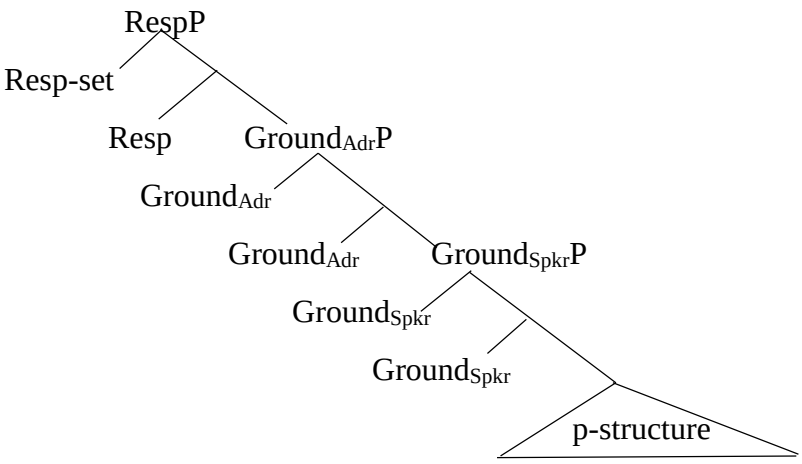



Figure 2.7. Interactional Spine Hypothesis Structure

The next section shifts the attention back to the Hindi enclitic particle *-to* under investigation and uses the tools provided in this section to formulate some preliminary issues with an analysis of *-to* as a topic marker (*a la* Kidwai 2000).

## 2.3 Hindi IS and particle *-to*

In the IS literature discussed above, a directionality of research can be deduced. An inquiry starts from basing itself in an understanding of IS and IS notions and then analyses grammatical encoding of these IS notions in linguistic data. Formal accounts

[37] Abbreviations: RespP- response phrase; Resp-set – response set; $Ground_{Adr}$P- Addressee's ground phrase; $Ground_{Spkr}$P- speaker's ground phrase; p-structure – propositional structure.

are proposed to account for linguistic devices (a syntactic movement, a morphological particle or a prosodic pattern) that are presumed to be the linguistic realizations of an IS notion. This line of thought is also followed by Kidwai (2000, 2004) who claims that the enclitic particle *-to* is a topic particle — a morphological strategy employed by the language to overtly mark its topic in a sentence. This claim implicitly assumes particle *-to* to inherently encode topicality. This section showcases issues with this analysis of *-to* as a topic marker and provides evidence to dissociate topicality as an inherent property from the particle *-to*. To accomplish this, a reverse line of inquiry that starts from a linguistic datum and moves towards its information structural interpretations is motivated at the end of this chapter.

### *2.3.1 Hindi -to is NOT Japanese -wa or Korean -(n)un*

The only existing generative account of enclitic particle *-to* in Hindi (Kidwai 2000, see section 1.2.5) provides a formal treatment of it on lines similar to Japanese topic particle *-wa* (which is claimed to be functionally similar to Korean topic particle *-(n)un*). A parallel is drawn between a thematic and a contrastive interpretation obtained for *-wa* marked XPs (Miyagawa 1987) in Japanese and *-to* marked topic XPs in Hindi.[38]

However, Hindi *-to* is not the same linguistic device as Japanese *-wa* or Korean *-(n)un*. In terms of empirical distribution, these languages allow their sentences to have more than one occurrence of particles *-wa/-nun*. Two NPs can be marked by these 'topic' particles within the same clause. Such a construction is called a "double subject" construction (Kiss 1998). This is exhibited for Korean in example (50) that contains two *-(n)un*-marked NP arguments (data from Roberts 2011).

(50) Mary-nun John-un chohaha-n-ta
Mary-TOP John-TOP like-PRES-DECL
'As for Mary, she likes John (but not others).'

[38] Kidwai (2000) assumes a Reinhart-style 'aboutness' definition of a sentence topic.

Marking more than one XP with particle *-to* in a clause in Hindi leads to ungrammaticality. Example (51) shows a Hindi sentence with two *to*-marked XP arguments.

(51) * merɪ=ko=to dʒon=to pəsənd hɛ
Mary=DAT=TOP John=TOP like AUX.PRS.3SG
'Mary likes John.'

A second difference between these languages is the range and locality of distribution of these particles. While Japanese and Korean marks NPs, predicates and adverbial constituents with *-(n)un/-wa* (Heycock 2007), Hindi permits a much wider domain of constituents that can get marked with *-to* (see section 1.4 of chapter 1). For a quantified NP, while Japanese and Korean can mark the noun with these particles, Hindi permits a more local marking of only the quantifier too, besides a permissible structure of full quantified NP being marked with *-to* (see section 1.4.2 for a discussion of these distributional facts).

Example (52) exhibits a topical Japanese quantified NP 'every room' where the noun *heya* 'room' is *-wa*-marked (example from Portner and Yabushita 2001).

(52) heya-wa subete sansetto biichi ni menshite-imasu
room(S)-TOP every sunset beach LOC facing-be
'Every room faces sunset beach.'

Example (53) exhibits an existentially quantified NP in Hindi with *-to* marking the existential quantifier *koi* (someone) and not the full existentially quantified NP (which is shown in example (54).

(53) koi=to bəʈʃʈʃa gʰər gəja
each=TOP child home go.PFV.MSG
'Some child went home.'

(54) koi bəʈʃʈʃa=to gʰər gəja
each child=TOP home go.PFV.MSG
'Some child went home.'

A third difference between Japanese *–wa* and Hindi *–to* particles is that Japanese *–wa* can mark a constituent in a constituent question but Hindi *–to* cannot. Example (55)

shows –wa marking of 'John' in a constituent question in Japanese (data from Heycock 2008). The same question in Hindi (56) is ungrammatical when enclitic particle –*to* marks 'John' in it.

(55) John wa nani o yatta no?
John WA what ACC did Q
'What did John do?'

(56) * dʒon=ne=to kja kija
John=ERG=TOP what do.PFV.MSG
'What did John do?'

Thus, basing an analysis of Hindi particle *-to* (as a topic marker for the language) on account of only its semantic interpretational similarity with the traditionally accepted topic markers *-wa/-(n)un* is not justified, especially when they exhibit dissimilar syntactic properties.

### *2.3.2 Investigating Hindi Topics with Diagnostic Tests*

This section uses the diagnostic tests proposed for testing the topicality of some entity X in a sentence (see section 2.2.2) to investigate grammatical reflex of IS notion of topic for Hindi. I adopt Reinhart's definition of sentence topic to be what the sentence is 'about'. For aboutness topic (AT) interpretation, I employ *What about X* test. Contrastive topics are topics that correspond to questions of sub-enquiry (Büring 2005) or one in which the hearer answers a question that is different from the one being asked by the speaker (Sener 2010). The only diagnostic test that implies a contrast with entities specified in the preceding discourse is *As for X* test, which I employ to test contrastive topic (CT) interpretation.

These tests are applied on minimal pairs of sentences that vary on the parameter of marking X with *-to* particle. In each example, first the unmarked XPs are tested and then *to*-marked XPs.

1. *What about X* test[39]

Example (57 Q) establishes a context that substitutes variable X in the test template for *khaana* 'food' and accords an AT interpretation to it. Example (57 A.1) and (57 A.2) are the two constructions that are felicitous responses to this question. Thus, both the unmarked XP in (57 A.1) and the marked XP in (57 A.2) can be AT topics in Hindi.

(57) Context: Two people are planning a house party for their friends. They are having a discussion about what all things are needed for this party.

Q. kʰane=ka kja
food.OBL=GEN what
'What about food?'

A1. kʰana bahər=se lenge
food outside=ABL take.FUT.3SG
'(We) will get food from outside.'

A2. kʰana=to bahər=se lenge
food=TOP outside=ABL take.FUT.3SG
'(We) will get food from outside.'

2. *As for X* test

Example (58) shows a discourse context that requires 'Rahul' to be made the CT for the sentence that continues the discourse. Since both unmarked XP (in example (58a)) and *to*-marked XP (in example (58b)) felicitously maintain discourse cohesion, both of them are CT topics in Hindi

(58) Context: A mother is telling about the sleeping habits of her children, Ravi and Rahul, to a friend. She says "Ravi falls asleep very easily…

a. ɔr rahʊl ke lije sona ek pəreʃanı hɛ
and Rahul for sleep.INF one problem AUX.PRS.3SG
'As for Rahul, sleeping is a problem.'

b. ɔr rahʊl ke lije=to sona ek pəreʃanı hɛ
and Rahul for=TOP sleep.INF one problem AUX.PRS.3SG
'As for Rahul, sleeping is a problem.'

[39] The closest translation of this test into Hindi is 'X ka kya?' which may not capture the exact intuition behind 'what about X?' test. However, since the Hindi equivalent also fixes X as the topical constituent about which the response utterance is expected, this test is applied to Hindi data.

However, there is an issue with the applicability of *As for X* test to diagnose contrastive topics for Hindi. The translation of this test into Hindi implies a switch/shifted topic rather than a contrastive topic. To avoid this ambiguity that creeps in because of the translation, I follow a diagnostic template adopted by Sener (2010), based on a *Tell me about X* test proposed by Neeleman and van de Koot (2008), to identify the AT constituent, the CT constituent and the focus constituent in a congruent question-answer pair.

Example (59) shows a context that fixes 'mummy' as the AT for speaker B's response by speaker A's utterance. The speaker B, by answering a question that is different from what speaker A has asked, makes 'papa' the CT of his utterance. The constituent that corresponds to the wh-variable in question is traditionally assumed to be the focus part in the answer. Thus, *kaafi daant* is the focal constituent of speaker B's utterance. Crucially, the constituent that gets CT interpretation can be both unmarked as well as marked with particle *-to*.

(59) Context: Two brothers are talking about how their parents received the news of one of them failing an exam in school.

Speaker A: məmmɪ=ke bare=mẽ batao
mummy=GEN about tell.IMP.MSG
'Tell me about Mummy.'
unhõ=ne kja kəha
she.HON=ERG what say.PFV.MSG
'What did she say?'

Speaker B: [məmmɪ=ka]$_{AT}$ pəta nəhɪ pər
mummy=GEN know NEG but
'I don't know about Mummy but'
[papa=ne=(to)]$_{CT}$ [kafɪ ɖãṭ]$_{F}$ sunajɪ
papa=ERG=(TOP) many scolding listen.CAUS.PFV.FSG
'Papa scolded me a lot.'

Based on the result of the above two diagnostic tests, both unmarked XPs and *to*-marked XPs are CT topics in Hindi. In addition, the first test has shown that both unmarked and marked XPs are AT topics for the language. Thus, the discourse role of topic (both AT and CT) can be performed by a *to*-marked XP as well as an unmarked XP in Hindi.

This raises a language design issue. Assuming topical interpretation to be available at CI interface in the architecture of grammar, if an unmarked XP can get a topical interpretation at CI interface (without any recourse to a morphological strategy), then claiming the particle *to* to encode that same discursive meaning yields a redundancy. For language design to be optimal and efficient, particle *-to* should have an independent motivation to be selected from the lexicon in the computation and also, *to*-marked topic XPs should have some distinction from unmarked topic XPs.

Evidence that *to*-marked XPs are not the same topical phenomenon as unmarked XPs comes from data like example (60). Using the *Tell me about* test from Sener to construct a template for identifying AT, CT and focus constituent in a sentence, example (60) fixes 'Rahul' as the AT constituent, 'Shyam' as the CT constituent and *Dosa* as the focused constituent for speaker B's response to speaker A's question. Response 1 is felicitous in the given discourse-context whereas as response 2 is infelicitous.

(60) Context: A and B are friends who are discussing about the party that happened at B's office. A only knows one of B's colleagues – Rahul. So, A asks B:

Speaker A: rahul=ke bare=mẽ batao
mummy=GEN about tell.IMP.MSG
'Tell me about Rahul.'
us=ne kja kʰaja
he=ERG what eat.PFV.MSG
'What did he eat?'

Speaker B: [rahul=ka]$_{AT}$ pəta nəhı pər
Response 1 Rahul=GEN know NEG but
'I don't know about Rahul but.'
[ʃjam=ne]$_{CT}$ [ɖosa]$_{F}$ kʰaja
Shyam=ERG dosa eat.PFV.MSG
'Shyam ate a dosa.'

Speaker B: [rahul=ka]$_{AT}$ pəta nəhı pər
Response 2 Rahul=GEN know NEG but
'I don't know about Rahul but.'
# [ʃjam=ne=to]$_{CT}$ [ɖosa]$_{F}$ kʰaja
Shyam=ERG=TOP dosa eat.PFV.MSG
'Shyam ate a dosa.'

The only difference between minimal pairs of response 1 and 2 is that the CT constituent of response 1 is unmarked whereas the CT constituent of response 2 is *to*-marked. Thus, the source of infelicity in the enclitic particle *-to*. This distinction is not expected from the result of the diagnostic tests performed in the beginning of this section, according to which both unmarked and *to*-marked XPs are AT and CT topics without anything different about *to*-marked topics or to particle *-to*. Example (60) exhibits that felicitous licensing of *-to* is constrained by certain conditions. The next step is to diagnose these licensing conditions for particle *-to*.

Reinhart's notion of 'aboutness' relation for analysing topic is not useful to diagnose the difference between non--*to*- marked topic XPs and their marked counterparts that leads to infelicity in the situation described above. A solution can be found in Krifka's conception of topic (see section 2.2.1) that makes use of discourse-sensitive notion of Common Ground (CG). CG was earlier theorized as a shared set of mutual beliefs but was later defined as a shared set of acceptances rather than beliefs (as one may accept a proposition without believing it). Presuppositions are associated with a speaker's belief about common beliefs.[40] An entity that a speaker presupposes must be a part of the Common Ground between the speaker and the hearer. In case of mismatch, the hearer has the strategy of 'presupposition accommodation' available to him to keep the conversation moving. Accommodation has been defined as the process by which something becomes Common Ground in virtue of one party recognizing that the other takes it to be CG. This factor controls the dynamics of discourse by controlling the way that CG changes in response to what happens in discourse.

In example (60), the CG established between speakers A and B has only one person that A knows that goes to B's office (as made explicit by the context). This referent is {Rahul}. Speaker A presupposes that Speaker B knows that {Rahul} is the only entity in CG between them. Speaker B's response 1 gives AT status to 'Rahul'. It also introduces a new entity to the CG — {Shyam} which gets contrastive topic interpretation. Response 2 of speaker B is problematic because it attaches particle *-to*

[40] Presupposition can be defined as – *sentence S presupposes that P if and only if S is either true or false only if it is true that P*, where S and P are propositions. This concept is discussed more in chapter 3.

to an entity ({Shyam}) that was not part of the Common Ground between the interlocutors. Infelicity of speaker B's response 2 is removed if the addressee (here speaker A) accommodates B's presupposition that the relevant set of alternatives available to answer A's question is {Rahul, Shyam} as opposed to just the singleton set {Rahul}. This strategy of presupposition accomodation is available provided {Shyam} had a prior mention in discourse, which may not be recent, which enters this entity into the larger CG. Thus, particle *-to* can mark only those entities that may be actively or passively available in CG.

To conclude, enclitic particle *–to* can only mark those entities that have been mentioned in the discourse and thus, are presupposed to be available in the common ground between the speaker and the hearer. This argument gets support from empirical data like example (61) in which *-to*-marking a specific indefinite NP (a non-d-linked constituent) leads to ungrammaticality.

(61) * merɪ ek dost=kɪ=to pɪʧʰle məhɪne ʃadɪ ho gəjɪ
my one friend=GEN=TOP last month wedding happen go.PFV.FSG
'A friend of mine got married last year.'

Uttering example (61) in an out-of-the-blue context (with no implied CG between the interlocutors) is not grammatical. Its English counterpart in example (62), however, is grammatical.

(62) [A friend of mine]$_{\text{TOPIC}}$ got married last month.

Such new topics are attested in literature (See Krifka 2008) where they function to introduce new entities in the discourse. However, pragmatic requirements of particle *-to* restrict its marking of such non-d-linked expressions.

In conclusion, this section has raised empirical and theoretical issues for analyzing *-to* as only a topic marker. This assumption of one-to-one correspondence between topicality and this particle is wrong and topicality needs to be dissociated from this this particle to understand what actually is encoded by this particle and its semantic import. The next chapter deals with this issue: the non-topic interpretation yielded by particle *-to* in certain contexts. An inquiry of this type can only be undertaken when it

starts from linguistic data and the interpretations yielded by it, and then moves towards theory to propose an analysis.

# CHAPTER 3

## NON-TOPIC INTERPRETATION AND PARTICLE *-to*

In the preceding chapter, two issues were raised for an analysis of *-to* as a topic particle; a third (and crucial) issue will be discussed in this chapter. Firstly, it was claimed that Hindi *-to* is a different linguistic device from canonical topic particles like Japanese *-wa* and Korean *-(n)un*. Although functionally similar in marking thematic and contrastive topical constituents, I have provided empirical data that exhibits different syntactic distributional properties between *-to* and *-wa*/*-(n)un* (see section 2.3.1). Since there is an underlying variation in the syntax of these morphological particles, this can be taken to indicate a (mutually) distinct machinery underlying these particles, with an overlap in the observed discourse-semantic interpretative effects. This motivates a deeper investigation of particle *-to* before locating it at the syntax-semantics-pragmatics interface.

Secondly, from a Minimalist perspective, a theoretical argument was presented that pertained to the core organizing principles of economy and efficiency for language design. If the unmarked XPs can receive the exact same topical interpretation as *to*-marked XPs at the C-I interface, then employing a separate morphological strategy (of *-to*-marking) to achieve the same interpretative effect — i.e., to signal the information structural role of *topic* for a constituent of the sentence— seems redundant and the computational design becomes inelegant. Thus, there must be some difference between *to*-marked topical XPs and unmarked topical XPs in Hindi. Section 2.3.2 exhibited a licensing constraint that is unique to *to*-marked topical XPs (wherein the referent of a *to*-marked topical XP must be obligatorily available in the Common Ground established between the interlocutors, and hence given in the discourse) and that does not hold for unmarked topical XPs. From this, it was concluded that the enclitic particle *-to* is obligatorily discourse linked since it cannot mark discourse-new entities. This property of the particle *-to* cannot be derived from an analysis of it as a

grammatical reflex of topicality.[1] Based on these two factors, I had suggested that particle *-to* cannot be analysed as just a topic marker and a linguistic unpacking of this enclitic particle *-to* is necessitated on these grounds.

A third issue, explored in this chapter, is that there exist configurations where a *to*-marked XP does not serve as the sentence topic in Hindi. Thus, a topical XP may or may not be marked by enclitic *-to* and in cases where an XP is marked by *-to*, it may or may not get a topical interpretation. This provides evidence for rejecting the claim that particle *-to* is *directly* linked to the IS notion of topic by virtue of it being (as previously assumed) a topic marker and for motivating an *indirect* relation between *-to* and topicality (where obtained).

In this chapter, I investigate such non-topic interpretations yielded by particle *-to* using this methodology: I segregate the set of XPs that this particle marks (recall its empirical distribution discussed in chapter-1) into a topical type and a non-topical type.[2] A topical type XP is an XP that satisfies the definitional criteria implicitly assumed for sentential topics by any of the prevalent frameworks that provide an analysis for it.[3] Thus, a topical type XP could be either a referential or a non-referential nominal entity (NP/DP) or a propositional unit (vP/CP) i.e. an individual, an event or a situation — that can fulfil the 'aboutness'-topic criteria. To include frame-based topics, this category of topical-type XPs is expanded to include spatio-temporal adverbs, domain-restrictive adverbs like 'healthwise' etc. (AdvP) as well as conditional clauses (CP). In the Question Under Discussion (QUD) framework, a sentential topic is not an XP in the sentence *per se* but rather the most salient or the immediate QUD at any point in discourse (where the discourse itself is modelled as a

[1] Sentence topics need not be obligatorily *given* (in the sense of discourse-old) as they can introduce new entities in a discourse too (cf. Krifka 2008, Büring 2016, Rochemont 2019 etc.) and thus, be potentially non-discourse-linked. Hindi unmarked topical constituents can be discourse-new.

[2] A note about terminology: the word 'type' here is merely a categorizing label and is not intended to mean 'semantic type' (like <e> for entities, <t> for truth-values etc.).

[3] Whether *topic* be what the remaining part of the sentence is 'about' (Reinhart 1981, Lambrecht 1994, Krifka 2008 etc.); whether *topic* be the 'frame' within which the predication holds (Chafe 1976, Jacobs 2001); whether *topic* be the salient 'Question Under Discussion' (von Fintel 1994 etc.). See section 2.2.1 for an expanded discussion of these frameworks.

partially structured set of questions to which an assertion is an answering move). Thus, the notion of QUD-based topics is not compatible with the concept of topics as *to*-marked (or unmarked) topical XPs and hence, cannot be directly incorporated into the set of topical-type XPs that I am aiming for here.[4] The typology of topics based on their discourse-semantic properties —switch/shifted topics, continued topics, contrastive topics[5] — all mark nominal NP/DPs as their topical constituent.[6] A non-topical type XP would be any XP that does not satisfy the topic templates stated above.[7] Hypothetically, a non-topic interpretation can be yielded by a *to*-marked XP either if the XP itself is of a non-topical type or if the XP is of a topical type but is still not the topic of the sentence. Hindi data provides empirical evidence for *to*-marked XPs instantiating both these types of configurations.

---

[4] Similarly for Japanese topic particle *-wa*, Portner and Yabushita (1998) have claimed that 'topic-as-an-entity' analysis is better suited for *wa*-marked phrases and thus, QUD does not have much currency in Japanese literature.

[5] By the term 'contrastive topic', I mean a topical constituent that the sentence is still 'about', but one which stands in a contrastive relation with the pre-designated topic for that sentence. This definition of CT is conceptually different from a Büring-style analysis of CTs (2003, 2016) that extends the QUD approach to d-(iscourse) trees to formalise CTs as generating a set of CT-alternatives over a set of F(ocus)-alternatives for a sentence that is CT-F marked.

[6] There are certain analyses where a prosodic C(ontrastive) T(opic) marking is found on non-nominal elements like a negation particle (example (i) below from Büring 2016) or a 'finite verb' (example (ii) from Jacobs 1997).

(i) ich habe $\text{NICHT}_{\text{CT}}$ getrunken, weil ich $\text{TRAURIG}_{\text{F}}$ bin
I have not drunk because I sad am
'I didn't drink because I am sad.'

(ii) man $\text{MUSS}_{\text{CT}}$ das buch NICHT mogen (,aber man KANN)
one must the book-acc not like but one can
'You must not like the book, but you may.'

This rise-fall-rise accent pattern (or 'B-accent'), that indicates CT-marking of a constituent, is not linked to topichood in such cases (cf. Molnár 1998) but rather to open questions indicative of contrast. Thus, NPs are only the topic types that are uncontroversially CT-marked, where the grammatical marking encodes a contrastive topic interpretation.

[7] Note: Since the diagnostic tests for topichood (see section 2.2.3) are primed for 'aboutness'-based topics, they cannot be used to diagnose frame-based or QUD-based topics. Hence, these tests cannot be employed as a tool for filtering out non-topical type XPs from across the board topical types XPs.

Example (1) provides one such instance where an AdjP *theek* 'fine' is *to*-marked and this *to*-marked XP is not the topic of the sentence.[8] The set of entities that could have been potential topics for this sentence include the spatial adverb 'outside' (as a frame-setting topic) and the subject NP 'weather' (as an 'aboutness' topic) — {*baahar, mausam*}—and neither of them are *to*-marked in (1). Thus, in this context, the notion of topicality is dissociated from the particle *-to* and this provides further evidence for pushing forward an analysis of particle *-to* that does not require it to be a topic particle.

(1) Context: Two colleagues are discussing whether to step out for lunch or not (since bad weather had been predicted in the news for that day). One person wants to head out and says to the other:

| bahər | mɔsəm | [$_{AdjP}$ ʈʰik]=to | hɛ |
|---|---|---|---|
| outside | weather | fine=TOP | AUX.PRS.3SG |

'Outside, the weather is fine.'

Based on the three issues outlined above, I claim in this chapter that Hindi enclitic particle *-to* is not a topic marker itself but is rather a linguistic device that can signal either a topical or a non-topical interpretation for the entity that it marks.[9]

Similar to Kim's (2013, 2015) analysis for Korean *-(n)un,* I propose that Hindi *-to* is a salience marker. Salience has been defined by Clamons *et al* as "the cognitive prominence of the referent of an element in a discourse relative to all other elements in the discourse" (Clamons *et al* 1993:520) and by Chiarcos as "a graded notion that expresses the availability of entities and/or the degree of attention they are assigned" (Chiarcos 2009:134). A salient entity is in the focus of attention for the discourse participants. Between the two types of salience – inherent or imposed – I claim that particle *-to* encodes the latter i.e., imposed salience. This meaning of particle *-to* is not

[8] In Hindi, *theek hona* is an adjectival predicate ('to be fine') and in example (1), particle *-to* marks the adjective part of this predicate construction.

[9] I retain the topic marker gloss TOP for particle *–to* throughout this chapter until section 3.2.4 where I conclude that this particle is not a topic marker *per se* since it can yield a non-topical pure contrastive interpretation too. From that juncture onwards, I gloss this particle as TO through the length of the thesis.

truth-conditional in nature but rather *use*-conditional (used in the sense intended by frameworks that posit multiple dimensions of meaning like Potts (2003, 2007) etc.). This particle has an *instructional* meaning (cf. Gutzmann 2013) in the sense that it 'instructs' the hearer to increase the salience of the constituent that it marks. This imposed (discourse-) salience interacts with syntactic, semantic and pragmatic factors to yield different interpretative effects for a 'to-marked XP.

This proposal that instructing the hearer to increase the salience of an entity can lead to that entity receiving a topical interpretation receives support from Molnár (2006). She attempts a uniform account of topichood and claims, "Topics serve, namely, in one sense or another to optimally restrict the domain of the main predication in the sentence drawing the speech participants' attention to a certain entity. They can fulfil this function in two different ways: Either in an unmarked way by choosing a salient entity already in the focus of attention of the hearer/speaker, or by directing attention to this entity by highlighting the entity, in which case the co-occurrence with focus should be thinkable" (Molnár *et al* 2019: 28). In the cases where particle *-to* marks the sentence topic in Hindi, it employs the second mechanism that Molnár predicts can signal the topical status of a constituent— i.e., it directs the hearer's attention by imposing (discourse-)salience on an entity and thereby, highlights the *to*-marked entity in the sentence.

A *to*-marked topical XP can be a switch/shifted topic, a contrastive topic but not a continued topic. Incompatibility of particle *-to* with continued topics provides another argument for the salience-based analysis of particle *-to*. A constituent established as the topic in the preceding discourse becomes a salient entity (and as a result, is in the focus of attention). If the continued topic is marked with *-to* in the subsequent sentence, this leads to imposing salience on an already salient entity— which is pragmatically redundant. In example (2), this redundancy leads to the unacceptability of (2b), in contrast to the felicitous (2a). In (2b), the speaker B's second sentence has the continued topic (referent of *wo* 'she' has its antecedent as 'Riya', the preceding topical constituent) marked with enclitic *–to*, thereby marking as salient an already salient entity. In contrast to this salience-based analysis, a topic particle analysis

cannot account for the source of infelicity that is obtained when continued topics are marked with *-to* in Hindi.[10]

(2) Speaker A: Did Riya go to the party?
Speaker B: Yes. She was late though.
Speaker A: 'When did she come back?'

Speaker B: (a) wo gjarəh bədʒe wapəs aji
she eleven o'clock return come.PFV.FSG
(b) # wo=to gjarəh bədʒe wapəs aji
she=TOP eleven o'clock return come.PFV.FSG
'She came back at 11'o clock.'

Returning to the issue at hand, the next two sections provide empirical data that illustrates the non-topic interpretation yielded by *to*-marked XPs. Section 3.1 exhibits data for the non-topical type XPs being marked with particle *-to* and section 3.2 discusses the topical type *to*-marked XPs that are not sentence topic constituents. Section 3.2.4 claims 'contrast' to be the discourse-sematic import of particle *-to* in such non-topic interpretation cases.

### 3.1 *-to* and Non-topical type XPs

The distribution of the enclitic particle *-to* with respect to the types of constituents it can mark in Hindi was discussed in section 1.4 of chapter 1. Using the methodology stated above, these *to*-marked constituents can be segregated as being of a topical type or a non-topical type. In this section, examples (3)-(9) exhibit a range of non-topical type constituents that particle *–to* marks — an adjective, a determiner, a numeral, a quantifier and a postposition. In all the cases specified below, an inference can be drawn when a sentence with a *to*-marked XP is uttered in the given discourse context. These inferences are indicative of the semantic-pragmatic import of this particle in a sentence.[11]

[10] For some Hindi speakers, the (b) sentence in example (2) is not completely infelicitous but there is an observable degradation in the acceptability judgment from (2a) to (2b).

[11] These inferences are analysed in section 3.2.4 after the concepts required to do so are discussed.

Examples (3) and (4) show sentences where the particle *-to* marks either an adjective (*kali* 'black') or a determiner (*meri* 'my'), respectively, within a referential NP. Neither an adjective nor a determiner can be topical constituents in any approach to IS. The parenthetical at the end of the sentences indicates that these utterances are not informationally complete. An inference (apart from the asserted content) can be drawn by uttering these sentences, as specified in (3) and (4).

(3) Context: A person had misplaced a few shirts some days back. Today, he informs his parents:

| meri | [$_{Adj}$ kali]=to | kurti | mil | gəji | (pər...) |
|---|---|---|---|---|---|
| my | black=TOP | shirt | find | go.PFV.FSG | but |

'My black shirt was found (but...).'

Assertion: The speaker's black shirt has been found.
Inference: There exists (at least one, maybe more) other shirt(s) whose status is unknown in the current discourse context.

(4) Context: Two sisters had lost their matching shirts some time back. Today, one sister says to the other:

| [$_{D}$ meri]=to | kali | kurti | mil | gəji | (pər…) |
|---|---|---|---|---|---|
| my=TOP | black | shirt | find | go.PFV.FSG | but |

'My black shirt was found (but…).'

Assertion: The speaker's black shirt has been found.
Inference: There exists (at least) one other black shirt (the other sister's black shirt) in the current discourse context, whose status is unknown.

Examples (5) and (6) show an NP where the particle *-to* marks the numeral *do* 'two' and not the full NP subject. This numeral is neither what these sentences are about nor the frame within which their predication holds. Thus, these examples exhibit another instance of non-topical XPs that particle *-to* marks. It is observed that these two examples vary in terms of the inference that is evoked by *–to*-marking of a numeral within different discourse contexts.

5) Context: A teacher finished a tutorial quickly and came back to the staff room. His colleague asks him the reason behind his short class duration, to which the teacher responds:

| [$_{NUM}$ do]=to | bətʃtʃe | aje |
|---|---|---|
| two=TOP | child.PL | come.PFV.MPL |

'Two children had come.'

Assertion: Two children came.
Inference: Just two children had come (counter to speaker's expectation of more children coming to class).

(6) Context: A teacher had by mistake arranged for an extra class on a Sunday. By the time he realised this, it was too late to cancel it. On the day, he did not expect any of his students to turn up. Upon his return home from class, his partner asks him how many students came. He replies by saying:

| [NUM do]=to | bətʃtʃe | aa | gəje |
|---|---|---|---|
| two=TOP | child.PL | come | go.PFV.MPL |

'Two children had come.'

Assertion: Two children came.
Inference: At least two children did come (counter to speaker's expectation of lesser number of children coming to class).

The status of topichood of quantified DPs has been controversial in the IS literature. Since *referentiality* is taken to be one of the core criteria of topichood (as "only entities and classes of entities that can be presupposed to exist (…) can be predicated about" (É. Kiss 1992: 68)), quantified DPs are generally excluded from the set of topics as they are non-referential entities (cf. Frey 2004: 97). However, this correlation has been questioned by linguists like Horn (1989), Krifka (2008) and Büring (2016) who provide empirical evidence that exhibits that a language specific grammatical encoder of topic marking (like Japanese *-wa* marker or German fall-rise CT contour) is found with quantified DPs too. Krifka (2008) provides a theoretical explanation for ascribing topic interpretation to the quantified DPs. He claims that since a topic constituent "identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG [Common Ground] content", for a sentence with a quantified expression "[t]he quantifier in such sentences expresses the extent to which the comment holds for the elements of the set" (Krifka 2008:42). In (7), the quantifiers 'every' or 'most' express the extent to which information regarding the set of 'zebras' is updated in the CG. Crucially, in this analysis, the quantified DP is the topical constituent (and not the quantifier alone).[12]

[12] Quantified expressions were briefly discussed in the section 1.4.2.2 of chapter 1. A quantified DP can be semantically analysed as a quantifier-variable pair that creates a tripartite structure of the sentence it occurs in — [quantifier][restrictor][nuclear scope]. For the sentence in (i), its tripartite

(7) {Every zebra/Most zebras}$_{\text{TOPIC}}$ in the zoo were sick.

In Hindi, particle *-to* can mark either the quantified NP or the quantifier itself. However, there is a restriction on the distribution of the particle *-to* with different types of quantifiers (see section 1.4.2.2). The existential quantifier in Hindi (the wh-indefinite *koi/kuch* 'some') robustly gets marked with particle *-to* whereas the distributive universal quantifier (*har* 'every') cannot be marked with particle *-to*. Example (8) exhibits a sentence where particle *-to* marks an existential quantifier ('some') inside a quantified phrase ('some fruit').[13] Even if the quantified NP can get a topic interpretation (as reasoned in Krifka's framework), a quantifier itself cannot be a topical candidate in any approach to IS. Thus, the quantifier Q is another non-topical type entity that gets marked by particle *-to* in Hindi without yielding any sentence topic interpretation to it. Similar to the case of examples (3)-(6), an additional inference is also evoked by uttering (8) in the given discourse context.

(8) Context: Two boys were lost in a forest. They became hungry after a point of time and after a lot of searching they were about to give up the hunt when suddenly, they see a fruit tree from afar. One boy says to the other:

---

structure is given in (ii). Similarly, a semantic analysis has been proposed by Han (1998: 12) that analyses sentence topics as a topic operator-variable structure. These create a parallel tripartite division, as shown in (iii). Partee (1991) and Diesing (1992) have shown that focal elements map onto the nuclear scope and topics onto the restrictor scope. When a *-(n)un* marked quantified DP gets a topic interpretation, it falls in the restrictor scope of an operator-variable chain (Han 1998).

(i) Some student failed the exam.

(ii) $\exists$x [x=student]$_{\text{RESTRICTOR}}$ [x failed the exam]$_{\text{NUCLEAR SCOPE}}$

(iii) Top(x)[x is John][x likes Mary] for 'John likes Mary'

[13] A parallel example with the universal quantifier being marked with *-to* leads to ungrammaticality, as shown in sentence (i) below:

(i) Context: Two boys were told to find as many types of fruits as they could from a specific fruit garden. After completing the hunt, one boy says to another:

| | | | | |
|---|---|---|---|---|
| * həme | [hər]=to | p$^{h}$əl | mil | gəja |
| us.OBL | every=TOP | fruit | find | go.PFV.MSG |

'We found every fruit.' *(intended)*

| | | | | |
|---|---|---|---|---|
| həme | [koɪ]=to | pʰəl | mil | gəja |
| us.OBL | some=TOP | fruit | find | go.PFV.MSG |

'We found some fruit.'

Assertion: The speaker (and the hearer) has found some fruit.
Inference: The speaker (and the hearer) has found at least one/some fruit (counter to their expectation of not finding anything).

Example (9) shows *-to* marking of a postposition *neeche* 'under'. This constituent is not the topic of the sentence as it is neither a referential entity 'about' which the rest of the sentence is nor is it a permissible frame of predication. Thus, postpositions are another type of non-topical constituents with which particle *-to* can cliticise.[14] Like the preceding examples, this sentence with a *to*-marked XP evokes an inference within the given discourse context.

(9) Context: Geeta has lost her cat and she has been given directions to search every nook and corner of her garden, given the nifty nature of cats in general.

| | | | | |
|---|---|---|---|---|
| gita=ne | peɽ=ke | nitʃe=to | dekʰa | (pər..) |
| Geeta=ERG | tree=GEN | under=TOP | see.PFV.MSG | (but..) |

'Geeta looked under the tree.'

Assertion: Geeta searched at the bottom of the tree.
Inference: It is uncertain whether she looked at other places (provided that she was specifically told to check other alternative locations in the garden).

To sum up this section, the data given in examples (3)-(9) corroborates the central argument of this thesis — that particle *-to* is not a topic marker. This particle can

[14] Note that a PP marked with Hindi particle *-to* is different from an English PP preposed or topicalized, like in (i) below (cited from Prince (1984: 214); underlined constituent is unstressed, capital letters indicate nuclear stress of the sentence), in terms of both the constituent that gets grammatically marked as well as the interpretative effect of that marking.

(i) With Rosa, Felix went to the BEACH.

In a preposed English PP, it is the NP within the PP (i.e. 'Rosa') that gets a contrastive topic interpretation (Reinhart 1981). In the case of Hindi, it is the non-topical postposition that gets lexically marked by particle *–to* and it is either the full PP or the postposition (and not the noun) that gets associated with the semantic import of this particle. The inference in (9) is about the discourse status of an alternative to a *–to*-marked PP 'under the tree'. Another inference is possible where the relevant alternative is just the *–to*-marked postposition 'under' as in the case of 'It is uncertain whether she looked *above* the tree.'

mark non-topical constituents (like adjectives, determiners, quantifiers and postpositions) that can never receive a topic interpretation at the C-I interface as these constituents do not fulfil the definitional criteria entailed by the notion of topic. Conversely, if topicality was indeed an inherent property of this particle, then it would have forced a topic interpretation on the *to*-marked XPs in examples (3)-(9). But this is evidently not the case since the topical constituents in all the sentences in (3)-(9) are their respective subject NPs (as can be verified by preceding these utterances with the *Tell me about X* diagnostic template).

The next section exhibits the *-to*-marking of topic-type XPs in Hindi that do not receive the sentence topic interpretation at the CI interface.

## 3.2 *to*-marked Topic-type XPs that are NOT topics

For this section, I focus on the referential entity-denoting NPs out of the list of topic-type XPs mentioned in the introduction of this chapter.[15] Section 3.2.1 diagnoses which NPs have a topical bias (that are sentence initial) and which do not (that are not sentence initial) and thus, can be used as a test case to evaluate any potential non-topic interpretation. Section 3.2.2 tests the in-situ object NPs in Hindi for any evoked interpretative effect. Sections 3.2.3 and 3.2.4 claim that these non-topical in-situ object NPs receive a contrastive interpretation, that is facilitated via an uncertainty or a scalar implicature evoked in these cases.

### *3.2.1 Diagnosing Non-topic NPs*

Within a typological bifurcation of the world's languages as being either topic-prominent or subject-prominent or both or neither (Li and Thompson 1976), Hindi has been classified as a subject-prominent language.[16] Thus, as a standard assumption,

[15] The reason for this selection is that I employ the *Tell me about X* test during the course of analysis in this sub-section and these NPs exhibit the clearest results for the diagnostic topichood test.

[16] Li and Thompson (1976) themselves do not explicitly mention Hindi in the list of languages they typologize. They only mention the superset 'Indo-European' to be subject-prominent.

subject NPs receive a default topic interpretation in a Hindi sentence. Kidwai (2004) has given a syntactic proposal for deriving the typological parameters of subject and topic prominence of a language and motivated a syntactic feature based account for Hindi subjects receiving a topic interpretation.[17] Thus, *prima facie,* subject NPs are biased to receive a topic interpretation. This is exhibited in example (10) where the subject NP 'Rahul' is pre-designated as the sentence topic constituent (speaker A's utterance employs the *Tell me about X* template).[18] This indicates that a non-subject NP marked with particle *-to* would be the ideal test candidate to analyse the non-topic interpretation yielded by *-to* marking of a topic-type XP.

(10) Speaker A: mʊdʒʰe rahʊl=ke bare=mẽ bətao
I.DAT Rahul=GEN about tell.IMP.2SG
'Tell me about Rahul.'

Speaker B: [rahʊl=ne]$_{TOPIC}$ seb kʰaja
Rahul=ERG apple eat.PFV.MSG
'Rahul ate the apple.'

On a closer analysis, not just the subject NP but any NP that is in the sentence-initial position is biased for a topic interpretation. I propose this based on the examples (11)

[17] Kidwai (2004) argues for a generative treatment of (sentence) topic interpretation within a minimalist architecture of Universal Grammar. She reformulates Chomsky's (1996) universal thematization requirement as a requirement for each sentence to have a topic. This topic interpretation is syntactically sourced from the EPP feature that a functional head gets 'merged with' in the numeration. For a subject-prominent language like Hindi, this EPP feature is merged with the T(ense) functional head and any XP that occupies the [spec,TP] position and checks its EPP feature receives the topic interpretation at the C-I interface of derivation process. Following standard syntactic assumptions, subject arguments are the XPs that move into [spec,TP] position and thus, get topic interpretation by virtue of their checking the EPP feature.

Although this analysis works for some unmarked topic XPs, it cannot account for all *to*-marked XPs. Kidwai herself notes that her framework raises questions for the topic particle *-to* (see Kidwai 2004:2). Neither are all –to-marked XPs subjects and nor do they all receive a topic interpretation.

[18] Speaker B's response is in the SOV base order in example (10). If the speaker A had fixed object NP as the topical entity for its subsequent utterance, then speaker B's response would be infelicitous in that situation. This is exhibited below:

Speaker A: mʊdʒʰe seb ke bare mẽ bətao
'Tell me about the apple.'
Speaker B: # rahʊl=ne seb kʰaja
'Rahul ate the apple.'

and (12). For an unaccusative verb 'to break', the theme argument 'mirror' moves into the syntactic subject position and the grammatical subject is an instrumental case marked adjunct. In both (11) and (12), the subject NP 'mirror' is the pre-designated topic for speaker B's response fixed by speaker A's utterance. Speaker B's response is felicitous in (11) but infelicitous in (12) which I propose is because the pre-designated topical NP is not in the sentence-initial position in (12).

(11) Speaker A: mʊdʒʰe aɪne=ke bare=mẽ bətao
I.DAT mirror.OBL=GEN about tell.IMP.2SG
'Tell me about the mirror.'

Speaker B: aɪna rahʊl=se ʈuʈa
mirror Rahul=INS break.PFV.MSG
'The mirror was broken by Rahul.'

(12) Speaker A: mʊdʒʰe aɪne=ke bare=mẽ bətao
I.DAT mirror.OBL=GEN about tell.IMP.2SG
'Tell me about the mirror.'

Speaker B: # rahʊl=se aɪna ʈuʈa
Rahul=INS mirror break.PFV.MSG
'The mirror was broken by Rahul.'

Thus, any NP in a sentence-initial position is biased to yield a sentence-topic interpretation in Hindi. In the base order, subject NPs occupy the sentence-initial position. Thus, subject NPs are the entities that are most frequently associated with (an aboutness) topic interpretation in a sentence. This tendency is attested cross-linguistically in the IS literature.[19] Many linguists have proposed language-specific analyses that identify topic interpretation with the constituent occupying the leftmost position in the sentence (a functional projection of Ref(erential) Phrase as proposed by Szabolcsi (1997) for Hungarian) or a sentence–initial hierarchical position (occupied by A(boutness)-topics in Italian as claimed by Bianchi and Frascarelli 2010). The preference of topics to occupy the left-peripheral position is also noted by

[19] But this is not an absolute correlation since exceptions to it exist. For example, Chafe (1976: 50) gives data from Chinese multiple subject construction (see example (i) below) where what the predicate of the sentence is 'about' is not the sentence-initial NP *nei-xie shumu* but the second argument *shu-shen*.

(i) nei-xie shumu shu-shen da
those tree tree-trunk big
'Those trees the tree trunk is big'

Givón 1992, Jacobs 1997, Vallduví and Engdahl 1996, É. Kiss 1987, 2002 etc. For a language like Japanese that has multiple NPs being marked with a topic marker *-wa*, the first *-wa*-marked NP occupying the sentence-initial position is interpreted as the thematic (non-contrastive) topical constituent (Heycock 2008).[20]

This section has diagnostically tested and proved that the sentence-initial subject NPs have a topical bias. Consequently, they are filtered out from the set of topical-type XPs that can get a non-topic interpretation. An object NP is a possible candidate for a topic interpretation and an in-situ object NP is not biased for a sentence topic interpretation (Since Hindi has the base order SOV). If Hindi sentence-medial object NPs can be *to*-marked, then they would provide evidence for a topical type *to*-marked XP that is still not the topical constituent in the sentence. Such a configuration would strengthen the argument for the particle *-to* not being a topic marker by itself. The next section 3.2.2 provides data to exhibit that such a configuration actually exists in Hindi.

### *3.2.2 to-marking of in-situ object NPs*

In example (13), the two sentences form a minimal pair that only vary with respect to object NP *kela* ('banana') being unmarked in example (13a) and being *to*-marked in example (13b).[21] Since the discourse context is the same for both the sentences and

[20] No other *-wa*-marked NP that is not in a sentence-initial position is interpreted as a contrastive topic. This is exhibited in the example (i) below (example (4) from Heycock 2008:11) where *watasi-wa* is the thematic topic and *tabako-wa* and *sake-wa* are contrastive topics.

(i) watasi wa tabako wa suimasu ga sake wa nomi-masen
I wa cigarette wa smoke but alcohol wa drink-NEG
'I smoke but I don't drink.'

[21] An anonymous reviewer for Triple A7 Workshop (2020) had suggested that the sentences in example (13a) and (13b) should be preceded by a QUD like 'What happened?' or 'What about Ravi?' to be properly anchored in discourse. I accept their suggestion that the discourse coherence improves by having a preceding QUD of the type they mention, which I had left implicit in this example. This also favours the point that I am making here that crucially, neither of the QUDs mark the object NP 'banana' as the topical constituent. A preceding QUD of 'What happened?' leads to a stage topic interpretation for the response as a thetic sentence (see section 2.2.2). A preceding QUD of 'What

the only variable factor is particle *-to*, thus the only probable source of infelicity of sentence (13b) is *-to*-marking of its object NP. The previously diagnosed licensing condition for *-to*-marking of an XP —that the denotation of that XP be available in the Common Ground between the interlocutors (see section 2.3.2)—is fulfilled in this example. This indicates that there must be another pragmatic licensing constraint that is specific to the *-to*-marking of an object NP.

(13) Context: During the school Board examinations period, Ravi's parents leave a banana for him to eat in his room during his break. Upon seeing the banana peel thrown by him in the dustbin, his mother says to his father:

(a) rəvi=ne kela kʰa lija
Ravi=ERG banana eat take.PFV.MSG
'Ravi has eaten the banana.'

(b) # rəvi=ne kela=to kʰa lija
Ravi=ERG banana=TOP eat take.PFV.MSG
'Ravi has eaten the banana.'

Example (14) exhibits that a minimal modification of the context that precedes the sentence containing a *to*-marked object NP leads to an alteration in the pragmatic status of example (13b) and makes it felicitous. The discourse context in example (14) has a C(ommon) G(round) that contains two items — an 'apple' and a 'banana' — rather than one item like the CG of example (13). Both the unmarked and the *to*-marked object NPs are acceptable in this context. The sentences in example (13a) and (13b) are repeated as (14a) and (14b) respectively, keeping the (implicit) QUD constant.

(14) Context: During the school Board examinations period, Ravi's parents leave a banana and an apple for him to eat in his room during his break. Upon seeing the banana peel thrown by him in the dustbin, his mother says to his father:

(a) rəvi=ne kela kʰa lija
Ravi=ERG banana eat take.PFV.MSG
'Ravi has eaten the banana.'

---

about Ravi?' leads to 'ravi' being fixed as the pre-designated topical constituent in the response utterance. In this example, a non-topical NP 'banana' gets *to*-marked.

Assertion: Ravi ate the banana.
Inference: None

(b) rəvi=ne kela=to kʰa lija
Ravi=ERG banana=TOP eat take.PFV.MSG
'Ravi has eaten the banana.'

Assertion : Rave ate the banana.
Inference 1: The speaker does not know whether Ravi ate the apple.
Inference 2: The speaker knows that Ravi did not eat the apple.
Inference 3: The speaker knows but does not want to give the information regarding the status of apple to the hearer (maybe in the case where the speaker knows that the hearer is only interested in knowing about the status of 'banana' and no other fruit).

A comparison of examples (13) and (14) indicates two things. First, a felicitous licensing of particle *-to* on an object NP requires that there must exist at least one alternative entity to the entity denoted by the object NP in the Common Ground between the discourse participants. Thus, the cardinality of the set of alternatives should be more than one.[22] Second, a sentence containing an unmarked object NP has only an assertion part whereas a (felicitous) sentence containing a *to*-marked object NP generates an inference part as well, besides the assertion part. Thus, when an entity from the set of alternatives is marked with particle *-to*, it triggers some specific inferences regarding the discourse status of the remaining members of this set. These inferences could have an epistemic base (based on the speaker's knowledge/belief

[22] This pragmatic constraint is similar (but not identical to) the presuppositional view of Korean particle *-(n)un* proposed by Han 1998. According to her analysis, "the marker *–(n)un* presupposes a non-empty set (including a singleton set)…NP-*nun* presupposes a set whose member is at least the entity picked out by the NP" (Han 1998: 5). She gives a semantic formalization of this as:

(i) presupposition of α-(n)un, where α is an individual and X is a set variable over individuals:

$\exists X[(\alpha \epsilon X) \Lambda (|X| \geq 1)]$

In her system, the topic, the contrastive topic and the contrastive focus reading of *-(n)un*-marked NPs all have the presupposition in (i) in common. See Kim (2013: 70-73) for a critique of why topicality and contrast cannot be derived from an analysis of *-(n)un* as a presupposition marker. In my analysis, only the non-topical XPs have this set of alternatives (where the cardinality of the set is $|X|>1$) obligatory requirement. Topical XPs can optionally get a contrastive interpretation that is contingent on the existence of the set of alternatives in the discourse domain. Additionally, since *-to* can mark even non-nominal categories— like Adj, Num, Q, P etc — this presuppositional analysis, contingent on a set of individuals, cannot be extended to these categories.

system) or could arise because of some non-epistemic reason too like the speaker's desires or politeness concerns. In example (14b), the first and second inferences are epistemic whereas the third inference shows one type of a non-epistemic inference that could be generated when this sentence is uttered in the given discourse context. This chapter will discuss the epistemic inferences at length in sections 3.2.3 and 3.3 but will briefly touch upon the importance of the non-epistemic inference in section 3.3.3. The inference 1 in (14) is an example of an uncertainty inference (the speaker does not know/ is uncertain about the status of other entities in the set of alternatives). The inference 2 in (14) is an example of a scalar inference (discussed in section 3.3.1). Before moving ahead with the analysis of these epistemic inferences, I test the type of inference generated by the *-to*-marking of the non-topical object NP in Hindi. The next section (section 3.2.3) diagnoses the type of inference triggered by particle *-to* using a battery of tests, after providing a brief overview about the different types of inferences in section 3.2.1.

### *3.2.3 Investigating the Inference triggered by particle -to*

#### 3.2.3.1 Types of Inference: An Overview[23]

Strictly speaking, an inference (or an implication) is a relation that a sentence bears with respect to another sentence by the virtue of it being licensed either by the informational content of that other sentence or by the conversational expectations that are implicit in uttering that other sentence. Inferences have traditionally been classified into three types — entailments, presuppositions and implicatures. Each inference type and their respective diagnostic tests are discussed below. This battery of diagnostic tests is then further applied to the inferences generated by a *to*-marked object NP in the section 3.2.3.2.

*1. Entailment*

An entailment is a strong implication relation that is formally defined as: *sentence A* entails *sentence B* if and only if whenever *A* is true, *B* has to be true too. This

[23] This section is based mostly on Chierchia and McConnell-Ginet (1990) and Coppock and Champollion (2019).

inference type is exhibited via example (15) that involves an entailment relation between a pair of English sentences. In every circumstance/situation where sentence A is true (i.e. in a situation where the speaker's father indeed cooked their breakfast and after the completion of this event, the speaker washed their utensils), sentence B will be true too (i.e. it would be necessarily true that the father would have cooked the breakfast). Thus, *A* entails *B* in example (15).

(15) Sentence A: After my father cooked our breakfast, I washed the utensils.
Sentence B: My father cooked the breakfast.

A diagnostic test used to evaluate entailment relations is the 'Defeasibility Test' since entailments are strong inferences that are not defeasible. In this test, for a pair of sentences A and B (where A entails B), '*A* and *not B*' should be contradictory.[24] To prove the entailment relation predicted for the pair of sentences in example (15), a defeasibility test is applied to them in (16).

(16) *A*: After my father cooked our breakfast, I washed the utensils.
*B:* My father cooked the breakfast.
*not B:* It is not the case that my father cooked the breakfast.
*A and not B:* # After my father cooked our breakfast, I washed the utensils and it is not the case that my father cooked the breakfast.

In a situation where sentence A is evaluated as true, the speaker's father did cook their breakfast and the speaker did wash their utensils. In a situation where negation of sentence *B* is true, the speaker's father did not cook the breakfast. A proposition cannot be both true and false at the same time within a possible world. Thus, *A and not B* is indeed contradictory confirming that whenever *A* is true, *B* has to be true too. This proves that *A* entails *B* in example (15).

*2. Presupposition*

Presupposition as an inference relation that can be formally defined as: sentence *A* presupposes sentence *B* if and only if:

(i) *A* implies *B*

[24] Conventionally, negation of a proposition P is paraphrased as 'It is not the case that P'.

(ii) The truth of *B* is taken for granted as the background for considering *A*.

Example (17) shows a pair of English sentences where sentence A implies sentence B and also, the truth of sentence B is taken for granted for computing the semantic value of sentence A. Thus, sentence A presupposes sentence B. The existence of a unique hostel president is assumed to be uncontroversial and taken to be true as background information over which sentence A is asserted.

(17) Sentence A: The hostel president has knee-length hair.
Sentence B: There is a unique hostel president.

It is a characteristic property of presuppositions that they project through a set of contexts. These contexts include an affirmative declarative, a negative declarative, an interrogative and an antecedent of a conditional. For a presupposition P, this set is conventionally called as the P-family of contexts. Presuppositions are diagnosed through a 'Projection Test' i.e., sentence *A* presupposes sentence *B* if and only if not only *A* but also other members of P-family imply (and assume as background) *B*.

To verify the predicted presupposition relation in example (17), the pair of sentences A and B are evaluated via a projection test. Example (18) exhibits the P-family of contexts for sentence A of example (17).

(18)
a. Affirmative Declarative : The hostel president has knee-length hair.
b. Negative Declarative : The hostel president does not have knee-length hair.
c. Interrogative : Does the hostel president have knee-length hair?
d. Antecedent of Conditional: If the hostel president has knee-length hair, then she might have to spend a lot on hair-care products.

In all the contexts (18a)-(18d), it is assumed to be true that there is a unique hostel president. Thus, this proposition i.e., sentence B of example (17) is the underlying presupposition for sentence A of example (17) since it successfully projects through all the P-family of contexts in example (18).

*3. Implicature*

A theory of implicatures has its origin in the work of Grice (1975, 1978), who stated that co-operative discourse participants typically follow certain norms of conversation (called 'maxims of conversation') that are assumed to be mutually known to them. These maxims of conversation are the "rules that govern allowable conversational moves" (Simons 2012: 2). They are — Maxim of Quality (try to make your contribution one that is true), Maxim of Quantity (make your contribution as informative as is required), Maxim of Relation (be relevant) and Maxim of Manner (be perspicuous). Grice stated that the discourse participants typically adhere to an over-arching principle — a Cooperative Principle that requires one to "make a conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of talk exchange in which one is engaged" (Simons 2012: 2). Implicatures are the inferences that are evoked based on a shared presumption of interlocutors that they obey these maxims. In certain cases, an apparent violation of a maxim leads to some implicature generation.

An implicature can be defined as: sentence *A* implicates sentence *B* if one takes *B* to be a part of what the speaker of *A* meant by uttering the sentence *A*. In example (19), an implicature can be drawn from sentence *A* to sentence *B*. The speaker of sentence *A* takes the informational content of sentence *B* to be a part of the full information that he wants to convey by uttering sentence *A*. The co-operative speaker gives a response that is 'informative' and 'relevant' for the discourse at hand by uttering sentence *A* that implicates sentence *B*.

(19) Sentence *A*: I used to walk 5 kms daily
Sentence *B*: I do not walk 5 kms daily.

Implicatures are weaker inferences (as compared to entailments or presuppositions) that tend to disappear under negation. They can be defeated or reinforced depending on the addition of appropriate contextual information. The diagnostic tests that have been proposed to evaluate implicatures operate on this core property. These tests include a 'Defeasibility Test' and a 'Redundancy Test'.

(i) Defeasibility Test: According to this diagnostic, since implicatures are defeasible, if sentence *A* implicates sentence *B* then '*A* and *not B*' should not be contradictory (unlike entailments). This test is applied to the pair of sentences in example (19) in a step-by-step method in (20).

(20) *A*: I used to walk 5 kms daily
*B*: I do not walk 5 kms daily.
*Not B:* It is not the case that I do not walk 5 kms daily.
≈ I walk 5 kms daily.[25]
*A and not B:* I used to walk 5 kms daily. In fact, I still walk 5 kms daily.

In example (20), the context is enriched (with the use of 'in fact' and 'still') to connect sentence *A* and sentence *Not B*. The pair of sentences for *A and not B* are coherent and not contradictory. The inference generated by the first sentence is defeasible, so much so that it can exist with the negation of that inference. This proves that sentence *B* is indeed an implicature drawn from sentence *A*.

(ii) Redundancy Test: According to this diagnostic, since implicatures are reinforceable, if sentence *A* implicates sentence B then '*A* and *B*' should not be redundant. Example (21) exhibits application of this test to the pair of sentences in example (19).

(21) *A*: I used to walk 5 kms daily
*B*: I do not walk 5 kms daily.
*A and B:* I used to walk 5 kms daily. But I do not walk 5 kms daily anymore (because of the pandemic situation).

The pair of sentences for *A and B* are not redundant. The inference generated by the first sentence is reinforced using the connective 'but' without being redundant. This re-affirms the result of the preceding test that sentence *B* is indeed an implicature that is drawn from sentence *A* in example (19).

To conclude, this section outlined the three types of inferences that can be drawn from a sentence and the diagnostic tests available for each, based on their characteristic properties.

[25] I use the sign ≈ to indicate logical equivalence in the meaning conveyed by two sentences.

3.2.3.2 Testing the type of inference generated by *to*-marked object NP

In section 3.2.2 it was shown that *-to*-marking of a sentence-medial object NP leads to triggering of some inferences, which are absent in their unmarked counterparts. To investigate the type of inference triggered by particle *-to*, I run the relevant pair of sentences through the diagnostic test machinery. For the ease of discussion, I repeat example (14b) from the section 3.2.3 as example (22) below.

(22) Context: During Board examinations period, Ravi's parents leave a banana and an apple for him to eat in his room during his break. Upon seeing the banana peel thrown by him in a dustbin, his mother says to his father:

| | | | |
|---|---|---|---|
| rəvi=ne | kela=to | $k^h$a | lija |
| Ravi=ERG | banana=TOP | eat | take.PFV.MSG |

'Ravi has eaten the banana.'

Assertion : Rave ate the banana.
Inference 1: It is uncertain whether Ravi ate the apple.
Inference 2: Ravi did not eat the apple.

In this section, I test the inference 2 and will extend the result obtained to inference 1. For expository purpose, I label the *to*-marked sentence in example (22) as sentence *A* and the inference under discussion (inference 1 given in Hindi in (23)) as sentence *B.*

(23)

| | | | |
|---|---|---|---|
| rəvi=ne | seb | nəhi | $k^h$aja |
| Ravi=ERG | apple | NEG | eat.PFV.MSG |

'Ravi did not eat the apple.'

*1. Testing for Entailment*

Entailments are diagnosed using the defeasibility test. Since entailments are not defeasible, if a sentence *A* entails a sentence *B* then *'A and not B'* should be contradictory. Applying this test to sentences A and B in (24):[26]

(24) *A*: rəvɪ=ne kela=to $k^h$aja
'Ravi ate the banana.'

[26] I do not provide the interlinear glosses for this section as they are not relevant for the discussion at hand.

| | |
|---|---|
| *B*: | rəvɪ=ne seb nəhɪ kʰaja<br>'Ravi did not eat the apple.' |
| *Not B*: | It is not the case that Ravi did not eat the apple.<br>≈ Ravi ate the apple. |
| *A and not B*: | rəvɪ=ne kela=to kʰaja. us=ne seb bʰɪ kʰa lɪja<br>'Ravi ate the banana. He also ate the apple.' |

The pair of sentences in *A and not B* exhibit that it is possible to defeat the inference (sentence *B*) generated by sentence *A*. With a same situation, sentence *A* is 'true' and sentence *B* is 'false' (since the negation of sentence B is 'true'). Thus, this proves that the inference being tested is not an entailment.

*2. Testing for Presupposition*

Presuppositions project through the P-family of contexts. Applying this projection test to sentence *A* and its inference (sentence *B*) in (25), which lists the P-family of contexts for sentence *A*:

| | |
|---|---|
| *A*: | rəvɪ=ne kela=to kʰaja<br>'Ravi ate the banana.' |
| *B*: | rəvɪ=ne seb nəhɪ kʰaja<br>'Ravi did not eat the apple.' |
| *Affirmative Declarative*: | same as *A* |
| *Negative Declarative*: | rəvɪ=ne kela=to nəhɪ kʰaja<br>'Ravi did not eat the banana.' |
| *Interrogative:* | * kja rəvɪ=ne kela=to kʰaja ?[27] |

[27] This information-seeking yes-no question with a polar *kya* 'what' is ungrammatical in Hindi with a *to*-marked object NP. Neither can the wh-constituent questions be marked by particle –*to* in Hindi (see section 2.3.1). However, a confirmatory question (where the speaker indicates that she believes the affirmative sentence to be true but wants the hearer to verify it) is acceptable with -*to*-marking in Hindi. This data point is credited to a recent discussion between Ashwini Deo and Rajesh Bhatt in FASAL 2021 conference. An example for this is provided below, uttered with a rising intonation. An analysis for this type of constructions is proposed in chapter 5.

(i) Context: A teacher knows which of her students had prepared dances for the school festival but the teacher could not attend the festival herself. Later, she wants to confirm whether her not-so-confident student was able to perform on stage or not. She asks her colleague:

‘Did Ravi eat the banana?’

*Antecedent of Conditional* * əgər rəvɪ=ne kela-to kʰaja, təb us=ko dʒəldɪ bʰukʰ nəhɪ ləgegɪ[28]
‘If Ravi ate the banana, then he will not feel hungry soon.’

Since the negative declarative is the only grammatical form (apart from test case affirmative declarative) that permits *-to*-marking of the object NP, the projection test is applied to this construction in the context. The potential candidate for presupposition (sentence *B*) is neither implied nor assumed to be ‘true’ as background information when the negative declarative form is considered. Thus, the negative declarative context does not let *B* project through it as a presupposition. Hence, the inference in *B* is not a presupposition for sentence *A* either.

*3. Testing for Implicature*

Implicatures, being relatively weak inferences, are defeasible or re-inforceable depending on the type of contextual enrichment. A defeasibility test, which requires *A and not B* to be non-contradictory for implicatures, is exhibited in example (26). By using the additive particle *bʰɪ* ‘also’, the inference (sentence *B)* drawn from sentence *A* is defeated in the acceptable, non-contradictory *A and not B* discourse unit. Thus, this proves that the status of the inference drawn in this case is that of an implicature.

(26) *A*: rəvɪ=ne kela=to kʰaja
‘Ravi ate the banana.’

*B*: rəvɪ=ne seb nəhɪ kʰaja
‘Ravi did not eat the apple.’

*Not B*: It is not the case that Ravi did not eat the apple.
≈ Ravi ate the apple.

*A and not B*: rəvɪ=ne kela=to kʰaja. us=ne seb bʰɪ kʰa lɪja
‘Ravi ate the banana. He also ate the apple.’

---

sɪta=ne=to dans kɪja
sita=ERG=TOP dance do.PFV.MSG
‘Did Sita dance?’
‘Sita danced, right?’ (*intended meaning*)

[28] This construction is ungrammatical because enclitic particle *-to* cannot mark an XP in the antecedent clause of a conditional sentence.

This result is verified using another test— the Redundancy Test. Since implicatures are reinforceable, the *A and B* structure should not sound redundant (unlike in the case of entailment). This test is applied to sentence unit *A and B* in example (27). Using the connective *məgər* 'but', the inference is reinforced without it being redundant. Thus, *B* is an implicature drawn from *A.*

(27) *A*: rəvɪ=ne kela=to kʰaja
'Ravi ate the banana.'

*B*: rəvɪ=ne seb nəhɪ kʰaja
'Ravi did not eat the apple.'

*A and B*: rəvɪ=ne kela=to kʰaja. məgər u=ne seb nəhɪ kʰaja
'Ravi ate the banana. But he did not eat the apple.

Since the inference 2 (from example (22)) is diagnosed to be an implicature, I directly perform the defeasibility test on inference 1. As shown in example (28), inference 1 (labelled as sentence B here) is defeasible in the *A and not B* pair of sentences. Thus, this inference is also an implicature drawn from sentence A.

(28) *A*: rəvɪ=ne kela=to kʰaja
'Ravi ate the banana.'

*B*: mʊdʒʰe pəta nəhɪ kɪ rəvɪ=ne seb kʰaja ja nəhɪ
'It is not certain/ I do not know whether Ravi ate the apple (or not).'

*Not B*: It is not the case that it is not certain that Ravi ate the apple.
≈ It is certain/ I know that Ravi ate the apple.

*A and not B*: rəvɪ=ne kela=to kʰaja hɛ, ɔr mʊdʒʰe pəta hɛ kɪ ʊs=ne satʰ mɛ̃ seb bʰɪ kʰaja hɛ
'Ravi ate the banana. And I know that he also ate the apple with it.'

To summarise, this sub-section analysed that the type of inferences triggered by enclitic particle *-to* when it marks a non-sentence-initial NP is that of an implicature. This claim is verified using the diagnostic tools of defeasibility test, projection test and redundancy test. An analysis for these implicatures is provided in section 3.3.

### *3.2.4 Contrast as the non-topical interpretation of to-marked object NP*

Sections 3.2.2 and 3.2.3 have established that Hindi particle *-to* can mark a topical-type XP without yielding a topic interpretation to the denotation of that constituent. A sentence-medial object NP was selected as the test case. After investigation, it was concluded that:

(i) *-to*-marking of an object NP requires a set of alternative entities to be available in the Common Ground (a pragmatic licensing constraint)

(ii) *-to*-marking of an object NP can evoke implicatures regarding the discourse status of the remaining members of the set of alternatives (an interpretative effect).[29]

There are two notions in the IS literature that require a set of alternatives as a necessity for their interpretation —*contrast* and *focus* — with some overlap in the two notions. 'Contrast' can be understood pre-theoretically as a relation that exists between two sentences $s_1$ and $s_2$ if the sentence $s_1$ "contains an element $\alpha$ that can be construed as an *alternative* to $\beta$ in $s_2$, where being *construed as an alternative* reflects the notion of *juxtaposition* and *comparison*" (Repp 2016: 2). Different theories of contrast posit differing requirements on what can constitute as an 'alternative' to some constituent or whether such alternatives should be explicit in the preceding sentence or whether they can be implicit (in the discourse) too. For example, Kiss (1998) requires the set of alternatives to be a contextually restricted set such that the members of the set are clearly identifiable by the participants of a discourse. A broader perspective than this is taken up by Vallduví and Vilkuna (1998) or Katz and Selkirk (2011), who equate contrast with only the presence of alternatives, with no added stipulation on the set membership or the implicitness/explicitness of the alternatives. In their view, since alternatives are different from each other, this is sufficient criteria to establish a contrast between them.

[29] By 'discourse status' I mean whether the predicate holds true or not for the members of the set of alternatives in the discourse context. The assertion (in example (22)) indicates that predicate holds true for the *to*-marked entity.

Another view on the notion of contrast requires an alternative to exist such that by substituting the original element at issue in the sentence, the sentence turns out to be false (eg. Halliday 1967, Chafe 1976, Neeleman and Vermeulen 2012). This understanding of the notion of contrast introduces an exhaustive meaning into it since it involves exclusion of the other alternative from the set of elements that can make the sentence true. Yet another view on contrast is to link it to the belief system of discourse participants (e.g., Halliday 1967, Frey 2006, 2010). The alternative selected out of the set of alternatives by the speaker is unexpected or remarkable in some way. Thus, defining 'contrast' is ridden with the same issues as defining other notions of information structure like 'topic' or focus' (see chapter 2) — that of multitude of theoretical approaches with different defining criteria.

Parallel to the theories of contrast that equate it with alternatives, there is a framework of focus that equates it with alternatives. The Alternative Semantics theory developed by Rooth (1985, 1992), and adopted by Krifka (2008), requires focus-marking to inherently involve contrast. A focussed constituent introduces a set of entities (with focus being signalled by a pitch accent) and the focussed entity is selected from this set of alternatives. This is theoretically implemented by requiring the semantics to establish a focus semantic value separate from the ordinary semantic value such that a focus-marked element β contrasts with an element α if the ordinary semantic value of α is a member of the focus semantic value of β. This means that the denotation of α belongs to the set of focus alternatives of β. In this approach, focus is exhaustive since focus-marking of β in a sentence exhausts the set of alternatives (none of the other alternatives return a true value when substituted in the same sentence). This theory of focus cannot account for the sub-types of focus like 'new information' focus (that involves no alternatives) or a separate 'contrastive' focus.

Since felicitous *-to*-marking of an object NP is contingent upon the availability of a set of alternatives in the discourse context (see point (i) in the introduction of this section), this particle could signal either contrast or focus for the object NP. I claim that *-to* marking of object NPs in Hindi involves *contrast* and not *focus* as their IS interpretation. I adopt the definition of contrast as proposed by Kim (2013) to account for how contrast can be indicated by a constituent being marked by particle *–to* in a

sentence. Kim defines contrast as "a relation between discourse referents that are partitioned with respect to some semantic property P such that it is established (either via assertion or implicature) that the value 'true' results when P is applied to one part of the set and 'false' or 'unknown' when applied to the other" (Kim 2013: 50).

The particle *-to* marks an object NP when a set of alternatives to the denotation of that NP is available in the discourse context. This set of alternatives might have been mentioned directly in the preceding sentences (i.e., an explicit alternatives set) or they might be inferable from the context because they involve relations between "kinds and their representatives, plural individuals and their atomic parts, generalized quantifiers and elements of their witness sets" etc., thus forming an implicit alternatives set (Repp 2016: 6). Example (29) exhibits both the types of alternative sets that can license *-to*-marking of object NP. Speaker A explicitly mentions the members of the set of alternatives whereas speaker B mentions the plural term 'subjects' from which the atomic set members 'Maths' and 'Science' can be inferred, provided that this set has been introduced in the context before since this set membership is context-specific.

The preceding sections (3.2.2 and 3.2.3) had established that certain implicatures are evoked when *-to*-marks an in-situ object. The implicatures generated by uttering a *-to* marked sentence after speaker A or B's question in example (29) are specified in (29a). These implicatures fulfill the second criteria of Kim's definition of contrast — that of establishing via implicature (in this case) that the relevant semantic property P ('studied by Geeta') returns a 'true' value when applied to one part of the set ('Maths') and returns a 'false' or 'unknown' value when applied to the other part of the set of alternatives ('Science'). Thus, a partitioning of the set of discourse referents can be implied by *-to*-marking of the object NP in the sentence in example (29a). The application of the semantic property P returns varying truth values when applied to the set of discourse referents, as indicated by (29b).[30]

[30] Note that a question-answer pair is one type of discourse structure that can establish a contrastive relationship between elements. A narrative discourse, where the same speaker provides the alternatives set before uttering a contrast-marked sentence, is another. The contrast observed for the *to*-marked

(29) Speaker A: 'Did Geeta study Maths and Science?' (Explicit Alt. Set)
Speaker B: 'Did Geeta study both the subjects?' (Implicit Alt. Set)

(a) gita=ne mæt$^{h}$s=to pəɽ$^{h}$a
Geeta=ERG Maths=TO read.PFV.MSG
'Geeta has studied Maths.'

Assertion: Geeta studied Maths.
Implicature 1: The speaker doesn't know whether Geeta has studied Science or not.
Implicature 2: Geeta did not study Science.

(b) Set of alternatives={Maths, Science}
Let predicate P be such that P(Geeta,x) = study(Geeta,x)

P(Geeta,Maths) = true (from Assertion)
P(Geeta,Science) = unknown (from Implicature 1)
P(Geeta, Science) =false (from Implicature 2)

A *to*-marked object NP, thus, receives a contrastive discursive interpretation. This contrast is the non-topical interpretation that a *to*-marked entity can receive. Particle -*to* is not a focus marker (in the Alternative Semantics sense) because a *to*-marked entity does not indicate the exhaustification of the set of alternatives (which is a crucial characteristic of the focus-marked entity for its set of focus alternatives). Because of this reason, I disagree with the analysis proposed in Han (1998) that claims that an object NP marked with a morphological topic marker (in Japanese, Korean and Hindi) gives a contrastive focus reading.[31] Example (30) is cited from Han (1998:5), who claims that this sentence comes with an implicature that Ram does not like other fruits.

(30) ram=ne seb=to k$^{h}$a lɪja
Ram=ERG apple=TO eat take.PFV.MSG
'Ram ate the apple, (but not other fruits)'

Han (and Bhatt) do not take into account the observation that a sentence like example (30) generates another implicature— the uncertainty implicature (i.e. the speaker is not certain that Ram ate other fruits)—besides the implicature mentioned in their

object, like in example (29a), is not contingent on the type of discourse that provides the alternatives set.

[31] She credits Rajesh Bhatt for Hindi judgment data (Han 1998:5)

analysis. A contrastive focus reading obligatorily comes with an exhaustivity implicature — only the CF marked constituent holds true of the predicate and no other alternative holds true of the predicate, i.e., CF-marking of a constituent exhausts its alternatives.[32] The presence of an uncertainty implicature indicates that the speaker does not know or believe that alternative entities are exhausted in the discourse. Thus, the uncertainty implicature triggered by a *to*-marked object NP in example (30) cancels the exhaustivity implicature that is obligatorily required for a contrastive focus reading. Thus, an in-situ object NP when marked with particle *-to* gets a contrastive and not a (contrastive) focus interpretation.

Most theories about focus take the focal constituent in an answer to be congruent to the wh-variable in its preceding question (see section 2.1.3). However, a study by Heycock (2008) provides evidence to claim that this congruence between a wh-variable and focus is only heuristic in nature and not a definitional criterion of focus. Kim (2013, 2016) also maintains that relational newness (and not alternative indication related to a preceding wh-variable) should be the relevant definition of focus.[33] Thus, the *to*-marked object NP in example (31) is neither topical (i.e., not designated by the *Tell me about X* test) nor focal (since *to*-marked entities are discourse-given and hence, not relationally new), but rather contrastive in interpretation.

[32] In a QUD-based approach (Lee 2000, Büring 2003, Wagner 2012 etc.), a sentence containing a Contrastive Focus constituent is preceded by a disjunctive alternative question in the discourse tree. A sentence containing a Contrastive Topic constituent is preceded by a conjunctive alternative question. CFs are associated with an exhaustivity implicature because CF answers one of the alternatives from the preceding question. Since one disjunct is answered, this satisfies the alternative question and the other disjunct is exhausted. CTs are associated with an anti-exhaustivity implicature because a CT answers only one conjunct from the conjunctive alternative question. The remaining conjunct questions thus remain as open questions in the discourse. This is the traditional analysis behind CFs triggering an exhaustivity implicature and CTs triggering an anti-exhaustivity implicature. Wagner (2012) coined the term 'disputability implicature' for the implicature that accompanies a CT because of the unresolved question that is disputably open in the discourse context.

[33] Kim (2016) gives this relational newness as the argument for claiming that Korean particle *-(n)un* cannot mark contrastive focus, since nun-marked constituents are not relationally new. For an expanded discussion on this topic, see Kim (2013, 2016).

Example (31), *prima facie*, could pose as a counter-evidence to my claim of in-situ object NPs being contrastive and not contrastive focal constituents.[34]

(31) Context: A group of friends are talking about what cities their partners (who are not physically present) have visited.

Speaker A: Tell me about John. Which cities did he see?

Speaker B: ʊs=ne dɪllɪ=to dekʰɪ hʊɪ hɛ
he=ERG delhi=TO see PFV.FSG AUX.PRS.3SG
Assertion: 'He has seen Delhi.'[35]
Implicature: I am uncertain about other cities.[36]

Kim's (2013) definition of contrast treats contrast as a primitive notion of IS, which can be combined with a topical (or a focal) constituent to compositionally build-up a contrastive topic (or a contrastive focus) interpretation. A Contrastive Topic (CT) in this framework can be defined as "CT is a topic whose alternatives are evoked in the context and is distinguished from the alternatives with respect to some semantic property P in that some truth value results when P is applied to the topic and the opposite truth value or 'unknown' when applied to the alternatives" (Kim 2013:57).[37]

---

[34] Note that this example also provides support to the claim that in-situ object NPs, when marked with particle *-to*, get a contrastive interpretation that is dissociated from topic interpretation. This example employs a *Tell me about X* test to predesignate a topical constituent (unlike the implicit QUD in section 3.2.1 that was used to diagnose the topical constituent). The topic of speaker B's sentence is 'John', anaphorically referred to by the topical pronoun 'he', as the sentence is 'about' this entity. A contrastive interpretation is accorded to the *to*-marked object NP 'Delhi', which evokes the uncertainty implicature.

[35] 'Delhi' is presumed to be available in Common Ground or it is made available in it at utterance time by the strategy of presupposition accommodation by the speakers of the discourse (see section 2.3.2). Only then can particle *-to* felicitously mark it. A pragmatic constraint on the licensing of this particle requires that this particle can only mark those entities that are *given* in the Common Ground and hence, discourse-old.

[36] It is an observation that in a group of native Hindi linguists (from IIT-Delhi and JNU, New Delhi), they accepted the sentence in (31) but they preferred overtly stating the inference rather than keeping it implicit. They deemed the sentence 'informationally incomplete'. I suggest that this is because of non-exhaustivity of alternatives that a contrastive inference assumes and a perceived violation of Gricean maxims of Quantity and Relevance in this case.

[37] Similarly, a Contrastive Focus can be defined as "CF is a focus whose alternatives are evoked in the context and is distinguished from the alternatives with respect to some semantic property P in that

The advantage of adopting Kim's framework for Hindi is that a contrastive interpretation has no overlap with a topical interpretation and it provides tools to compositionally build up a CT interpretation (contingent on the availability of alternative in the discourse context). This machinery is relevant for the analysis being developed here since the conclusion reached so far in this chapter is that a *to*-marked non-topical but topic-type XP can have a contrastive interpretation. This contrastive interpretation is compositionally more basic than a contrastive topic interpretation yielded when *-to* marks a sentence-initial topical constituent that evokes a set of alternatives (which then results in a differing truth value when the alternative is substituted in the sentence). The *to*-marked subject NP in example (32) can have both a topic or a contrastive topic interpretation, dependent on whether the alternatives set is available, evoked and contrasted with in the context.

(32) Context: Two friends are talking about the whereabouts of one person's child (in case of (i))/children (in case of (ii)). That person says to the other person:

ram=to səkul gəja
ram=TO school go.PFV.MSG
'Ram went to school.'

(i) Topic Interpretation: Ram went to school.
(ii) CT Interpretation: Ram went to school (but I do not know about the other child/ the other child did not go to the school).

The contrast-based analysis proposed in this section for topical type XPs that are not topics can be extended to explain the inferences evoked in all the cases of the non-topical type constituents (Adj, D, Num, Q and P) discussed in section 3.1. Example (3) and (4) involved a *-to* marked adjective and possessive determiner — both of which had a relevant alternative in the discourse context with which the status of the predicate in the sentence was inferentially juxtaposed. Thus, a contrastive discourse relation is implied between two adjectives and two determiners in (3) and (4) respectively. Examples (5) and (6) involved *-to*-marking of a numeral. Numerals are inherently scalar items. By *–to*-marking of a numeral, a contrastive relation is evoked between that numeral and other numerals (the value of which is decided by the speaker's expectation in the discourse context). In case the contrast is with a lower

some truth value results when P is applied to the focus and the opposite truth value or 'unknown' when applied to the alternatives" (Kim 2013:59), where focus is defined by its discourse function of marking a relationally new entity in a sentence.

element on the scale of numerals, then an 'at least' type interpretation is evoked. Similar to the case *–to*-marking of numerals, *-to*-marking of the existential quantifier *koi/kuch* 'some' also evokes an 'at least' type interpretation as seen in example (8). The existential quantifier *koi/kuch* is implied to be in a contrastive relation with the negative quantifier 'no one/nothing', which consequently yields an 'at least' type interpretation. Example (9) involved a *to*-marked postposition. Here also, an alternative is discernable in the context with which the *–to*-marked postposition is juxtaposed with. Thus, in all these cases, *to*-marked non-topical type XPs evoke contrastive interpretative effects.

A potential issue can be raised here that enclitic particle *-to* could have two different flavors — a topical *-to* and a contrastive *-to*— and that these flavors are encoded by different morphemes in the lexicon.[38] This issue arises because the contrastive cases with particle *–to* might seem to have properties similar to that of a distinct contrastive *-wa* proposed by those analyses that assume Japanese topic marker to have two separate *-wa*s (a thematic *-wa* and a contrastive *-wa*) in the lexicon (Heycock 2008, Tomioka 2016). (33) lists out the properties of contrastive *-wa* that differentiate it from thematic *-wa*, cited from Tomioka (2016: 14).

(33) a. $Wa_{contrast}$ typically displays the focus prosody for contrastive focus; an f0-boost and the subsequent f0-compression. No such prosody for its thematic variant.

b. $Wa_{contrast}$ can be attached to a discourse-new item, such as an answer to a wh-question.

c. Unlike $wa_{theme}$, $wa_{contrast}$ need not be in the sentence initial position.

d. While $wa_{theme}$ can only attach to (quasi-) nominal phrases (eg. NP, CP, PP), there is no categorical restriction for $wa_{contrast}$.

Hindi *to*-marked XPs that receive either a contrastive or a contrastive topic interpretation are not obligatorily prosodically marked with an f0 boost-compression contour, as specified in (33a). This prosodic strategy may optionally co-occur with

[38] See section 1.3 and the discussion about the theoretical choice between a simple syntax and complex semantics versus the complex syntax and simple semantics. By positing two different morphemes for particle –to would imply the second type of theoretical approach. This thesis pushes for a single morpheme analysis of particle *–to*.

morphological strategy of *-to*-marking for an XP that gets a contrastive interpretation but it is not a necessary criterion to mark one.[39] Unlike the property in (33b), any *to*-marked entity should be discourse-old and hence *given* in the context. An answer constituent to a wh-question may be discourse-old (in terms of being available in the CG of the discourse participants) but what is new (for the hearer) about it is that when the predicate applies to the answer-constituent, it returns a true truth value. The property specified in example (33c) can be read as a tendency of languages to license topical interpretation in the sentence initial position. This explanation is conceptually different from positing different morphemes that encode different discursive interpretations. The property in example (33d) differentiates between the type of XPs that a thematic *-wa* or a contrastive *-wa* can mark. A thematic topic status is attributed to quasi-nominal XPs because of the way 'themes' are conceptualized —being 'about' some entity. In comparison, contrast is more free in its distribution because contrast is not a discourse role that constituent takes but is defined a discourse relation that exists between different entities. This relation can exist between any constituent type with which an alternatives set can be established. Using Occam's Razor, this thesis puts forth a proposal that does not increase the number of particle *–to*s beyond necessity, since an alternate proposal (salience based analysis) can efficiently account for the linguistic behavior of this particle in the language.

In summary, this section has argued that Hindi particle *-to* can mark a topical type XP and still not yield a topical interpretation. Contrast is proposed to be the non-topical interpretation yielded by *-to*-marking of an in-situ object NP. This contrastive interpretation is contingent upon the availability of a set of alternatives in the discourse context and is instantiated through the mechanism of implicatures that are evoked when a sentence containing a *-to*-marks a non-topical XP is uttered. The next section attempts an analysis of these implicatures. Since particle *to*-marked XPs can have simple contrast as a possible, non-topical interpretation, this deduction strengthens the core argument of this thesis that particle *-to* is not a morphological marker of IS notion of *topic*. This particle is only indirectly linked with a topic,

[39] This observation was also stated in the presentation on Indo-Aryan languages in the LISSIM-X Workshop on Information Structure in Solang Valley, 2017.

contrastive or a contrastive topic interpretation —which are further dependent on syntactic, semantic and pragmatic considerations.

## 3.3 Analyzing Implicatures evoked by *-to*-marking of NPs

The overview about implicatures in section 3.2.3.1 had introduced the notion of implicatures as a sentence's interpretation that is inferable from the speaker's utterance beyond what is actually said by them. It had also briefly mentioned the maxims of conversation that are assumed to prescribe the nature of reasonable conversation and which further become the basis from which implicatures arise. This section discusses implicatures in detail — by expanding upon the maxims as laid out by Grice and some reformulations of them; by categorizing implicatures as either conventional or conversational; by showing standard calculation of implicatures from various maxims or their violations. This theoretical machinery is then applied to analyse the implicatures that can be evoked in the case of particle *-to* in Hindi.

The original version of the maxims of conversation (as stated in Grice's seminal work) have been revised and amended by linguists working in semantics and pragmatics to account for the observed linguistic phenomenon. The Gricean maxims are listed in (34) below (cited from Tomioka 2020:774):

(34) a. Quality: Try to make your contribution one that is true. Do not say what you believe to be false. Do not say that for which you lack adequate evidence.

b. Quantity: Make your contribution as informative as is required. Do not make your contribution more informative than is required.

c. Relation: Be relevant.

d. Manner: Be perspicuous. Avoid obscurity of expression Avoid ambiguity. Be brief. And be orderly.

These maxims have been observed to be of unequal status cross-linguistically.[40] Proposals have been made to both condense as well as augment the list of the maxims. Sperber and Wilson (1995, 2004) have made use of only one maxim (i.e., maxim of relevance, in an enriched and more generalized form) in their Relevance Theory. A 'Principle of Informativeness' by Atlas and Levinson (1981) and a politeness strategy by Brown and Levinson (1978) have been proposed as additions for an expansion of the set of maxims. Horn (1984, 2004) has propagated a Neo-Gricean approach in which all the maxims and sub-maxims (apart from quality maxim) have been reduced to two fundamental principles — the Q-principle and the R-principle. These revised principles are stated in (35) below (cited from Tomioka 2020:782, which slightly modifies the ones in Horn 1984:13).

(35) a. The Q-Principle (hearer-based): Make your contribution sufficient. Say as much as you can (given R).

b. The R-Principle (speaker-based): Make your contribution necessary. Say no more than you must (given Q).

The relevance of these principles will become apparent in the discussion of Q-implicatures or scalar implicatures in section 3.3.1 below.

Implicatures are further categorized into two types— conventional implicatures and conversational implicatures. Grice introduced conversational implicatures as, "an implicature, the supposition of which is necessary for maintaining the assumption that the speaker is co-operative" (Grice 1975). These implicatures arise due to the interaction of the semantic content of the proposition uttered by a speaker within a context and the co-operative principle (the condensed form of all maxims, see section 3.2.3). The general mechanism of deriving a conversation implicature is provided in example (36), where 'p' and 'q' are sentence variables and S, the speaker. The steps of derivation are taken from Hara (2006b: 3), who further cites them from Levinson (1983: 113-114).

[40] For example: unless maxim of quality is maintained (i.e., what a speaker says is true), the other maxims are of no consideration.

(36) a. S has said that p.

b. There is no reason to think that S is not observing the maxims, or at least the co-operative principle.

c. In order for S to say that p and be indeed observing the maxims or the co-operative principle, S must think that q.

d. S must know that it is mutual knowledge that q must be supposed if S is to be taken to be co-operating.

e. S has done nothing to stop me, the addressee, thinking that q.

f. Therefore, S intends me to think that q, and in saying that p has implicated q.

An example of a conversational implicature is given in (37).[41] The implicature generated by speaker B's utterance can be derived as: speaker B's utterance does not provide a direct response to speaker A's query regarding the movie. Since speaker B is assumed to be co-operative and hence, observing the maxims of rational conversation, (s) he must think that the movie is boring. Speaker B assumes it as mutual knowledge that a person can sleep while watching a movie only if it is true that the movie is boring enough to induce sleep. Speaker B does not stop speaker A from inferring this implicature from the uttered sentence. Therefore, speaker B actually intends speaker A to think that the movie is boring and thereby, provides an answer to his query indirectly.[42]

(37) Context: Two friends, A and B, are talking about a movie that B had gone to see last night.

[41] Conversational implicatures are further categorized as Generalized Conversational Implicatures (GCI) and Particularized Conversational Implicatures (PCI), where the former does not need some prior contextual information for its calculation while the latter is context-specific and depends on some particular feature that is specific to discourse participants or their situation. Example (37) shows a GCI since it is calculated based on the general assumption that people can sleep only during a not good movie. This example would have exemplified a PCI if this assumption was speaker B specific and not holding across as a general tendency.

[42] Note that there is a violation of maxim of Manner in this example since the speaker B does not give a direct answer to speaker A's query by his actual utterance but provides an indirect reply by his implicated proposition. This violation itself leads to implicature generation under the assumption that the discourse participants are still co-operative and obeying maxims of conversation.

Speaker A: How was the movie? Was it good?
Speaker B: I slept within the first half an hour.

Assertion: The speaker slept within 30 minutes of movie's start.
Implicature: The movie was boring.

In contrast to such implicatures that are calculated based on the truth conditional content of the sentence and a pragmatic principle (the cooperative principle), conventional implicatures are implicatures that are triggered because of the conventional meaning associated with a particular lexical item in a sentence. A canonical example of a lexical item that evokes a conventional implicature is the English connective 'but' (Grice 1975). A sentence having 'but' as a connective implicates a contrast between the two conjoined units, with the second conjunct being 'unexpected' given the first conjunct. This is exhibited in example (38). This 'unexpectedness' between the two conjuncts is implicated by the use of 'but'.[43] Replacing 'but' by its truth-conditionally equivalent connective alternative 'and' makes no change in the asserted part but the implicature disappears. This is taken as evidence for the claim that the implicature is triggered by the lexical item 'but'.

(38) He is from a village but he speaks English fluently.

Assertion: He is from a village. He is fluent in English.
Implicature: A person from a village is not expected to speak English fluently.

Returning to the issue of Information Structure, many linguists have proposed formal accounts that link Japanese particle *-wa* or Korean particle *-(n)un* and the implicatures licensed by the C(ontrastive) T(opic) marking of an entity with them (for Japanese see Hara (2006a, b), Tomioka (2010a, b) etc.; for Korean see Lee (2000, 2006, 2007), Kim (2013, 2018, 2019) etc.). An example is given for Japanese in (39a,b), cited from Hara (2006a: 1). The sentence in example (39b) exhibits a type of implicature that is

[43] The added meaning arising out of the use of 'but' may, *prima facie*, appear like a lexical entailment of this connective. However, lexical entailments can be rejected or denied but conventional implicatures cannot be directly rejected by responding with "No, that's not true." (Tomioka 2020: 774). The unexpectedness meaning triggered by 'but' cannot be denied directly — thus proving its status as conventional implicature.

made available by CT marking of an entity (here 'John') with *-wa* in Japanese.[44] Hara claims that her analysis for Japanese is extendable to Korean CT-marking with *-(n)un*.

(39) a. dare-ga paatii-ni ki-ta-ka?
who-Nom party-Dat come-Past-QP
'Who came to the party?'

b. John-wa ki-ta
John-Top come-Past
'As for John, he came.'

Implicature: It is possible that it is not the case that John and Mary came.
≈ I don't know about others.

Keeping the context same as the preceding example for Japanese, example (40) shows that the same implicature is evoked in Hindi when particle *-to* marks the subject NP. This entity gets a contrastive topic interpretation (because of: (a) sentence-initial entities receiving a default topic interpretation for a 'topic-comment' structure sentence in Hindi; and (b) an alternative to John, i.e. Mary, being available in the context and evoked).[45]

(40) Speaker A: parɪ=mẽ kɔn aja
party=LOC who come.PFV.MSG
'Who came to the party?'

Speaker B: dʒon=to aja
John=TO come.PFV.MSG
'John came'
Implicature: I do not know about others.

Recall that section 3.2.2 had shown that a similar implicature is licensed in Hindi when particle *-to* marks an in-situ object NP (see examples (22), (29), (30)). Based on the reasoning given in section 3.2.1, I had claimed that in-situ object NPs are not

[44] Note that this implicature is not licensed when the subject is marked with '-ga', the nominative marker, in the same sentence. This is shown in in (i) below:

(i) John-ga ki-ta
John-NOM come-Past
'John came.' (complete answer)

[45] Note that an unmarked subject NP in the same example still gets a contrastive topic interpretation based on the criteria listed for getting these interpretations. However, no such implicature is evoked in the unmarked case. This is an example of a contrastive non-implicational topic (see Tomioka 2020).

topical entities for a sentence. The non-topic interpretation that a *to*-marked object NP receives is — pure contrast. Since the implicatures under investigation are evoked in both the cases where particle *-to* marks a contrastive topic subject NP and where it marks a contrastive object NP, I propose that contrast (and not CT) is the necessary criteria for generation of these implicatures, with nothing hinging theoretically on the notion of topic itself (since topicality itself has no contribution in the derivation mechanism of these implicatures). This is a point of difference between the account being proposed for Hindi *-to* in this thesis and the earlier accounts proposed for Japanese/Korean particles *-wa*/*-(n)un*. Other analyses (implicitly or explicitly) assume these particles to be a CT marker and these implicatures to be evoked because of CT marking of a constituent.[46] Section 3.3.1 proposes a derivation mechanism for the uncertainty and scalar implicatures obtained by *-to*-marking of an object NP in Hindi. Section 3.3.2 unifies the derivation of both these implicatures in a single procedure, based on the influential proposal by Sauerland (2004).

Grice has proposed a set of diagnostic properties for conversational implicatures that distinguishes them from conventional implicatures. These features are described in section 3.3.3. In the pragmatic literature on Japanese and Korean, the uncertainty implicature and the scalar implicature have been analysed either as conventional implicatures (i.e., conveyed conventionally by *-(n)un*/*-wa*) or as conversational implicatures (i.e., arising out of the semantic content of the sentence containing *-(n)un*/*-wa*-marked phrase and the application/violation of maxims according to the context).[47] This section concludes by evaluating the status of implicatures evoked in the case of particle *-to* in Hindi.

[46] Kim (2018, 2019) is an exception who makes a similar claim for separating CT meaning as an inherent meaning of Korean particle *-(n)un*.

[47] For example, Hara (2006 a,b) analyses Japanese *-wa* as conventionally encoding the CT meaning. Her analysis uses von Stechow's Structured Meaning Approach. A *-wa*-marked phrase presupposes the existence of stronger scalar alternatives (based on asymmetric entailment) and implicate the possibility that it is not the case that the stronger alternative holds true. In contrast, Kim (2018) derives these implicatures using the standard recipe of scalar implicatures involving maxims and the co-operative principle.

### *3.3.1 Deriving Uncertainty and Scalar Implicatures*

Recall that sections 3.2.2 and 3.2.3 have claimed that the *-to*-marking of object NPs evoke (minimally) two types of implicatures — an uncertainty implicature and a scalar implicature. An uncertainty implicature, in simple terms, implies that the speaker is only certain about the contents of the asserted proposition but is uncertain about some other information that might be relevant for the current discourse context. On the other hand, a scalar implicature is a type of quantity implicature that has been defined as "an implicature that is triggered by a violation of (maxim of) Quantity based on the use of an informationally weaker term on an implicational scale" (Blome-Tillman 2013:12).[48] An implicational scale or a Horn scale (from Horn 1972) has been defined as "a set of lexical items that form a linear ordering according to their informational/logical strength" where strength is calculated on the basis of an asymmetric entailment relation. Some of the traditional Horn scales are given in example (41).

(41) a. <all, most, some>
b. <and, or>
c. <always, often, sometime>

Since this implicature has been established to be linked to the core semantic notions of quantification as well as logical connectives, this is also called as Q-implicature (this terminology being attributed to Horn 1984, 2004). Q-implicatures are traditionally categorized as a type of conversational implicatures as they arise out of the interaction of the semantic content of the sentence uttered and the assumption that the discourse participants are co-operative, obeying the maxim of Quantity. Geurts (2010) has provided a 'standard recipe' for generation of Q-implicatures. This is outlined in (42) below (cited from Tomioka 2018: 24):

(42) a. The speaker S says φ.

b. S could have made a stronger and/or more informative claim by saying ψ.

[48] The term 'quantity implicatures' also includes non-scalar implicatures that do not involve any pre-established scales but make reference to the quantity of information that is conveyed by an utterance (see footnote 9 on Tomioka 2020:778).

c. The reason for S's not saying ψ may well be that S fails to believe that ψ is true.

d. Assuming S is knowledgeable or has a strong opinion about the truth/falsity of ψ, one can conclude that S believes that ψ is false.

An example of a scalar implicature involving a quantifier 'some' is given in (43).

(43) Assertion : Some students came to school.
Implicature: Not all students came to school.

The derivation of this implicature, based on the criteria of asymmetric entailment as informational/logical strength, is explicated in example (44). In the semantic scale <all, some>, the quantifier 'all' is an informationally stronger item than the quantifier 'some' because in a sentence like 'x men came' (where x is a place-holder for the quantifier), 'all men came' entails 'some man came' but 'some man came' does not entail 'all men came'. Since scalar implicatures arise when the speaker believes the sentence containing a stronger alternative of a scalar item to be false (see point (42d) in Geurt's recipe), the implicature takes the form of the negation of the sentence containing the stronger alternative. Thus, a sentence containing 'some x' would evoke a scalar implicature that involves 'not all x', where x is the domain of entities over which the quantifier quantifies. This derivation is shown in (44).

(44) (i) Let proposition P : All students came to school
proposition Q : Some students came to school

(ii) P entails Q (since whenever P is true, Q is true too)
Q does not entail P (since when Q is true, P can be either true or false)

(iii) => scalar item in P ('all') is informationally stronger than scalar item in Q ('some') i.e., 'all' > 'some' in a scale <all, some>

(iv) Q implicates that it is not the case that P is true.
i.e., 'Some students came to school' implicates 'Not all students came to school.'

The semantic scales based on entailment listed above are one type of *poset* (partially ordered set) relations (Ward and Birner 2001: 4).[49] There can be semantic scales that involve other poset relations like part/whole, type/subtype etc. Besides these, a context-dependent pragmatic scale (Hirschberg 1985, Matsumoto 1995) can also evoke scalar implicatures. Example (45) is cited from Lee (2007: 161). In the context of this example, a pragmatic scale is assumed to be licensed by the relative prestige value that is associated with each Beatles' member's autograph. Uttering a sentence containing a lower ranked element ('Harrison') from the pragmatic scale evokes a scalar implicature involving negation of a higher ranked element ('Lennon') of the scale. This implicature is easily cancelled in a discourse-context where such a pragmatic scale is not available, maybe where the speaker holding all member's autographs at the same level of prestige.

(45) Q: Which Beatles' autograph do you have?
A: George Harrison's.

Implicature: I do not have Lennon's.
Autographic Prestige Scale: Starr <Harrison<{Lennon, McCartney}

Returning to the issue of implicatures triggered by Hindi particle *-to*, let us consider example (46). In the given discourse context, uttering a sentence with a *to*-marked object NP in example (46a) generates the implicatures as specified in (46b) and (46c). The sentence in (46a) asserts that Ravi has completed reading one book (out of the set of two books available in the context) and marks this object NP with particle *-to*.

(46) Context: Ravi was told to read two books — *The Search* and *The Calhouns*— on a weekend to be able to get a gift. On Sunday night, his sister tells their parents (who are aware of the conditions required to receive the gift) that:

(a) Assertion: rəvɪ=ne *The Search*=to pəɽʰ lɪ
Ravi=ERG The Search=TO read take.PFV.MSG
'Ravi has read *The Search*.'

(b) Uncertainty Implicature: I do not know whether he read *The Calhouns*.
(c) Scalar Implicature: He did not read *The Calhouns*.

[49] Elements in a partially ordered set (poset) relationship can have values ranked in three ways with respect to each other: an element A can have lower value than B, A can have higher value than B, or the "two can be of equal rank, or "alternate values" sharing a common higher or lower value but not ordered with respect to each other" (Ward and Birner 2001:4).

Deriving the uncertainty implicature

An uncertainty implicature, as stated in example (46b) is evoked because a co-operative speaker is assumed to obey both the Q-principle (*say as much as you can*) and the R-principle (*say no more than you must*). Since the context marks information regarding both the books as relevant, the speaker is obligated to provide information regarding both. But since the speaker only knows, and thus can assert with certainty, the status of one book – he implicates his uncertainty regarding the other book.

Deriving the scalar implicature

In the given context, a semantic scale is formed between the elements (*The Search, The Calhouns*) that enter into a poset relationship. The relevant scale for implicature calculation, based on the informative value of the set elements, is outlined in (47). Both the books (as atomic elements) have equal value on the scale. The atomic elements themselves are unranked in the scale with respect to each other. However, the scalar item containing both the books together has a higher informational value. This is because a sentence where the predicate holds true of both the books would asymmetrically entail a sentence where the predicate holds true of either of the books. Thus, both the books together have more informational strength as a scalar item than each book separate.

(47) Scale: {*The Search*}, {*The Calhouns*} < {*The Search* AND *The Calhouns*}

A sentence containing a lower valued scalar item evokes an implicature such that it is not the case that the sentence containing the higher valued scalar item is true. This mechanism is outlined step-by-step in (48) below. Thus, asserting the sentence in example (46a) leads to evoking the scalar implicature in example (46c).

(48) (i) Let proposition P : I read x.
proposition Q : I read y.
proposition R: I read x and y.

(ii) R entails P (whenever x and y both are read, x is read)
R entails Q (whenever x and y both are read, y is read)
P does not entail R (reading x does not entail reading both x and y)
Q does not entail R (reading y does not entail reading both x and y)

(iii) => On a scale, R i.e. (read (x AND y)) is informationally stronger than P i.e. (read(x)) and Q i.e. (read (y))
=> scale: x, y < x AND y

(iv) P implicates that it is not the case that R is true
~ 'I read x' implicates that 'It is not the case that I read x AND y'[50]
~ 'I read x' implicates that 'It is not the case that I read x OR It is not the case that I read y'[51]
~'I read x' implicates that 'It is not the case that I read y'

The derivation of a scalar implicature hinges on the licensing of an implicational scale (semantic or pragmatic) between the set elements. The implicature for example (46a) is licensed because an 'inclusion' relation based poset forms a semantic scale, as specified in (47). However, there are contexts where no such implicational scale exists between alternative entities in the discourse. In such contexts, contrastive marking of an alternative entity in the sentence does not generate a scalar (or an uncertainty) implicature. Example (49) is cited from Kim (2018: 13, ex 9), who employs this data as an evidence for his claim that the contrastive implicatures evoked by *-(n)un*-marked entities in Korean (or *-wa* in Japanese) are not conventionally conveyed by particle *-(n)un*/*-wa*.[52] Although Kim's boyfriend is contrasted with Lee's boyfriend in example (49c) with respect to the property of buying a ring, no pragmatic scale seems to exist between {Kim's and Lee's boyfriends} and {Kim's boyfriend} in the discourse context. Although a semantic "member-set" relation based poset can be questioned to exist between {Kim's and Lee's boyfriends} and {Kim's boyfriend}, but "Kim does not intend the semantic scale to be evoked by using C(ontrastive) T(opic) in her second utterance… Thus, in the given context, the alternative to Kim's boyfriend can and must be represented as a singleton set {Lee's boyfriend}" (Kim 2018: 17). Any possible contrastive implicature (of scalarity like 'But your boyfriend has bought a ring for you' or uncertainty like 'But I don't know whether your boyfriend bought a ring for you') are "overridden and thus canceled by context" (Kim 2018: 13).

---

[50] This step involves application of standard De Morgan's Law: ¬ [x Λ y] = ¬ [x] V ¬ [y]

(here ¬, Λ, V stand for logical negation operator, conjunction operator and disjunction operator)

[51] At this step, one of the disjuncts is in contradiction with the assertion. Because of this, the other disjunct gets valued true.

[52] This analysis goes against Hara (2006a, b) and Lee (2007, 2008), who argue for contrastive implicatures induced by *-wa*/*-(n)un* to be conventional and not conversational implicatures.

(49) a. Kim: Where did you get that ring? It's beautiful!
b. Lee: My boyfriend bought it for me.
c. Kim: My boyfriend-*nun/wa* has never bought me a ring.

This section had focused on scalar implicatures alone, without giving any source or explanation for uncertainty implicatures that are evoked along with scalar implicatures for a *to*-marked object NP in Hindi, like in example (46). The next section outlines some formal proposals that derive both the uncertainty implicature and the scalar implicature within a single process. Section 3.3.3 explores the status of these implicatures in terms of the distinction between conventional implicatures and conversational implicatures.

### *3.3.2 Unifying uncertainty and scalar implicatures in a single process*

Sauerland's theory of scalar implicatures (2004) takes Gazdar (1979) as its starting point. Gazdar claimed that the computation mechanism of scalar implicatures involved one more step beyond the standard negation of a stronger scalar alternative sentence. Borrowing from Hintikka (1962), he employed an epistemic certainty operator K that indicates that the speaker is certain about the argument of K operator. He reformulated the theory of scalar implicatures such that for a sentence $\varphi$ and its stronger scalar alternative sentence $\psi$, the scalar implicature of a sentence $\varphi$ is the expression $K\neg\psi$ (and not the expression $\neg\ \psi$). This implicature can be paraphrased as 'the speaker is certain that $\psi$ is false.'

Sauerland (2004) proposed that the scalar implicatures are epistemically modalized. The computation of a scalar implicature of an assertion $\varphi$ with a scalar alternative $\psi$ proceeds not in one but in two steps. Following Grice's maxim of Relevance and Quantity, assertion of $\varphi$ leads to an implicature generation of the form "the speaker is not certain whether $\psi$ holds". Sauerland calls this the 'primary' implicature of $\varphi$ that has the form $\neg K\psi$. If the speaker assumes some additional knowledge, then an 'epistemic switch' can provide a Gazdar-type implicature of the form $K\neg\psi$ (i.e., the speaker is certain that $\psi$ does not hold). Sauerland calls this type of implicature the 'secondary' implicature of $\varphi$. The secondary implicatures are available provided that

"K¬$\psi$ is consistent with the conjunction of $\varphi$ and all primary implicatures of $\varphi$" (Sauerland 2004: 383). The two-step mechanism for deriving a scalar implicature is summarized in example (50).

(50) a. Scale: $\varphi < \psi$

b. Assertion : $\varphi$
Step 1 - Primary Implicature : ¬K$\psi$
Step 2 - Secondary Implicature: K¬$\psi$ (after epistemic switch)

Adopting Sauerland's proposal and implementing it for Hindi, the distinct types of implicatures triggered by *-to*-marking of object NPs can be combined in a single computational process. It is evident that the uncertainty implicature (as in example (46b) in the previous section) is a primary implicature of the form ¬K$\psi$. Provided the speaker assumes some additional information about the discourse context, the speaker can make an epistemic switch and compute the secondary implicatures of the K¬$\psi$ (the scalar implicature given in example (46c)).

Primary implicatures are also called as 'weak' Q-implicatures and secondary implicatures are also called as 'strong' Q-implicatures (see Tomioka 2020: 777-79). Weak implicatures have to be combined with an additional assumption —"the speaker S is knowledgeable and has a definite opinion about the truth/falsity of $\psi$, one can further assume that either S believes that $\psi$ is true or S believes that $\psi$ is false" (Geurts 2010)— to be strengthened to yield the strong Q-implicature. Sauerland (2004) calls this the 'Experthood Assumption'; Zimmerman (2000) calls this the 'Authority Assumption'; van Rooij and Schulz (2004) calls this the 'Competence Assumption'. This assumption does not hold true at all the times.

Much work has been done to understand the role of morphological particles like *-wa* and *-(n)un* and their role in licensing implicatures in a sentence containing them. Hara (2006), Schwarz and Shimoyama (2011) and Swada (2012) take a lexicalist perspective and propose that contrastive *-wa* is lexically specified for an 'at least' scalar meaning. This can be seen by *-wa*-marking of an NP in example (51) and *-wa*-

marking of a numeral expression in example (52), data from Tomioka (2020:788-89).[53]

(51) Context: Who passed the exam?

MARI-WA ukar-imasi-ta
Mari-TOP pass-POL-PST
'(At least) Mari passed.'

(52) Context: How many guests were at the dinner?

ZYUUGO-NIN wa i-masi-ta
fifteen-CLF TOP exist-POL-PST
'(At least) fifteen people were there, (as far as I can tell)."

In contrast to this, Kim (2018, 2019) argues against a Hara-style analysis of implicatures being conventionally generated by *-wa*/*-(n)un*. A third route is taken by Tomioka (2010a, 2020), who adopts a 'mixed approach' where both conventional and conversational meanings are involved. Regarding his proposal, Tomioka claims that "instead of assigning an implicature to *-wa* itself as its conventional meaning, this approach proposes a convention applying to *-wa* that leads to the scalar implicature" (Tomioka 2020: 790). He puts forth a hypothesis, called the Total Laissez-faire Hypothesis, that a set of scalar alternatives is generated by contrastive *-wa* and secondly, that the Competence Assumption applies optionally. He argues for the standard recipe of Q-implicatures to be sufficient for generating scalar implicatures, with no conventional meaning added by *-wa*. The convention that, however, can apply to *-wa* is that "Do not apply the Competence Assumption to the stronger alternatives generated by contrastive wa" (Tomioka 2020:790).[54] Because of this convention, the speaker does not make the stronger implicature and the hearer can reason both

[53] While the 'at least' scalar reading is obligatorily evoked by contrastive *-wa*-marking of numerals (like in example(47)), this may be optionally obtained in the other case, where the speaker may be ignorant or lacking confidence about the stronger scalar items.

[54] A 'grammatical theory' of Competence Assumption is proposed in Fox (2007). In this theory, an Exhaustive Operator Exh operates at the level of propositions syntactically. A wa-marked phrase (both topical and contrastive) has a wide-scope because the *-wa* particle functions to escape the introduction of exhaustive meaning at the level of propositions (cf. Tomioka 2010b).

epistemic (speaker's ignorance or their belief regarding falsity of stronger alternative) and non-epistemic (speaker's politeness concerns) reasons for doing so.

### *3.3.3 Conversational versus Conventional Implicatures*

Conversational implicatures were introduced as a distinct category from conventional implicatures in the beginning of this section. Grice (1972) identified certain characteristic features of conversational implicatures that can be used as a diagnostic to check the status of an implicature being conversational. These are given in example (53) and are cited from Tomioka (2020: 776).

(53) a. Calculability: The addressee must be able to follow the series of inferences that derives the implicature.

b. Detachability: Conversational implicatures are not tied to particular linguistic expressions. If a sentence generates an implicature $\psi$, then $\phi'$, a sentence that is distinct from $\phi$ but expresses the same meaning as $\phi$, gives rise to the same implicature $\psi$.

c. Indeterminacy: In a given situation, there may be more than one way to explain why the speaker made an utterance in a way that she did. Thus, there may be a disjunction of multiple possible implicatures associated with such an utterance.

d. Cancellability: Conversational implicatures can be negated without causing a logical contradiction.

I repeat the example (46) as example (54) below, to check the status of the implicatures at issue.

(54) Context: Ravi was told to read two books — *The Search* and *The Calhouns*— on a weekend to be able to get a gift. On Sunday night, his sister tells their parents (who are aware of the conditions required to receive the gift) that:

(a) Assertion: rəvɪ=ne *The Search*=to pəɽʰ lɪ
Ravi=ERG The Search=TO read take.PFV.MSG
'Ravi has read *The Search*.'

(b) Uncertainty Implicature: I do not know whether he read *The Calhouns*.
(c) Scalar Implicature: He did not read *The Calhouns*.

Section 3.3.1 had exhibited that both the uncertainty and the scalar implicature can be calculated by the addressee based on the semantic content encoded by the sentence and the assumption of adherence to or violation of maxims in a discourse context. The assumption of co-operative principle necessitates uncertainty implicature to be evoked and the application of competence assumption licenses the strong scalar implicature to be evoked.

Besides the two implicatures specified in example (54b) and (54c), another implicature can be licensed by uttering the sentence given in example (54a) in the same discourse context — the speaker knows but does not want to tell that Ravi has read the second book too (maybe because she does not want Ravi to get the gift). This implicature is based out of a non-epistemic reason.[55] Since three (or more) implicatures are permissible for the *to*-marked sentence, the implicatures are indeterminate (cf. point (53c)). This diagnostic property is taken to be the reason behind implicatures being reinforceable without being redundant (cf. Levinson 2000).

Cancellability (53(d)) is considered to be the most robust diagnostic to signal a conversational implicatures. Section 3.2.3 had diagnosed the inferences, like those in examples (53c) and (53d), to be implicatures using the defeasibility test. These implicatures can be easily cancelled by modifying the discourse context by addition of 'And he also read *The Calhouns*'/ 'And I know for a fact that he read *The Calhoun* too'.

Since the implicatures are calculable, indeterminate and cancellable, these provide strong evidence to claim that the implicatures evoked by uttering a *to*-marked object NP are conversational implicatures. The fourth criterial property of detachability is difficult to test for since "one needs to choose a paraphrase that has the same level of complexity and length. Otherwise, the Manner maxim kicks in, and a different implicature might be evoked" (Tomioka 2020: 776). To evaluate whether uncertainty and scalarity can be detached from a sentence containing a *to*-marked object NP, one

[55] Non-epistemic inferences were evoked in example (14), section 3.2.2.

needs to analyse the meaning contributed by particle *-to* so that it can be incorporated in the paraphrase on the basis of which the detachability property is evaluated.

The next chapter begins with a discussion the meaning for enclitic particle *-to*. Since this particle has neither a lexical meaning nor a truth-conditional contribution to the meaning of a sentence, our system requires the expansion of term 'meaning' to include non-truth conditional meaning too. This is implemented by invoking a multi-dimensional theory of meanings as proposed by Potts (2005), Gutzmann (2013) etc.

To conclude, this chapter has employed the terminology and tools developed in chapter 2 to demarcate the topical constituents from the non-topical constituents that the particle *–to* can mark. Adjectives, determiners, quantifiers, numerals and postpositions are the non-topical categories that this particle marks and that do not comply with the topic definition nor with the topichood diagnostic tests. After diagnosing in-situ object NPs to be not biased to receive topic interpretation, it was exhibited that a contrastive interpretation, mediated via an uncertainty implicature and a scalar implicature, can be evoked when these constituents are marked with particle *–to*. It was argued that these implicatures are conversational in nature that can be derived based on the cooperative assumption and the meaning conveyed by a sentence containing a *–to*-marked constituent within a discourse context. These implicatures are calculable, cancellable and indeterminate. The concept of contrast and topic are distinct interpretive notions that can compositionally build up a contrastive topic interpretation for a constituent.

This chapter has provided arguments to claim that particle *–to* is not a topic marker *per se* in the language and that neither topic nor contrast are inherently associated with this particle.[56] This particle can mark constituents that receive a topic

[56] A recent paper presentation by Dash *et al*. in FASAL-8 conference, held virtually in March 2021, has also proposed that particle *-to* is not a morphological topic marker or a contrast marker for Hindi. Their analysis is grounded within the Farkas and Bruce (2010) Table Model of Discourse that models the utterances in a context as a stack of at-issue propositions. The end goal in a discourse is to attain a stable state by resolving all the at-issue propositions by adding them to the common ground between the interlocutors. Dash *et al*. propose that this particle is an 'anti-exhaustivity' discourse category

interpretation, a contrastive interpretation or a contrastive topic interpretation at the interface. Thus, if a unified analysis of particle *-to* is on track, then the meaning of particle *-to* should be such that these discursive interpretations can be obtained by interaction of the meaning of this particle with the syntactic and semantic-pragmatic factors. The next chapter proposes this meaning to be 'increase of salience' of the entity marked by this particle.

marker that signals that the issue being addressed by the proposition containing the *–to*-marked constituent is not resolved in the current utterance. Their proposal is different from the one being proposed in this thesis as their framework and the theoretical assumptions are different. For example, they do not take into account the pragmatic licensing constraint of a constituent being discourse old for it to be marked by particle *-to* being ruling out as a permissible contrastive interpretation of a *–to*-marked object nominal.

# CHAPTER 4

## THE MEANING OF PARTICLE *-to*

The line of enquiry adopted in this thesis proceeded from observing a linguistic phenomenon (morphological marking of an XP by enclitic particle *-to* in Hindi, in chapter 1) towards the interpretative effects made available by it (topicality and contrast, in chapter 2 and chapter 3 respectively), without *a priori* assuming any one linguistic category being encoded by this particle.[1] To summarize the preceding chapters, this particle is argued to be neither a topic marker nor a contrast marker *per se* in the language. The next logical step in the linguistic unpacking of this particle is to analyse the meaning contributed by it for the semantic information contained in a sentence.

One condition that any proposal regarding the meaning of enclitic particle *-to* has to fulfill is that it must account for both the empirical distribution and the interpretative effects obtained by marking of a constituent via this particle.[2] Thus, the proposed

[1] This direction of research is credited to Matić and Wedgwood (2013), and subsequently Kim (2013, 2015 etc.). They critique directly linking an IS category with some grammatical reflex in language (and thereby giving it a 'universal' status), without adequately accounting for the full range of effects exhibited by that linguistic phenomenon. Thus, "if a unifying theoretical entity such as a cross-linguistic category is to be explanatory (or even useful), it should participate in a chain of causal reasoning from the existence of this entity, through any processes it triggers or participates in, and through its interactions with any relevant external factors, to the different effects we see in the data" (Matić and Wedgwood 2013: 137). While Matić and Wedgwood (2013) focus on the assumed-*to*-be-universal IS notion of focus, Kim (2013, 2015) questions the prevalent topic-as-base or contrast-as-base approach to analyzing particle *-nun* in Korean.

[2] Whereas the topic marker analysis of particle *-to* proposed by Kidwai (2000, 2004) does not account for its full empirical distribution (and assumes no topicality-independent contrastive effect), the other analyses outlined in section 1.2 of chapter 1 give a non-unified, polysemous account for this particle — Lakshmi Bai (1977) proposed three separate *-to*s: an emphatic *-to*, a request particle *-to*, a conjunctive *-to*; Montaut (2015) differentiated a topic marker and an argumentative particle functions of discursive *-to*, apart from a separate grammatical word *-to* that has a co-ordinator or a correlative function. More recent proposals like Dash *et al* (2021) or Deo (2021) analyse it in terms of its pragmatic functions

meaning should be, in a way, both 'general' (so as to be able to associate with different categories like nouns, verbs, adverbs, adjectives, numerals, postpositions *etc.*) and 'specific' (so that topicality and contrast can be derived from it). Additionally, it should not be restricted by clause type, since a *to*-marked constituent can occur in declaratives (example (1)), interrogatives (confirmational questions like example (2)) and imperatives (example (3)) as well.

(1) ʊo kɪtab=to laja
he book=TO bring.PFV.MSG
'He brought a book.'

(2) ʊo kɪtab=to laja
he book=TO bring.PFV.MSG
'Did he bring a book?' (with rising intonation)

(3) tʊm kɪtab=to lao
you book=TO bring.IMP
'You bring the book.'

According to Egg (2013), this requirement of a 'general' part and a 'specific' part of the meaning holds true for any uniform semantic treatment of a discourse particle. Discourse particles are particles that can be defined broadly as having "the function of fitting the propositional content of a sentence to the context of speech by giving an utterance its specific 'shade' (Hartmann 1998:60) or alternatively, by imposing restrictions on appropriated contexts for a given utterance" (Zimmerman 2011: 2013). These particles are involved in management of the information in Common Ground in the sense of Krifka (2008). That Hindi enclitic particle *-to* belongs to the category of discourse particles has been mentioned in Deo (2021), Bayer (2020) and Montaut (2015). Therefore, as a corollary, Egg's general condition on meaning of discourse particles can be extended to the meaning of enclitic particle *-to* too.[3]

('anti-exhaustivity' marker for the former and 'marking the strength of the question answered by the prejacent as weak' for the latter) without explicitly outlining the 'meaning' of this particle.

[3] However, Egg's interpretation of 'general' and 'concrete' is slightly different from the one intended above. For him, a semantic analysis of a discourse particle "must be sufficiently *specific* to allow the derivation of the interpretation of concrete examples" and "must be *general* to cover a wide range of concrete usages" (Egg 2013: 1).

To start off, the lexical item *-to* in Hindi can be stated to have no descriptive meaning associated with it. Although Montaut (2015: 3) has sketched an etymological link between this particle and a pronominal basis in Sanskrit for third person (Sanskrit *ta-* for 'that' or 'he'), no lexical meaning has been attributed to it in Hindi. Section 4.1. discusses the notions of truth-conditional meaning and at-issue content and exhibits that the meaning contributed by particle *-to* cannot be explained in these terms. In section 4.2, the concept of discourse salience is introduced and the relation of particle *-to* with the concept of discourse salience is analysed. In section 4.3, I propose that, similar to the semantic treatment of Korean particle *-nun* given by Kim (2013, 2015), the Hindi particle *-to* encodes a use-conditional (Gutzmann 2013, 2015) or an instructional (Portner 2005) meaning within the multidimensional models of meaning (cf. Potts (2007), Gutzmann (2013, 2015) *etc.*). The meaning contributed by this particle is that of signaling that the constituent being marked by it is 'singled out' or 'picked out' from the context for that sentence and as a result, made 'prominent' (von Heusinger and Schumacher 2019) or 'salient' (Kim 2015). Topic and contrast are not the primary meaning-based notions encoded by this particle, but are rather the secondary interpretative effects that can be derived by the interaction of this encoded meaning of *-to* ('convey salience on the constituent it marks from a speaker's perspective') with the discourse context. Section 4.4. diagnoses the status of particle *-to* as a Use Conditional Item (UCI) by testing it for the set of properties that characterize UCIs. This section also locates the type of UCI of particle *-to* within the typology of UCI proposed by Gutzmann (2013, 2015). Section 4.5 concludes and summarizes the chapter.

## 4.1 Truth-Conditional Meaning and At-issue Content

In the Fregean tradition, meaning is calculated based on the criterion of 'truth-conditionality'. Informally, a speaker is said to know the meaning of a sentence if they know in what possible state of affairs that sentence holds true. These are the truth conditions of a sentence and truth-conditional semantics deals with this aspect of meaning. In type theoretic terms, ordinary propositional content is represented as sets of worlds or as functions from worlds to truth-values (semantic type — <s,t>). Truth-conditional meaning content can, thus, be analysed as a proposition that is based on

possible worlds. Gutzmann (2015: 18) calls this the "t-proposition" encoded by a sentence (marked with superscript 't'). Therefore, a given sentence is true in a world provided that the world belongs to the set of worlds denoted by the "t-proposition" of that sentence. The truth-conditional meaning of a sentence 'sky is blue' in type-theoretic terms is shown in example (4a) and (4b).

(4) a. $|| \text{Sky is blue} ||^t = \{w: \text{Sky is blue in } w\}$

b. "Sky is blue" is true in a world w, iff $w \in ||\text{Sky is blue}||^t$

The meaning of a linguistic unit is truth-conditionally relevant if the presence or absence of it affects the truth conditions of the sentence it occurs in. In example (5), the truth conditions of sentence without particle *-to* (example (5a)) are the same as that of a sentence with particle *-to* (example (5b)). These truths conditions are specified in example (6). As both these sentences are truth-conditionally equivalent, it can be concluded that the meaning contributed by enclitic particle *-to* in a sentence is non-truth conditional in nature.[4]

(5) a. adʒ barɪʃ huɪ
today rain happen.PFV.FSG
'It rained today.'

b. adʒ=to barɪʃ huɪ
today=TO rain happen.PFV.FSG
'It rained today.'
'About/As for today, it rained (today).' (*intended*)

(6) a. "Aaj baarish hui" is true,
iff it rained today. [Truth Condition]

b. "Aaj-*to* baarish hui" is true,
iff it rained today. [Truth Condition]

Another meaning-based dichotomy that is proposed in the literature is the 'at-issue' v/s 'non-at-issue' content that can be encoded by a linguistic expression (cf. Potts 2003, 2005, 2007, 2012). The descriptive meaning of an expression has been referred to as the 'at-issue' content (Potts 2003) (or its conceptually equivalent terms — the 'proferred content' (Roberts 1998) or 'what is said' (Grice 1975)). The at-issue

[4] Truth conditional meaning is traditionally attributed to the semantic aspect of meaning whereas non-truth conditional meaning to the pragmatic aspect of meaning (cf. Gazdar 1978).

content is "typically the content that speakers offer as primary and also the content that they are most expecting to have to negotiate with their interlocutors before it is accepted into the common ground" (Potts 2007: 666). In contrast, meaning contributed by non-at-issue content is secondary and is usually not a candidate for negotiation in the discourse. Example (7) shows both an at-issue content and a non-at-issue content (a conventional implicature contributed by 'still') for the semantic interpretation of an English sentence. Apart from conventional implicature triggers like *but, therefore, still, even* etc., Potts extends the non-at-issue meaning to nominal appositives (example (8)), expressive attributive adjectives (example (9)) or the anti-honorifics particles in Japanese (example (10)). (The English examples are mine; the Japanese example (10) is cited from Portner 2007 (who credits it to Potts and Kawahara (2004)), with the anti-honorific particle glossed as 'ANTIHON'):

(7) John is *still* swimming.
At-issue: John is swimming.
Non-at-issue: John was swimming earlier.

(8) John, *a diligent carpenter*, got work in my house today.
At-issue: John got work in my house today.
Non-at-issue: John is a diligent carpenter.

(9) The *damn* dog peed at my doorstep again.
At-issue: The dog peed at my doorstep again.
Non-at-issue: The speaker holds an emotionally charged state of hostility or dislike towards the dog.

(10)

| | | | | |
|---|---|---|---|---|
| John-wa | [Mary-ga | nesugoshi-chimat-ta] | -koto-o | shitteiru |
| John | Mary | oversleep- ANTIHON -PAST | -fact | know |

At-issue: John knows that Mary slept.
Non-at-issue: It sucks that that Mary overslept.

Potts (2012: 2516) holds the view that the at-issue content consists of both truth-conditional and entailed content. Since the meaning contributed by Hindi particle *-to* is not truth-conditional, therefore, by extension, it should be non-at-issue too.[5]

[5] Deo (2021) proposes a formal pragmatic analysis of particle *–to* where she too concludes that the particle *-to* contributes a type of non-at-issue content. In her proposal, this particle "combines with a proposition but adds nothing to its truth-conditional content. Its non-at issue component simply specifies that the $CQ_c$ (which the prejacent answers) is weak— which means that it is informationally weak relative to the context" (Deo 2021: 18), here $CQ_c$ stands for the Current Question in context c. This is formally expressed in (1), where p-proposition variable, w-world variable.

However, Gutzmann maintains that truth conditionality and at-issueness are "orthogonal to each other and do not necessarily draw the same distinctions" (Gutzmann 2015: 269) and thus, should be evaluated separately.[6]

In the case of Hindi, the particle *-to* neither affects the truth conditions of the sentence nor encodes at-issue content, such that the discourse participants can negotiate about it. Therefore, the meaning encoded by this particle is both not truth-conditional and not at-issue for the interpretation of the sentence. This section has highlighted the fact that the meaning of particle *-to* cannot be accounted for by only considering a model of meaning that permits linguistic expressions to encode content at only one dimension. In section 4.2., the concept of discourse salience and its types is introduced and how particle *-to* interacts with it in a context is spelt out.

## 4.2. Particle -to and Discourse Salience

### *4.2.1. What the particle -to actually conveys?*

When a Hindi speaker utters the sentence in (5a), repeated here as (11a), he conveys the neutral proposition that it rained today.

(11) a. adʒ barɪʃ huɪ
today rain happen.PFV.FSG
'It rained today.'

---

(1) $[| =to|]^c = \lambda p\, \lambda w: p \in CQ_c \wedge WEAK_c\ (CQ_c).\ p(w)$

Deo's proposal is orthogonal to the one being developed in this thesis (and neither do I share all her judgments regarding the Hindi data that she bases her proposal on). Therefore, I do not discuss her proposal further here.

[6] Gutzmann explains this caveat by giving a counter-example where a non-at-issue content may be truth conditionally relevant, as in the case of a non-restrictive relative clause. This is exhibited via example (i) below. The non-at-issue content (the non-restrictive relative clause combined with its anchor 'John') can be evaluated for its truth value and is considered to be true iff John taught Linguistics 101 last semester.

(i) John, who taught Linguistics 101 last semester, caught a viral infection recently.
At-issue: John caught a viral infection recently.
Non-at-issue but truth-conditionally relevant: John taught Linguistics 101 last semester.

(11) b. adʒ=to barɪʃ huɪ
today=TO rain happen.PFV.FSG
'It rained today.'
'About/As for today, it rained (today).' (*intended*)

In contrast, a subtle interpretative difference that is perceived when a speaker utters the sentence in (5b), repeated here as (11b), is that the speaker is overtly 'singling out' the *to*-marked constituent and signaling it to be 'salient' or 'prominent'. This salience conveys the idea that the marked constituent is important for the current utterance and possibly, for the upcoming discourse too. Thus, by uttering (11b), the speaker is interpreted to have overtly single out the adverb 'today' and mark it as salient and important for the utterance.[7]

It is crucial to observe that this additional interpretation (of being salient for speaker conveyed by marking an adverb with particle *–to*) is also obtained when this particle marks constituents of different grammatical categories.[8] For expository purposes, this is exhibited for a subset of the possible marking sites through examples (12a), (13a) and (14a), in which particle *-to* marks the sentence-initial subject pronoun, in-situ object nominal and the verb respectively. The speaker singles out and marks salient *vo* in example (12a), *paani* in example (13a) and *laja* in example (14a). These *to*-marked constituents are important for the current sentence from the speaker's perspective. The intuition that the *to*-marked constituent is possibly important for the upcoming discourse too is made explicit by (12b) to (14b) — the possible discourse continuations for the *-to* marked sentences. The succeeding sentence can felicitously contain a pronominal form that refers back to the *to*-marked constituent ('he' in example (12b)) or it can include alternatives that are juxtaposed to the *to*-marked constituent ('milk' in example (13b) or 'boil' in example (14b)).

[7] The *intended* translation of example (5b), repeated here for example (11b), can be understood to convey this interpretation. To foreshadow — the fact that [About/As for X, Y] type of template has been previously used to diagnose sentence topics shows the direct causal link between X being marked as salient for a sentence and it subsequently being accorded a topical interpretation.

[8] This is relevant for the first part of Egg's (2013) requirement — the meaning of a discourse particle (for this thesis, particle *-to*) should be *general* to account for the wide range of its distribution.

(12) Context: A group of friends had gone for a trek. Most had forgotten to bring a water bottle along. Person A says about person B:

a. [**ʊo**]=to pani laja hɛ
he=TO water bring.PFV.MSG AUX.PRS.3SG
'He has brought water.'
'About/As for him, he has brought water.' (*intended*)

b. After all, **he** is the most sensible of all.

(13) Context: A group of friends had gone for a countryside picnic, where they had planned to make tea. Everyone was assigned to bring something. Person A says about person B (who was assigned to bring both water and milk):

a. ʊo [**pani**]=to laja hɛ
he water=TO bring.PFV.MSG AUX.PRS.3SG
'He has brought water.'
'As for water, he has brought water.' (*intended*)

b. He did not bring the **milk**.

(14) Context: Person A was asked to heat a glass of water and bring it to person B. When asked about how he performed this task, person B says:

a. ʊo pani **laja**=to hɛ
he water bring.PFV.MSG=TO AUX.PRS.3SG
'He has brought water.'
'As for having brought (water), he has brought water.' (*intended*)

b. But he did not **boil** the water.

The above examples may pose a *prima facie* problem to the central claim of this thesis since the discourse continuations listed in examples (12)-(14) are pragmatically felicitous even if the particle *-to* is assumed to be either a topic marker (that establishes the basis for topic chaining in example (12a) and (12b)) or a contrast marker (that licenses the alternatives to the *to*-marked entity made explicit in examples (13) and (14)). However, I take this data as evidence to push for a unified semantic account for particle *-to* — i.e., a common meaning core from which the information structural interpretative effects of topicality and contrast can be derived.

It is evident from the uniform interpretations observed for the *to*-marked sentences in examples (12)-(14), in comparison to their unmarked counterparts, that by using the particle *-to* the speaker is singling out a constituent as salient and marking it as

important for the current sentence. Since salience is a "cognitive, mnemonic attribute (rather than a linguistic one)" (Falk 2014: 3), an analysis for the interpretation conveyed by a *to*-marked constituent can be proposed in terms of a discourse model that involves the interaction of language with other cognitive modules like attention.

### *4.2.2. Discourse Models- A Background*

In literature, the psycho-cognitive effect of the use of a linguistic expression and its reflex in language has been modelled in the theories of 'referential activation' (Chafe 1976, Lambrecht 1994), 'attention' (Grosz et al 1995), 'accessibility' (Ariel 1990), 'salience' (Chiarcos *et al* 2011, Falk 2014), 'givenness' (Gundel *et al* 2013) etc. A common line of enquiry amongst them deals with how the form of a linguistic entity conventionally signals how activated it is in the discourse participants' mental models or how much attention it has at the utterance time or how accessible it is for retrieval from memory. To elaborate one influential proposal, Gundel *et al*'s (1993) Givenness Hierarchy posited six distinct, implicationally related cognitive statuses, where each status can be identified as a "necessary condition on a different type of pronominal or determiner form" (Hedberg and Zacharski 2007: 3). Each cognitive status provides the condition for appropriate use of a pronominal form in a language. The six cognitive statuses and their hierarchical order is provided in Table 5.1, along with the distribution of English pronominal forms that correspond to each cognitive status.[9]

| In Focus > | Activated> | Familiar> | Uniquely > Identifiable | Referential > | Type Identifiable |
|---|---|---|---|---|---|
| she | this N | that N | the N | this N | a N |
| | that N | | | | |
| | this | | | | |
| | SHE | | | | |

Table 4.1: Givenness Hierarchy and English Pronominals

[9] Table 5.1 is sourced from Mulkern (2007: 114), who adopts the Giveness Hiercharchy model to account for the context appropriate use of the various Irish pronominals.

According to this framework, the cognitive status of a *to*-marked NP should be minimally 'referential' in hierarchy.[10] An NP that signals its cognitive status as 'type identifiable' cannot be marked with particle *-to*. Example (15) is an infelicitous sentence in Hindi with a type identifiable NP ('a dog') marked by particle *-to*. This sentence becomes contextually felicitous in the absence of particle *-to* from the subject NP. This framework is not relevant for the interpretative difference observed in examples (12)-(14).[11]

(15) Context: A discussion about type of animals and their traits.

| ek | kʊtta=(#to) | ʊəfadar | dʒanʊər | hɛ |
|---|---|---|---|---|
| one | dog=TO | loyal | animal | AUX.PRS.3SG |

'A dog is a loyal animal.'

### *4.2.3. Discourse Salience: An Overview*

I adopt the theoretical framework of Discourse Salience, as elaborated in Chiarcos *et al* (2011), to account for the interpretative impact of enclitic particle *-to* in discourse. The term 'salience' has been defined and operationalized differently in the various sub-fields of linguistics (cf. Boswijk and Coler 2020). For discourse pragmatics, salience has been understood to be a cover term of properties of mental states like 'activation' (Chafe 1976), 'discourse prominence' (Pustet 1997) and 'topicality' (Givon 1983). Chiarcos *et al* (2011) outlined a working definition of salience as "the degree of relative prominence of a unit of information, at a specific point in time, in comparison to the other units of information" (Chiarcos *et al* 2011: 2).[12] This

[10] In chapter 2, the same empirical observation was formulated as a pragmatic licensing constraint that the denotation of the *to*-marked entity must be available in the common ground between interlocutors. By virtue of an entity being mutually accepted by discourse participants (and thus, added to the CG), the referent of that entity is part of shared knowledge.

[11] There are two issues with adopting the Givenness Hierarchy to account for the *to*-marked constituents. Firstly, since this framework is built to account for only pronominal forms, it cannot be employed to account for cognitive status of non-pronominal categories, like adverb, adjective, verb etc., that particle *-to* marks. Secondly, in the case of *to*-marked NPs, it does not provide any insight as to how *-to*-marking dynamically affects the cognitive status of an NP (since each NP, based on the choice of its linguistic form, conveys its allocated cognitive status in the discourse).

[12] This working definition is attributed to be borne out of discussion of a *Salience in Discourse* themed 6th MAD International Workshop (2005) in Germany 2005.

definition has the advantage of extending salience to be a property of not only the entities in discourse but also of discourse segments beyond entities. Chiarcos (2009, 2011) proposes a corpus study-based Mental Salience Framework that formulates salience metrics that can predict contextually adequate realization preferences. This framework views salience to be a notion that facilitates the attraction of attention in discourse. Attention is considered to be an epiphenomenon of salience since the amount of salience that an entity has determines the attention allotted to it (cf. Chafe 1976 "centre of attention").[13] Unlike the static cognitive statuses of Givenness Hierarchy, salience is a dynamic concept in the sense that the relative salience that an entity has changes (or has the potential to change) throughout the discourse.

Within multidimensional models of salience, distinct types of saliences have been segregated, *e.g.* inherent salience versus imposed salience (Mulkern 2007), backward looking salience versus forward looking salience (Chiarcos 2009, 2011), speaker salience versus hearer salience (Falk 2014).[14] Inherent salience of an entity is determined by the preceding discourse and it signals the centrality of the entity or its 'aboutness' with respect to the discourse. Imposed salience relates to "the amount of prominence or foregrounding given to an entity for the purpose of signaling how the speaker intends the hearer to subsequently rank discourse entities relative to one another" (Mulkern 2007:119). Backward looking salience or hearer salience is conceptually similar to inherent salience whereas forward looking salience or speaker salience is similar to imposed salience. From the perspective of a speaker discourse model and a hearer discourse model, Chiarcos (2011) relates hearer salience to the accessibility or the givenness status of an entity and the hearer knows this hearer salient information. Retrieval of this information is easy for the hearer. In contrast, speaker salient information is "speaker-private and relevant, e.g. new for the hearer, not predictable or something the speaker wants to put special emphasis on" (Chiarcos 2011: 107). Speaker salience signals the 'importance' or 'newsworthiness' of an

[13] Building upon this framework, Kim (2015) defines discourse salience as "the cognitive prominence of the meaning of any part of an utterance made by discourse participants, the degree of which is determined by the amount of attention allotted to it" (Kim 2015: 93).

[14] Chiarcos *et al* (2011: 9) claim these salience dichotomies to be conceptually linked to Givon's (1983, 2001) notion of 'anaphoric topicality' and 'cataphoric topicality'.

entity. With backward looking salience, an anaphoric referential entity 'looks backward' towards its antecedent and with forward looking salience, a referential entity signals whether or not it will be picked up in the upcoming discourse.

In a recent paper by von Heusinger and Schumacher (2019), three characterizations of discourse prominence are proposed that are explanatorily sufficient to reconstruct the prevalent prominence-based notions of discourse models within these three definitions. In their work, salience is identified with discourse prominence that is defined as "a relational property that singles out one element from a set of elements of equal type and structure" (von Heusinger and Schumacher 2019: 119).[15] The advantage of this definition is that the selection or singling out of a constituent itself, by virtue of how prominence is defined, marks it as prominent or salient in the discourse. They also propose the term "prominence lending cues" that are those features of language that "boost the prominence value of their respective referent to a certain extent" (von Heusinger and Schumacher 2019: 119).

### *4.2.4. A Salience-based Proposal for particle -to*

I propose that Hindi enclitic particle *-to* is a morphological 'prominence lending cue' that conveys 'speaker salience' for the constituent it marks.[16] This proposal is similar to Kim's (2013, 2015) account for Korean particle *-nun*, who claims that this particle signals 'imposed salience' on the constituent it marks. Also, von Heusinger and Schumacher (2019) claim that Japanese topic marker *-wa* or syntactic operations like topicalization are both prominence lending cues that raise an entity to a higher prominence status. In their approach, these cues perform the 'forward-looking' function of a forward-looking or speaker salient entity. However, this similarity in the

[15] Besides this 'singling-out' definition of prominence, they propose two more definitions ('dynamicity' and 'structural attraction') that characterize what discourse prominence is.

[16] I choose 'speaker salience' over its conceptually similar variants 'forward looking salience' or 'imposed salience' since the interpretative difference observed for *-to*-marking of a constituent (see section 4.2.1.) is functionally closest to how Chiarcos (2011) conceptualizes speaker salience. A *to*-marked constituent has forward looking salient information and has imposed salience but these are not the first intuitive interpretations obtained when a speaker marks an entity with *-to*. However, these overlapping terms are used interchangeably in the literature and in this thesis.

proposals for Hindi *-to*, Korean *-nun* or Japanese *-wa* particles is not surprising but rather, indicative of a universal tendency of how languages employ the lexical resources available to them to mark discourse salience or prominence of a constituent. This further motivates a cross-linguistic (re-) investigation of the assumed-to-be 'direct' link between such particles and topicality since they all appear to be mediated by manipulation of the degree of salience in discourse.

In example (16), the particle *-to* marks the adjective 'red'. According to the claim made above, the speaker of the sentence singles out the adjective and marks it as salient or prominent. This salience is effected by virtue of the 'singling-out' characterization of prominence outlined in von Heusinger and Schumacher (2019). This is speaker-salience since it is speaker-private and it marks the constituent that speaker considers to be important for the current utterance and wants to put emphasis on (cf. Chiarcos 2011).

(16) Context: Rahul was given a green apple and a red apple to eat.

| | | | | |
|---|---|---|---|---|
| rahʊl=ne | lal=to | seb | kʰa | lija |
| Rahul=ERG | red=TO | apple | eat | take.PFV.MSG |

'Rahul had eaten the red apple.'

Mulkern (2007) proposes that 'contrast' and 'emphasis' are the two motivations for imposing salience on an entity. A speaker marks an entity for imposed salience or forward looking salience so as to either establish a contrastive relation with some other entity in the upcoming discourse or to emphasize an entity to be the most salient one so that it can be established as the new topic of discourse (Mulkern 2007: 123). Therefore, the information structural notions of 'topic' and 'contrast' can be analysed as the interpretative effects that are a consequence of marking an entity with imposed salience.

In example (16), salience is imposed on the adjective 'red' by singling it out and marking it with particle *-to* in the sentence. This constituent is important for the current sentence for the speaker's mental model of discourse. Like other examples in section 4.2.1., a discourse continuation of the example (16) with 'But he did not eat the green apple' exhibits the forward looking function of the *to*-marked constituent.

To summarise this section, the Hindi enclitic particle *-to* signals the speaker salience on the constituent it marks for the interpretation of that sentence. This forward looking salience could either mark a constituent as the topic for the upcoming discourse or mark it to be in a contrastive relationship with some constituent in the upcoming discourse. The next section introduces those models of meaning that can formally incorporate the meaning of particle *-to* in terms of its relation with discourse salience.

### 4.3. Multi-dimensional Models of Meaning and Particle *-to*

In this section, I propose that the meaning encoded by enclitic particle *-to* in Hindi is to *convey speaker salience* for the constituent it marks. This proposed meaning cannot be analysed within a unidimensional model of meaning in which the meaning of a linguistic expression can manifest on only the descriptive dimension. Multidimensional semantics (cf. Kaplan 1999, Kratzer 1999, Potts 2005, Gutzmann 2015) postulate a distinct dimension of meaning —the 'expressive' or the 'use-conditional' dimension—at which a linguistic expression can contribute a conventional non-truth conditional meaning. Within this framework, a linguistic expression can contribute either a truth-conditional meaning or a non-truth conditional meaning or both. The full meaning of a sentence is modeled as a tuple or a pair of truth-conditional (T) and expressive meaning (C) i.e, <A,C>.[17] Relevance Theorists (cf. Sperber and Wislon 1995, Bezuidenhout 2004) also propose a similar dichotomy in the type of the meaning a lexical item can encode— a conceptual meaning (that encodes concepts) versus a procedural meaning (that encodes procedures for interpretation).

In the following sub-sections, these various theoretical models areoutlined and how the meaning of particle *-to* can be accounted for in each framework is subsequently

[17] For Potts, 'C' is the 'set' of expressive meanings since a sentence can encode more than one expressive meaning and they all combine in a non-complex way without interacting with each other *i.e.*, as members of a set.

expressed. The meaning encoded by particle *-to* is formulated as a kind of procedural meaning (section 4.3.1), an expressive meaning (section 4.3.2) and a use-conditional meaning (section 4.3.3).

### *4.3.1. Procedural Meaning*

Relevance Theory (Sperber and Wilson (1997, 2002), Blakemore (2002), Bezuidenhout (2004) etc.) is a cognitive-pragmatic theory that takes Grice's maxim of relevance as its basis. This theory claims that "the expectations of relevance raised by an utterance are precise enough, and predictable enough, to guide the hearer towards the speaker's meaning" (Wilson and Sperber 2002: 250). Verbal comprehension and communication proceeds by a hearer decoding both the linguistic meaning encoded by an utterance and any other evidence of speaker's intention that can yield the interpretation of the speaker's meaning via an inference process.

Within this tradition, lexical items can encode a 'conceptual' meaning or a 'procedural' meaning or both. Common nouns, verbs, adjectives, adverbs and prepositions are the lexical items that encode a conceptual meaning. A lexical item with a conceptual meaning can 'potentially' add to the truth-conditional meaning of an utterance that contains it.[18] A lexical item with a procedural meaning encodes procedures or "instructions that constrain the inferential phase of verbal communication" (Bezuidenhout 2004: 1). For example, the lexical item 'but' encodes a procedure for interpretation — it signals either a contrast or a denial of expectation. By the use of such procedures, a speaker indicates some information about the context of utterance. These procedures guide the hearer towards an intended contextual effect, which results in a reduction of the overall effort needed to process the discourse. Bezuidenhout (2004: 3) lists the three contextual effects that the speaker guides the hearer towards via these procedural items:

[18] Bezuidenhout (2004: 2) gives the following counter-example to illustrate that a sentential adverb like 'sadly', although it has some conceptual meaning, is not relevant for the truth-conditional content of the sentence.

(i) Sadly, John's mother died last night.

| | |
|---|---|
| Contextual Effect 1: | contextual implication (those inferences that follow from contextual assumptions and the content of an utterance, but not from either of them individually) |
| Contextual Effect 2: | strengthening of an existing assumption |
| Contextual Effect 3: | contradiction of an existing assumption |

In the terminology of this framework, Hindi enclitic particle *-to* has a procedural meaning. A speaker marks a constituent with this particle to guide the hearer towards a contextual implication regarding that *to*-marked constituent. The procedure encoded by particle *-to* is to signal speaker salience on the constituent it marks. This imposed salience facilitates the hearer to infer that the *to*-marked constituent is important for the current utterance and possibly the upcoming discourse too. This inference results in a reduction of the processing effort on the hearer's part since the hearer is already expecting the *to*-marked constituent to be relevant for the succeeding sentence too.

### *4.3.2. Expressive Meaning*

Building upon Kaplan (1999), Potts (2005, 2007) develops a theoretical account of expressive meanings that is later expanded by Gutzmann (2013, 2015) in his theory of use conditional meaning. 'Expressives' in the narrow sense are defined as "expressions that express some emotional and evaluative attitude with a high degree of affectedness" (Gutzmann 2013: 4). Typical examples of expressives include attributive adjectives (cf. example (9) in section 4.1) and pejorative epithets like 'bastard' in example (17).[19] Lexical items that have both a truth-conditional meaning component and an expressive meaning component (that mostly conveys a negative attitude by the speaker towards the proposition) are called 'expressively coloured expressions' (Gutzmann 2013). An ethnic slur, like *chinki* (specific to Indian context) that is a derogatory term for a person of north-east Indian origin, is an example of an expressively coloured expression in (18).

[19] Expressive content is not limited to lexical items and can be contributed by a particular intonation or syntactic structure as well.

(17) That bastard John triggered the fire alarm again.

(18) ʊo=to ek ʧɪnkɪ hɛ
she=TO one chinki AUX.PRS.3SG
'She is a chinki.'

Descriptive: That person is of north-east Indian origin.
Expressive: The speaker holds a negative attitude towards people from north-east India.

Based on the claim by Potts and Kawahara (2004) that expressive meanings are performatives, Portner (2007) introduces the term 'instructional meaning', as a kind of expressive meaning, to account for vocatives, sentence topics and force markers.[20] Linguistics expressions with an instructional meaning encode "instructions for interpretation" of that sentence within the context (Portner 2007:1). For sentence topics, Portner proposes the expressive meaning given in (19a). This meaning can be formally represented as a function from contexts to propositions (given in (19b)), for a topical constituent 'x' in a context 'c' in a world 'w'.[21] The context c is a pair <speaker, world>.

(19) a. "(I report that) my mental representation of x is active."

b. $[|\text{ TOP }|]_c^C = [\lambda x\, \lambda c\, \lambda w.$ speaker(c)'s mental representation of x is active in w]

The instructional meaning of a sentence topic includes the speaker's (and not the addressees') mental activation status because the interpretation of a topic is constant across a matrix clause (that has both a speaker and an addressee) and an embedded clause (has only a speaker (the deictic center of the clause) and no addressee).[22]

[20] Portner (2007: 7) explains that expressive meaning, by being categorized as performatives, imply that "they are automatically true once understood".

[21] In Portner's semantic analyses, interpretation function for regular content is represented as $[|\ |]_c$ and interpretation function for expressive content is represented as $[|\ |]_c^C$ .

[22] Contrary to the general opinion that topics are a root phenomenon, Portner suggests that topical interpretation is available for clauses embedded under attitude verbs like 'think', 'say' etc. The deictic center for root clauses is the speaker of the sentence and the deictic center of embedded clauses under attitude verbs is the attitude holder. In example (i), the mental representation of Maria can be active for either the speaker of the proposition or the deictic center of embedded clause–John.

(i) John said that, as for Maria, she is nice.

Portner notes that the addressee's mental representation of x (the sentence topic) is activated as an indirect effect or as a perlocutionary act of the speaker conveying that his mental representation of x is active (Portner 2007: 12).

Within this framework, the particle *-to* can be analysed as encoding an instructional meaning given in (20a). The speaker, by morphologically marking a constituent *x* with enclitic particle *-to*, conveys an instruction for the interpretation of the proposition *p* that contains '*x-to*'— that *x* has speaker salience in *p*. The expressive content of this particle can be formalised as (20b). In a context *c* in a possible world *w*, constituent *x* in the proposition *p* is marked for speaker salience by particle *-to*.

(20) a. "Conveys speaker salience on *x* in *p*"

b. $[|\ \textit{-to}\ |]_c^C$ = [λx λc λp λw. Conveys speaker(c) salience on x in p in w]

This meaning is formalised relative to a context variable *c* (that is a pair of <speaker, world>) to account for the empirical fact that enclitic particle *-to* can occur in embedded clauses too, since the embedded context and the root context vary. In example (21), particle *-to* marks the object NP *banana* in the embedded clause under the attitude verb 'say'. The speaker who imposes salience on this entity can be either the deictic center of the root clause (the actual speaker, pivot for the contrastive implicature 1) or the deictic center of the embedded clause (Rahul, pivot for the contrastive implicature 2).

(21) Context: Riya was given two fruits —an apple and a kiwi—to eat.

rahʊl=ne kəha kɪ rɪja=ne kɪʊɪ=to kʰajɪ
Rahul=ERG say.PFV.MSG that Riya=ERG kiwi=TO eat.PFV.FSG
'Rahul said that Riya ate the kiwi.'

Implicature 1: The speaker is uncertain whether Riya ate the apple.
Implicature 2: Rahul is uncertain whether Riya ate the apple.

### *4.3.3 Use conditional Meaning*

Gutzmann (2013, 2015) proposed that the criterion of truth-conditionality (cf. section 4.1.) and the criterion of conventionality (whether something is conventionally

encoded by an utterance or not) makes its permissible for natural language to express four kinds of meaning. This four-way distinction and their instantiations in language is represented in Table 4.2. Gutzmann's main claim is that a certain expression, which may be truth-conditionally irrelevant, can "impose conditions on the felicitous use of the sentence in which it occurs" (Gutzmann 2015: 7).[23] Such expressions that affect the use conditions of a sentence encode a 'use-conditional meaning', which is conventionally encoded but non-truth-conditional in nature.[24]

| | + Truth-Conditional | - Truth-Conditional |
|---|---|---|
| + Conventional | *Descriptive Meaning* | *Use-Conditional Meaning* |
| - Conventional | *Pragmatic Enrichment* | *Conversational Implicatures* |

Table 4.2. Conventions v/s Truth Conditions[25]

Parallel to the truth conditions (italicized in example (22a)) that need to be fulfilled for a sentence to be true, use conditions (italicized in example (22b)) are the conditions that need to be met for a sentence to be felicitously used in a context; example from Gutzmann (2015: 16).

(22) a. "Snow is white" is t**rue** iff *snow is white.*

b. "Oops!" is **felicitously used** iff *the speaker observed a minor mishap*.

A Use Conditional Item (UCI) is a linguistic device that conveys "the meaning that does not contribute to the truth conditions of a sentence, but instead, they affect the conditions in which the sentence can felicitously be uttered" (Gutzmann 2013: 33). Therefore, the interjection 'oops' in example (22b) is a type of lexical UCI. Kaplan (1999) and Gutzmann (2013, 2015) argue that use conditions of the UCI can be semantically modelled using the tools of formal semantics. In type theoretic terms,

[23] Gutzmann bases his theory on Kaplan's idea that "For certain expressions of natural language, a correct Semantic Theory would state rules of use rather than something like a concept expressed" (Kaplan 1999: 6).

[24] Gutzmann adopted the term 'use-conditional' for his theory, which was originally proposed by Recanati (2004b).

[25] This table is sourced from Gutzmann (2015: 5 table 1.1).

just as truth-conditional meaning is modeled as a set of possible worlds or as functions from possible world to truth values (cf. section 4.1.), use-conditional meaning is modeled as a set of contexts or as a function from contexts to truth values. Semantic type of context being 'c', the semantic type of use-conditional meaning is <c,t>. The use-conditional meaning can be analysed as a special kind of proposition, which is based on contexts instead of possible worlds, and Gutzmann (2015) labels it as the 'u-proposition' encoded by a sentence. The interpretation function for use-conditional content (marked by superscript 'u') gives the set of contexts in which the use conditions are met. This is represented for UCI 'oops' in example (23), where $c_s$ stands for speaker s in context c and $c_w$ stands for possible world w in context c.

(23) a. $\| \text{oops} \|^u = \{c: c_s \text{ observed a minor mishap in } c_w \}$

b. "Oops" is felicitously used in a context c, iff $c \in \| \text{oops} \|^u$

Hybrid Semantics is a semantic framework proposed by Gutzmann (2015) that assumes that "the entire semantic information an expression encodes lies in its contribution to both the truth and use conditions of a sentence" (Gutzmann 2015: 21). For any expression A, an interpretation function on it yields the pair of truth-conditional content (T-Content or TC) of A and the use-conditional content (U-Content or UC) of A. This is schematically represented in (24a) and exhibited for a sentence in (24b).

(24) a. $\| A \| = < TC(A), UC(A) >$

b. || That damn co-passenger vomited again || =
< {w: The co-passenger vomited again in w},
{c: $c_s$ dislikes the co-passenger in $c_w$ } >

Within this framework, the Hindi enclitic particle *-to* contributes meaning at the use-conditional dimension because it does not affect the truth conditions of a sentence (cf. section 4.1) but affects the use conditions for the felicitous use of a sentence in a context. Examples (25)-(26) exhibits the use conditions that needs to be fulfilled for particle *-to* to felicitously mark a constituent in an utterance. In the first sentence in example (25), the denotation of Rahul is marked salient by virtue of it being the sentence topic from the speaker's perspective (topics are discourse prominent *a la* von Heusinger and Schumacher 2019). The particle *-to* attempts to convey speaker

salience on an already speaker-salient entity in the second sentence on example (25). Thus, this speaker salience redundancy leads to pragmatic infelicity of the sentence in the context. Without this particle, the sentence is felicitous in the given context.

(25) Context: Two friends are talking about their third friend. One tells the other that:

rahʊl=ne pudına kʰərida.
Rahul=ERG mint buy.PFV.MSG
'Rahul bought mint leaves.'

ʊs=ne=(#to) ʊs=kı ʧəʈnı bənajı
he=ERG=(#TO) that=GEN chutney make.PFV.FSG
'He made chutney out of it.'

Example (26) exhibits the use condition that the conceptual meaning encoded by the constituent that *-to* marks (in this example, denotation of the proper noun 'Braveheart') must be hearer-known.[26] Since the context specifies that the hearer knows only one movie, conveying speaker salience on a movie name that is not mutually known leads to pragmatic infelicity. This was observed as a pragmatic licensing constraint on particle *-to* in chapter 2. The sentence becomes felicitous in the absence of enclitic *-to*-marking.

(26) Context: A person knows the name of only one movie 'Titanic'. He asks his friend whether he has watched Titanic. To this his friend responds:

mɛ=ne titanic nəhı braveheart=(#to) dekʰı hɛ
I=ERG titanic NEG braveheart=(#to) see.PFV.FSG AUX.PRS.3SG
'Not Titanic, I have watched Braveheart.'

Combining the individual use conditions highlighted in example (25) and (26), the use condition that needs to be fulfilled for a felicitous use of a *to*-marked constituent x in a sentence is sketched out in (27a). Within the framework of Hybrid Semantics, the interpretation function for the u-content of a *to*-marked constituent x is formalised in (27b) which gives the set of contexts c where the contextual speaker $c_S$ imposes salience on a contextual hearer ($c_H$) known constituent x within the possible world $c_W$.

[26] I use 'conceptual meaning' encoded by the constituent instead of 'referent' of the constituent so as to be able to extend this proposal to non-entity type constituents that *-to* marks like adverb, adjective, prepositions, verbs, nouns etc. All these categories encode a conceptual meaning (cf. section 4.3.1).

(27) a. "... x-*to* …" is felicitously used iff the speaker imposes salience on a hearer-known x

b. || x-*to* ||$^{u}$ = {c: $c_S$ imposes salience on a $c_H$-known x in $c_W$}

The full semantic information of a sentence like example (28a) is a tuple of its t-content (a set of possible worlds where the truth conditions are fulfilled) and u-content (a set of contexts where the use conditions are fulfilled) as shown in example (28b).

(28) a. rahʊl=ne kɪʊɪ=to kʰajɪ
Rahul=ERG kiwi=TO eat.PFV.FSG
'Rahul ate the kiwi.'

b. || Rahul-ne kiwi-*to* khaayi || =
< {w: Rahul ate the kiwi in w},
{c: $c_s$ imposes salience on $c_H$-known kiwi in $c_w$ } >

Informally, the meaning of an expression can be represented as a fraction with the u-content as the numerator and the t-content as the denominator (Gutzmann 2013), as represented for an expression A in example (29). This type of informal fraction representation for the sentence in example (28a) is given in example (30).

(29) $A = \frac{\text{U-content (A)}}{\text{T-content (A)}}$

(30) Rahul-ne kiwi-*to* khaayi = $\frac{speaker\ imposes\ salience\ on\ hearer\text{-}known\ kiwi}{\text{Rahul-ne kiwi khaayi}}$

This section had provided an account of Hindi particle *-to* as encoding a procedural meaning, an expressive or instructional meaning and a use-conditional meaning. The next section further develops the use-conditional analysis of particle *-to* by discussing a typology of UCI based on how they interact with the truth-conditional content and whether the diagnosed type of UCI that particle *-to* encodes qualifies all the characteristics of UCI as outlined in Gutzmann (2013) or not.

## 4.4. Use Conditional Items and Particle *-to*

### *4.4.1. Typology of UCIs*

Gutzmann (2013) proposes a typology of use conditional items based on three criteria — Dimensionality, Functionality and Resource Sensitivity. These three criteria regulate the interaction of the Use Conditional (UC) meaning of a UCI with the Truth-Conditional (TC) meaning of the sentence. Based on these three criteria, five types of distinct UCIs are categorized.

Dimensionality of a UCI indicates the dimension at which it contributes its meaning. A UCI that encodes only UC meaning and thus, is one-dimensional, is called as an 'expletive' UCI.[27] In contrast to this, a UCI that contributes meaning at both the UC and the TC dimensions is labelled as a 'mixed' UCI.[28] Expressive attributive adjectives qualify as expletive UCI as they only encode a UC meaning (example (31a)). Ethnic slurs are mixed UCI since they encode both a UC meaning (derogatory attitude towards an ethnic group) and a TC meaning (neutral, non-racist meaning of an ethnic group's label). The ethnic slur 'Kraut' for Germans is a mixed UCI in example (31b).

(31) a. That *damn* dog bit me.
b. Those *Krauts* won the game again.

The criterion of Functionality segregates UCIs into those that need an argument for their UC meaning to apply (a 'functional' UCI) and those that are sufficient stand-alone and do not need any argument to contribute a UC meaning (an 'isolated' UCI).[29] The former are functions from an argument to use conditions while the latter are not. The expressive adjective 'damn' in example (31a) is a functional UCI since it requires a nominal argument (in this case, 'dog') to contribute the UC meaning of the speaker's negative attitude towards the argument. In contrast, an interjection like 'ouch!' is an isolated UCI as it is functionally saturated and does not need any argument to "directly express the emotion of pain" (Gutzmann 2015:39).

[27] The term expletive UCI is attributed to Cruse (2004: 57).

[28] The term mixed UCI is adopted from McCready (2010).

[29] This binary division of isolated UCI v/s functional UCI id adopted from Potts (2005: 65).

A third criterion proposed by Gutzmann that further classifies functional UCIs is Resource-Sensitivity. Certain functional UCIs that consume their arguments by "shunting" them over in the use-conditional dimension are called as Shunting UCIs (from McCready 2010). An exclamative, like example (32), does not leave its argument unmodified in the TC dimension and instead shunts it over the UC dimension completely. No TC meaning is contributed by an exclamative. In contrast, non-shunting UCIs do not consume their arguments and leave their descriptive meaning unmodified in the TC dimension, which can be reused later in semantic derivation. The functional expletive UCI 'damn' in example (31a) is a non-shunting UCI since it does not shunt its argument to the UC dimension but leaves it unmodified in the TC dimension.

(32) How tall John is!
UC Meaning= It is unexpected how tall John is.
TC Meaning = ∅

These three criteria have been analysed as a binary valued feature system in Gutzmann(2015) — [±f (unctional)], [±2d(imensional)] and [± r(esource)s(sensitivity]. The featural decomposition of the various types of UCI outlined in this section is specified in Table 4.3, which is a slightly modified version of Gutzmann (2015: 40 table 2.1).

| | *f* | *2d* | *rs* |
|---|---|---|---|
| Isolated Expletive UCIs | - | - | |
| Isolated Mixed UCIs | - | + | |
| Functional, Expletive, Non-shunting UCIs | + | - | - |
| Functional, Expletive, Shunting UCI | + | - | + |
| Functional Mixed UCIs | + | + | |

Table 4.3. Types of UCI and their Featural Decomposition

Within the typology outlined in this section, the Hindi enclitic particle *-to* is an expletive UCI since it contributes only UC meaning and no TC meaning (i.e. it is

valued [+2d]); it is a functional UCI because it necessarily needs an argument on which it maps its UC meaning i.e. it conveys speaker salience on some argument. Thus, it is valued [+f] for the functionality feature. Since it leaves its argument unmodified in the descriptive dimension for further semantic derivation and does not 'shunt' it over to the UC dimension, it is a non-shunting UCI that is valued [-rs] for the resource-sensitivity feature. To conclude, the enclitic particle *-to* is a functional, expletive, non-shunting UCI that has the feature composition of [+f, -2d, -rs].

Gutzmann (2013) modified the fraction-based schemata that separated TC-content from UC-content (illustrated in example (29) in section 4.3.3) to reflect the type of UCI. The scheme proposed for expletive functional UCI is represented in (33), where ε stands for the UCI, α is argument of the function denoted by the UCI and S is the sentence. Since numerator includes only the u-content, the UCI with its argument fills this place. The denominator includes only the t-content, so it has the sentence expression without the UCI item (since it does not contribute to the TC-dimension). Such a schematic representation is exhibited for a *to*-marked sentence in Hindi in example (34).

(33) $S[\ldots\varepsilon\ (\alpha)\ldots] = \frac{\varepsilon\ (\alpha)}{S[\ldots\ \alpha\ \ldots]}$

(34) Rahul-ne kiwi-*to* khaayi = $\frac{kiwi-to}{\text{Rahul}-\text{ne kiwi khaayi}}$

### *4.4.2. Characteristics of UCIs*

Adopting and building upon Potts (2007b), Gutzmann proposes a list of characteristics that qualify a UCI and can be used as a diagnostic tool "to check whether an alleged UCI does indeed contribute use-conditional content rather than just rely on our intuition that it does not affect the truth-condition of an utterance" (Gutzmann 2013: 34). These properties are discussed and then the status of enclitic particle *-to* as a UCI is verified using these tools in this section.

1. Independence

This characteristic requires that the UC content "contributes a dimension of meaning that is separate from the regular truth-conditional content" (Gutzmann 2013: 38). This property has been decomposed into further sub-properties:

(i) UC content cannot be negated by ordinary negation.
(ii) UC content cannot be denied directly in dialogue.
(iii) UC content is not part of what is questioned by an interrogative.
(iv) UC content does not affect the descriptive content if not fulfilled.

In example (28a), repeated as example (35) here, the UC content encoded by particle -*to* is that *the speaker imposes salience on the hearer-known kiwi*. This UC meaning is tested for the property of being independent from the TC meaning through examples (36)-(38).

(35) rahʊl=ne kɪʊɪ=to kʰajɪ
Rahul=ERG kiwi=TO eat.PFV.FSG
'Rahul ate the kiwi.'

In example (36), negating the proposition by the negative particle *nahi* negates the truth conditional content of the proposition. The speaker salience being marked on 'kiwi' is not negated in this sentence.

(36) rahʊl=ne kɪʊɪ=to nəhɪ kʰajɪ
Rahul=ERG kiwi=TO NEG eat.PFV.FSG
'Rahul did not eat the kiwi.'

In example (37), denial of a sentence containing a *to*-marked 'kiwi' can deny any of the variables that compositionally contribute truth conditional meaning but the speaker salience of 'kiwi' is not the locus of denial.

(37) Speaker A: rahʊl=ne kɪʊɪ=to kʰajɪ
'Rahul ate the kiwi.'

Speaker B: nəhɪ
'No.'

In the question in example (38), what is being questioned is whether the truth value of the proposition ‘Ravi ate the kiwi’ is true or false and not whether ‘kiwi’ is speaker-salient or not.

(38) rahʊl=ne kɪʊɪ=to kʰajɪ
‘Did Ravi eat the kiwi?’ (with rising intonation)

In a context where the UC meaning of example (35) is not fulfilled i.e., the speaker salience on ‘kiwi’ is not conveyed despite *-to*-marking, the sentence would be infelicitous and not ungrammatical in that context. Therefore, this UC content does not affect descriptive meaning if not fulfilled. Therefore, based on the observations above, enclitic particle *-to* does exhibit the ‘independence’ property characteristic of UCIs.

2. Nondisplaceability

By this property, Potts meant, “expressives predicate something of the utterance situation” (Potts 2007b: 166). UCIs are tied to the utterer, the time of utterance and the place of utterance. They cannot be ‘displaced’ to refer to the things or events or attitudes that do not manifest in the utterance situation. This version of nondisplaceability as property is considered as too strong a constraint by Gutzmann (2013). He claims that for some UCIs, the UC meaning can be interpreted relative to the reported speech or attitude context rather than the speaker of the utterance context. UCIs are spread on a continuum with one end being relatively easy to shift and the other end having UCIs that never displace.

In the case of Hindi, example (21) in section 4.3.2. (repeated here as example (39)) had made this observation that the UC meaning can be displaced to the subject of the embedded context too. Since imposed salience on a constituent motivates a contrastive implicature in this context, the deictic centers of these implicatures indicate that the UC meaning could shift to the reported speech context for Hindi particle *-to*. Therefore, particle *-to* is not strictly nondisplaceable but can be displaced under appropriate contexts.

(39) Context: Riya was given two fruits —an apple and a kiwi—to eat.

rahʊl=ne kəha kɪ rɪja=ne kɪʊɪ=to kʰajɪ
Rahul=ERG say.PFV.MSG that Riya=ERG kiwi=TO eat.PFV.FSG
'Rahul said that Riya ate the kiwi.'

Implicature 1: The speaker is uncertain whether Riya ate the apple.
Implicature 2: Rahul is uncertain whether Riya ate the apple.

3. Perspective Dependence

Gutzmann (2013) proposes this property of UCIs to account for the special cases that were displaceable under Potts (2007b) conception of the property of nondisplaceability. According to this property, "expressive content is evaluated from a particular perspective. In general, the perspective is the speaker's, but there can be deviations if conditions are right" (Gutzmann 2013: 42). By this property, the interpretation of UC meaning can be evaluated with respect to some salient individual's perspective ("the contextual judge" in Potts (2007b: 173)). If the contextual conditions are such that the contextual judge parameter is set to a non-utterance-speaker individual, then UC meaning is interpreted relative to that individual's perspective.

In example (39), the UC meaning of particle *-to* can be interpreted depending on the contextual judge's perspective. Since this sentence involves a reported speech verb, the contextual judge could be either the speaker of the utterance or the speaker of the reported speech verb. Therefore, Hindi enclitic particle *-to* exhibits the perspective dependence property of UCIs.

4. Descriptive Ineffability

Gutzmann revised Pott's formulation of this property to state that "it is impossible to paraphrase expressive content using only descriptive expression without changing the modus of expressing" (Gutzmann 2013: 46).[30] Kaplan (1999) explains the difference

[30] Potts described this property as "speakers are never fully satisfied when they paraphrase expressive content using descriptive i.e., nonexpressive, terms" (Potts (2007b: 166). The term 'satisfied' is contentious here.

in modus of expressing between use-conditional content and truth-conditional content such that the former 'expresses' or 'displays' while the latter 'describes'.

The UC meaning of particle *-to conveys* speaker salience on a constituent (by virtue of the act of marking it with a particle *-to*); it does not *describe* speaker salience on a constituent. This difference in modus of expressing becomes apparent in example (40) where paraphrase of the UC content in descriptive terms is pragmatically infelicitous. Therefore, the particle *-to* is descriptively ineffable.

(40) rahʊl=ne kɪʊɪ=to kʰajɪ
'Rahul ate the kiwi.'
UC as a TC content: #Speaker imposes salience on kiwi.

5. Immediacy

This property requires that "like performatives, expressives achieve their intended act simply by being uttered; they do not offer content so much as inflict it." (Gutzmann (2013:46). Uttering a performative construction is sufficient to perform the speech act encoded in it (cf. Potts 2007b).

The discussion in section 4.3.2 had exhibited that the expressive meaning conveyed by particle *-to* can be analysed as a performative instructional meaning (cf. Portner 2007) i.e., by uttering a *to*-marked constituent in a sentence, the speaker *imposes* salience on it or salience is *conveyed* on it from a speaker's perspective. Therefore, particle *-to* also exhibits the characteristic property of 'immediacy' for a UCI.

In conclusion, the Hindi enclitic particle *-to* does not affect the truth conditions of a sentence but affects the use conditions of a sentence. It also exhibits the diagnostic characteristic properties of a UCI, namely 'independence', 'nondisplaceability', 'perspective dependence', 'descriptive ineffability' and 'immediacy'.[31] Therefore, this

[31] A sixth characteristic property of repeatability had been proposed by Potts, which was stated as "if a speaker repeatedly uses an expressive item, the effect is generally one of strengthening the emotive content, rather than one of redundancy" (Potts 2007b, 167). However, Geurts (2007) and Gutzmann (2013) contradict Potts's assumption and claim that this property cannot be used to diagnose UC

section has verified the status of enclitic particle *-to* as a UCI that contributes conventional non-truth conditional meaning in the semantic interpretation of the sentence it occurs in.

## 4.5. Conclusion

This chapter has proposed an analysis of the meaning conveyed by particle *-to* within the frameworks that posit multiple dimensions of meaning for linguistic expression. Various concepts of meanings were introduced and segregated on two broad categories of a meaning divide — lexical meaning, descriptive meaning, truth conditional meaning and conceptual meaning on one side; expressive meaning, non-truth conditional meaning, procedural meaning and instructional meaning on the other. The meaning of particle *-to* is proposed in terms of the concept of discourse salience or discourse prominence after analyzing the interpretative difference obtained in the case of a marked *-to* sentence (as compared to the unmarked sentence). I propose that this particle encodes a kind of use conditional or an instructional meaning i.e., *to convey speaker salience on the constituent marked by -to*. Adopting Mulkerns's (2007) proposal that 'emphasis' and 'contrast' are the motivations for imposing salience on a constituent by a speaker, I propose that a topical interpretation or a contrastive interpretation is available to a *to*-marked constituent as a derived effect of this imposed salience on it by the speaker. 'Conveying or imposing salience' is a type of performative that is conventionally encoded by this particle and is effected when a speaker utters a sentence containing a *to*-marked sentence. A formal definition of this meaning in type theoretic terms within the Hybrid Semantics framework is proposed. Additionally, this particle satisfies all the diagnostic properties that characterize a use conditional item. Within the typology proposed for UCI items, particle *-to* identifies as an expletive, functional, non-shunting UCI with the featural composition [+f, -2d, -rs].

---

content from TC content as this property is neither a sufficient nor a necessary condition for expressive content. Therefore, I do not include this property in the discussion of diagnostic characteristic properties of UCIs.

This salience-based, use-conditional analysis for the meaning of particle *-to* satisfies Egg's (2013) condition of both the 'general' and the 'specific' aspect of a discourse particle's semantic treatment. 'Convey speaker-salience on a constituent' is general enough that any constituent with a conceptual meaning can be marked salient in a sentence. This meaning is not constrained by any clause type restriction and can account for a *to*-marked constituent's occurence in declaratives, imperatives or confirmational interrogatives. This meaning is 'specific' enough to derive the IS interpretation of topicality and contrast on a constituent it marks. Since emphasis and contrast are the only motivating factors for a speaker to impose salience on some entity, no other interpretation besides 'topic' and 'contrast' can be derived for a *to*-marked constituent.

The closing question of the preceding chapter (i.e., how this particle can be paraphrased) is answered by the last section of this chapter. Since particle *-to* is a use-conditional item for Hindi, it is characteristically descriptively ineffable. Therefore, the meaning of this particle cannot be paraphrased in any descriptive terms without changing what it actually *expresses*. The next chapter focusses on the question of syntax of this particle.

# CHAPTER 5

# THE SYNTAX OF PARTICLE *-to*

The preceding chapters have motivated a re-investigation of the enclitic particle *–to* and claimed that this particle is not an (inherent) morphological topic marker. The IS notions of topicality and contrast are derived interpretive effects of marking of a constituent via this particle. The primary meaning encoded by this particle is an instructional or a use-conditional meaning — that of conveying speaker salience on the denotation of the constituent this particle marks. Thus, this particle signals that in the speaker's model of discourse, the marked constituent is being singled out to make it more salient or prominent for the current utterance and the upcoming discourse. A topical or a contrastive or a contrastive topic interpretation are the interpretive effects contingent on a pragmatic factor (presence of a relevant alternatives set in the Common Ground) and a syntactic factor (position in the sentence).

The dissociation of topicality from particle *-to* resonates with the current understanding of the generative procedure within the biolinguistics framework.[1] Contrary to the cartographic approaches (pursued, most notably, by Cinque and Rizzi), Chomsky, Gallego and Ott emphasize that "informational notions such as "topic" or "focus", like grammatical functions or thematic roles, are properties of configurations and their syntactic/discursive context, not of individual syntactic objects (Chomsky 1965; Hale & Keyser 1993); consequently, they should neither be represented in the lexicon, nor in the narrow syntactic derivation (cf. Uriagereka 2003; Fortuny 2008; López 2009; Gallego 2013a)" (Chomsky *et al.* 2019:250). In the case of *to*-marked XPs in Hindi, neither the XP nor the particle *-to* comes specified with a [topic] feature from the lexicon. Such a feature, besides violating Inclusiveness Condition (see section 2.1.1), would restrict the discursive interpretation of a *to*-

[1]Chomsky, in his UCLA Lectures (2019), terms the current generative enterprise to be a biolinguistic enterprise since its object of study ('the basic property of language' — every language generating an infinite array of expressions with each having a semantic interpretation and an instruction for externalisation at the interfaces) is a biological property unique to the species of humans.

marked XP to only topicality but that is evidently not what the Hindi data exhibits. Thus, any syntactic framework that operationalizes a [topic] feature to drive a syntactic operation (like Jiménez-Fernández (2020), Frascarelli and Hinterhözl (2007) etc.) is not relevant for the proposal being built here.

A syntactic analysis of particle *-to* must adequately account for its empirical distribution in conjunction with its semantic-pragmatic interpretation (since it is the output of narrow syntax that is transferred to the Conceptual-Intensional (C-I) interface for interpretation). The occurrence of this particle is observed in declarative clauses, interrogative clauses (limited) and imperative clauses, besides marking a range of lexical categories like nouns, adjectives, numerals, postpositions, verbs and adverbs (see chapter 1).[2] Example (1) exhibits a *to*-marked NP in an utterance that, depending on the intonational contour of the verbal complex, could be interpreted as either being a declarative (an assertion with a falling intonation) or an interrogative (a confirmational question with a rising intonation). Keeping the discourse context same, an unmarked utterance in example (1) with a rising intonation is assumed to give a polar (yes/no) question interpretation (Bhatt and Dayal 2020).[3] The syntactic contribution of particle *-to* in changing a construction with a polar question interpretation to a confirmational question interpretation requires an explanation too.

(1) Context: A student, Rahul, was supposed to sing a song and a ghazal for a school function. After the function was over, someone says:

rahʊl=ne gana=to gaja
Rahul=ERG song=TO sing.PFV.MSG
'Rahul sang the song.' - with a falling intonation at the end
'Rahul sang the song, right?' - with a rising intonation at the end

In this chapter, I adopt two frameworks — the Interactional Spine Hypothesis (ISH) framework developed by Wiltschko (2021) and the Problems of Projection (POP) approach to minimalist enquiry (Chomsky 2013, 2015) — to incorporate and explain

[2] A *to*-marked XP cannot occur in a constituent question in Hindi but can occur in confirmational questions that involve a speaker bias (see chapter 4).

[3] Bhatt and Dayal (2020: 1117) cite the prosodic analysis of Butt *et al*. (2017) and Biezma *et al*. (2017) who associate the declarative falling intonation with a L-L% contour and the polar question rising intonation with L/H-H% contour.

the syntactic behavior of this particle. I give a brief overview of the previous attempts to syntactically analyse this particle as a discourse particle in section 5.1. In section 5.2, the ISH framework is introduced and discussed. Section 5.3 motivates the adoption of ISH framework for Hindi and discusses the analysis proposed for particle *-to* within it. Section 5.4 discusses the feature composition of this particle and its interaction with the core syntactic operations of Merge, Agree and Move in the POP model. Section 5.5 summarizes this chapter.

## 5.1. Particle *-to* as a Discourse Particle: Syntactic Perspective

Discourse particles, as a closed class of lexical items that are involved in CG management and discourse coherence, were introduced in Chapter 4. German modal particles like *ja* and *doch* are the most extensively studies cases of discourse particles in the literature (cf. Zimmerman 2011, Grosz 2016, Gutzmann 2017 etc.). These particles are assumed to convey the epistemic attitude of the speaker towards the propositional content or to convey what the speaker assumes their interlocutor's epistemic state is regarding some proposition. For example, particle *ja* in example (2a) signals that the speaker considers the hearer to be aware of the fact that Max is at sea; particle *doch* in example (2b) conveys that the speaker considers the hearer to be unaware of this fact (German data from Zimmerman 2011:3 ex(1a) and(1b)). Syntactically, these particles are restricted to a 'middle field' *i.e*, to the right of verb-second position and to the left of non-finite verb position (Grosz 2016b:5).

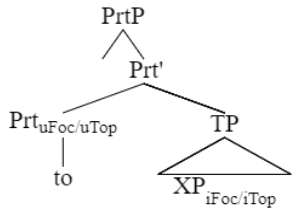
(2) a. Max ist *ja* auf See
b. Max ist *doch* auf See.
'Max is PRT at sea.'

A similarity between German particle *doch* and Hindi particle *–to* has been noted by Deo (2021) and Bayer (2018).[4] For the syntax of Bangla particle *to*, Bayer *et al.* (2014:10) propose a 'functional head in phrase structure' analysis wherein particle *to* is a functional head and focus-bearing constituents move into its specifier position.

[4] Strictly speaking, Bayer (2018) compares German *doch* with Bangla *to* and hints at a common core of these related but distant languages. He extends the comparative analysis to Hindi enclitic *–to* at certain points, under the umbrella term Indo-Aryan *-to*.

This particle (that is assumed to have an uninterpretable focus [uFoc] or topic [uTop] feature in their system) is merged with TP and projects a *to*P.[5] The semantic argument of this particle (with an interpretable focus [iFoc] or topic [iTop] feature) moves to the spec-*to*P position to value and delete the [uFoc] or [uTop] feature of this particle. (3) exhibits the projected tree structure for a Bangla sentence with an *XP-to* constituent

(3)




Since my theoretical assumptions regarding the discourse-related features is different from theirs, I do not discuss this proposal further. A syntactic approach to Hindi enclitic *-to* as a discourse particle had been proposed in Kidwai (2000) and briefly mentioned by Bayer and Struckmeier (2017) and Paul and Pan (2017). An overview of each of these is provided next.

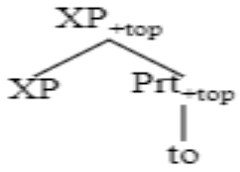
Section 1.2.5 had discussed Kidwai's proposal to adopt Bayer (1996) and Rothstein's (1991) notion of a minor functional head category for particle *-to*. This type of functional head subcategorizes for an argument without itself having any theta grid or binding a theta position. This functional head did not project its own category but its unique feature percolated up with the category of its complement. (4), repeated from chapter 1, shows the projected tree structure of a *to*-marked XP in this system. In Kidwai's analysis, particle *-to*, being a topic marker, is lexically specified for a

[5] Similar to Gutzmann's (2017) list of characteristic properties of the German discourse particles, Bayer *et al.* (2014:1) propose the list in (1) to be diagnostic of Bangla discourse particles, including Bangla particle *to*. Although I do not concur with their syntactic analysis, Hindi particle *-to* does exhibit the properties in (1b), (1e), (1f) and (1g).

(1) a. clause-type dependence (decl/interr/imp)
b. restriction to the root clause i.e., to direct speech
c. fixed position
d. unaccented
e. impossible in isolation
f. most are enclitic
g. contribute expressive meaning, not propositional meaning i.e. no interference with truth conditions

[+top(ic)] feature. This [+top] feature passes up to the XP node dominating this particle.

(4)




Besides the previously mentioned problem of postulating a discourse feature on a lexical item itself, this analysis cannot be maintained in view of the current minimalist assumptions about language faculty (cf. Chomsky 2000 and later works). With the aim to "reduce the postulated richness of UG" (Chomsky 2013: 38), it restricts the set of functional heads that operate in narrow syntax to C, T and v. Therefore, the category of minor functional head has to be eliminated from the grammar.

Bayer and Struckmeier (2017), in their review of the *status quo* of discourse particles, state that for a head-final language, discourse particles are often clitic-like elements that either attach to major constituents at the end of the phrase (like Hindi or Bangla) or they attach at clause-final position (like Chinese). Discourse particles in these languages are "generally on a par with "regular" functional heads such as interrogative particles" (Bayer and Struckmeier 2017:8).[6] An instantiation of this claim is developed in Paul and Pan (2017) that analyses sentence final particles in Mandarin Chinese as belonging to a three-layered split CP domain. The first layer $C_1$ is closest to TP and conveys 'tense', the second layer $C_2$ expresses 'force' and the highest layer $C_3$ expresses 'attitude' i.e., the speaker's attitudes and feelings. Sentence final particles, in their proposal, are functional heads that select and project. The grammar maintains the hierarchy between the co-occurring sentence final particles that belong to the three distinct CP layers. The Mandarin Chinese particle *ei* is an Attitude head ($C_3$ layer) which indicates that the speaker issues a reminder to the hearer even though the speaker knows that the hearer is upto date regarding the state

[6] In contrast, most German discourse particles are phrasal modifiers rather than syntactic heads in their view since finite verbs move across them without intervention effects to yield V2 surface order.

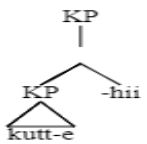
of affairs. This, they assume, is "reminiscent of German *ja* and *doch* as well as the particle *to* in Bangla and Hindi *to*" (Paul and Pan 2017: 4). Although the semantic interpretation of particle *-to* hinted at in this analysis is not in concord with the one being proposed in this thesis, but such a syntactic analysis that encodes a speaker-related domain, distinct from clause type domain, seems to be on the right track. This is discussed in section 5.3.

A review of the discourse particle-oriented approaches to Hindi particle *-to* has exhibited that a functional head analysis of it is theory-internally motivated.[7] This particle is not a C-like element (like in Chinese) nor is there any motivation to fix a syntactic position for it in the clausal spine with the rest of the sentential constituents moving around it. Thus, no previous proposal on particle *-to* provides a syntactic explanation for it. The next section discusses the Interactional Spine Hypothesis, the framework that I adopt to model the syntax of particle *-to*.

## 5.2. The Interactional Spine Hypothesis (ISH) Framework

Wiltschko, in her monograph titled 'The Grammar of Interactional Language' (2021), develops a framework that attempts to syntactically derive the form-meaning pair of linguistic units that regulate interaction cross-linguistically.

[7] Bajaj (2016), adopting Sharma's (2003) proposal that all discourse markers can adjoin to any part of the NP, proposes the following adjunction based syntactic tree for Hindi enclitic particle *-hii* marked oblique NP *kutt-e* 'dogs-oblique' .




She postulates (but doesn't expand on this) that this particle moves to the left periphery by a covert LF movement to get its semantic interpretation.

### *5.2.1. Introduction to the Framework*

The central premise of this framework is that Grammar configures both — the language used to express thought (i.e., the propositional language — utterances that convey thoughts about the world and can be evaluated as true or false) and the language used to regulate interaction (i.e., an 'interactional' language that includes discourse markers, intonational cues, confirmationals and response markers *etc.*).[8] The unit of analysis is not constrained to be of clause-size (as assumed in generative tradition) but can be smaller or bigger in size than a clause and can have either a morphological or a prosodic profile. These are termed as 'Units of Language' (UoLs). Such UoLs are integrated with the host clause (that encodes propositional content).

A UoL that has an effect on the use of language in interaction has a computable form and meaning. These UoLs often have a propositional function as well. They exhibit specific restrictions on their ordering in the sentence besides being prosodically integrated into it. In example (5a), 'eh' is a sentence-final discourse marker (or narrowly a 'confirmational') that is prosodically integrated with the propositional clause.[9] Evidence for this comes from the fact that a sentence-final intonation rise (indicated by ↗ symbol) can occur on the bare clause (example (5b)) or on the confirmational 'eh' (example (5c)) but not on both the propositional clause and the confirmational (example (5d)) and neither on just the propositional clause without integrating the confirmational (example (5e)). Interlocutors judge the complex expressions in examples (5d) and (5e) to be ill-formed.

(5) a. Context: Upon seeing his neighbour park a brand new car in front of his house one day, the speaker says to him:

Speaker: You got a new car, eh?

[8] Between the generativist v/s the functionalist approach to language (i.e., 'language as a means of expressing thought' v/s 'language as a means of communication'), this framework claims that language is *both* thought and communication. One of the arguments supporting this claim is that speakers have judgments regarding the well-formedness of interactional language too. Hence, it is part of the competence of the speaker (see Wiltschko 2021:3 for this discussion).

[9] Wiltschko notes that this confirmational 'eh' is a common feature of Canadian English.

b. You got a new car ↗
c. You got a new car, eh ↗
d. * You got a new car ↗, eh ↗
e. * You got a new car ↗, eh

One core assumption of this model is that the same computational system, that derives propositional language, derives interactional language too. Since sentences are hierarchically structured (under the standard generative theorizing), this model proposes that interactional language is hierarchically structured too, with the latter dominating the former.

Interactions involve a speaker and an addressee engaged in *doing* things (like updating each other's knowledge state about the state of affairs in the world or requesting/asking them to do so or to synchronize their mental states) by *saying* things.[10] Therefore, interaction embeds speech acts.[11] Clause types (i.e., declarative, interrogative, imperative, exclamative, subjunctive) are assumed to be indicators of the illocutionary points. However, UoLs, like discourse markers and sentence prosody, can modify the speech act category that was determined by the clause type. For example, the utterance in example (5a) had the declarative clause type (which is not associated with a 'directive' illocutionary act) but the peripheral discourse marker changed its illocutionary force from being a 'representative' to a 'directive'.[12]

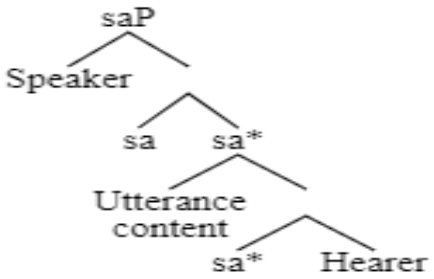
[10] Interaction is canonically manifested as inter-person but it can be intra-person too — the case of 'self-talk' where the same person assumes the role of both the speaker and 'the other' and this exhibits the characteristics of interactional language.

[11] Austin (1962) introduced the theory of speech acts to formalize the insight that any utterance is basically an act of speech. Speech acts are of three kinds: Locution (what is said); Illocution (what is intended by the speaker); Perlocution (what is effected in the addressee). They are associated with felicity conditions that restrict their context of use. Searle (1976) developed Searle's ideas further and decomposed the illocutionary act into further sub-units (like the illocutionary point and the direction of fit).

[12] A 'representative' illocutionary point commits the speaker to the truth of the proposition whereas a 'directive' illocutionary point is an attempt by the speaker to get the addressee to do something. Other illocutionary points are 'commissive' (speaker is committed to the future course of action), 'expressive' (speaker expresses their psychological state about p) and 'declaration' (speaking *p* leads to performing *p*) (Searle 1976:8).

Building up on the previous attempts to syntacticize the speech act layer (cf. Sadock 1969a, Ross 1970), Wiltschko and Heim (2016) proposed a neo-performative hypothesis that "speech act structure is part of the extended functional architecture" (Wiltschko 2021:24).[13] This speech act structure is developed from Speas and Tenny's (2003) proposal of an articulated speech act (saP) structure that selects for a root clause rather than a Rizzi (1997) type approach that introduces a 'Force' head inside an articulated CP layer. [14] The speech act structure proposed by Speas and Tenny is given in (6). Besides this, they also proposed a 'point of view' structure that is encoded via evidentials, speaker evaluative adverbs etc.

(6)




Wiltschko notes that while such a model has a structure corresponding to the locutionary act (the propositional structure/utterance content) and illocutionary act (the speech act structure), the perlocutionary act and the role of the addressee in regulating interaction is not incorporated in such a syntactic structure, besides the lack of marking of a 'hearer' projection. The problem that this issue raises is that "the transfer of the descriptive content of what is being said from the speaker to addressee is viewed as something that is non-negotiable" (Wiltschko 2021:36). In such a model, a speaker just by uttering a proposition *p* adds it to the Common Ground between the

[13] Ross (1970) proposed introducing an abstract speech act structure above the propositional structure within the deep structure in syntax and this speech act structure is removed from the surface structure by a transformational rule of deletion. The speech act structure has a subject (the speaker), an indirect object (the addressee) and a verb of communication that tells the illocutionary force. For example: the proposition 'I had food' is interpreted as 'I(speaker) tell you(addressee) that I had food' with the speech act structure being deleted later on. This is the Performative Hypothesis.

[14] Rizzi's 'Force' functional head is related to the syntactic clause typing and not the pragmatic illocutionary force *per se*.

interlocutors (the Stalnakerian version of CG). However, in natural language interaction, there is negotiation involved (cf. the 'Table' in Farkas and Bruce 2010).[15] The hearer has a more active role than previously assumed and they have to signal their agreement and thereby, let the proposition *p* be added to the CG or they can disagree and reject its addition to the CG. In an interactional sequence of discourse, the speaker can request a response from the hearer (an 'initiation') or the hearer can provide a response (a 'reaction'). In the absence of hearer's acknowledgement (of their understanding speaker's intention) or response, the interaction is not considered felicitous. This dimension of interactivity had not been incorporated in any formal theory previously.

### *5.2.2. Ingredients of the ISH Framework*

Wiltschko (2021) hypothesizes that an 'interactional spine' embeds the traditional sentence structure. This interactional spine has two layers— a 'grounding' layer and a 'responding' layer. "The core function of the grounding layer is for the speaker to configure the propositional content of the utterance so that the addressee can update their knowledge state to include it. The core function of the response layer is to manage the moves that serve to synchronize the interlocutor's knowledge states" (Wiltschko 2021:72). This is the Interactional Spine Hypothesis (ISH). Language in interaction is not fully regulated by grammar but only a part of it is. The other part being regulated by pragmatic principles like the Gricean cooperative principle that stipulates how a normal conversation should proceed. The part of interactive force that is configured by grammar is 'grounding' and 'turn-taking' (formal aspects of interactional language).

The ISH is based on a framework adopted for Universal Spine Hypothesis (Wiltschko 2014) that posited a universal spine consisting of four functions (classification, point-of-view, anchoring and linking). ISH extends this universal spine to include the

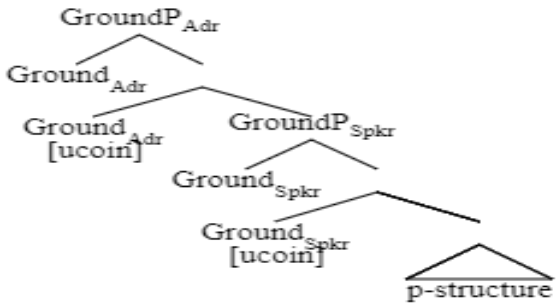
[15] In the Farkas and Bruce Table Model (2010), mentioned in footnote 56 in chapter 3, the items (utterances or syntactic objects and their denotations) *at issue* in a conversation are stacked at the Table. Once an issue is resolved, it is removed from the Table and added to the CG. The goal of a conversation is to empty the Table of all items.

grounding and responding functions. A UoL's distribution and interpretation is dependent on which function it associates with in the universal spine. Under the assumption that same formal mechanism operates across both domains (since Grammar configures both), each spinal head has two arguments in the complement and the specifier position. Wiltschko assumes each spinal head to be a transitive function that relates its two arguments. The formal notions of interpretability /uninterpretability and feature valuation (Chomsky 1995) are assumed here. Each spinal head has an unvalued 'coincidence feature' [*u*coin], a grammatical feature that is intrinsically binary, that has to be valued to be interpreted. This [*u*coin] feature is assigned either a positive value or a negative value by the substantive content of some UoL (Wiltschko 2021:79). This feature mediates the relation between the two arguments of a spinal head. A positive coincidence feature asserts that the argument in complement feature coincides with the argument in specifier position. A negative coincidence feature asserts that both the arguments do not coincide.

The Grounding Layer

The 'Ground' spinal head relates an utterance to the interlocutors' mental states. This grounding layer adds a subjective component to the propositional content embedded under it — that of "asserting that the propositional content is or is not in the knowledge state of the interlocutor" (Wiltschko 2021:82). Ground head takes the propositional structure as its complement. It introduces an abstract argument in its specifier position — the 'ground'. The 'ground' represents the mental state of the interlocutor and consists of "the mental representations of our thoughts about the world" (Wiltschko 2021: 82). This mental world corresponds to the concept of Common Ground. However, this CG is not the shared set of mutual beliefs (as originally proposed by Stalnaker) but rather it is relativized to each interlocutor. The speaker-oriented grounding layer asserts whether a given proposition is or is not in the speaker's ground. Similarly, the addressee-oriented grounding layer asserts whether or not the given proposition is in the addressee's ground from the speaker's perspective. (7) exhibits this articulated ground layer tree structure.

(7)



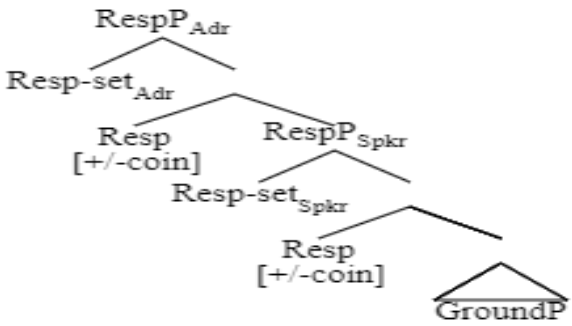

The ground consists of propositions, discourse referents as well as "individual interlocutor's publicly displayed attitudes towards the propositions" (Wiltschko 2021:89). This structure contributes the attitudinal aspect of meaning.[16] The grounding layer synchronizes the interlocutors' minds i.e., makes the interlocutor aware about the status of other's ground, including their attitude or belief state towards the proposition.

The Responding Layer

The Resp(onse)P head encodes the function of 'turn-taking' in interactional language. The complement of this head is the GroundP. The specifier of this head is an abstract argument — the 'response set'. This is "the set of items that an interlocutor has to respond to" (Wiltschko 2021:83). This set is conceptually similar to the notion of the 'Table' in Farkas and Bruce (2010). Similar to the GroundP layer, the RespP layer is also relativized with respect to both the speaker and the addressee. For a response set indexed to the addressee, the coincidence value of the Resp head asserts whether the utterance is or is not placed in addressee's response set (thereby, marking it as an initiating move by the speaker in an interaction). Similarly, for a response set indexed to the speaker, the co-incidence value of the Resp head asserts whether the utterance is or is not placed in speaker's response set (thereby, marking it as a reacting move by the speaker in an interaction). (8) represents the articulated responding layer as proposed in ISH.

[16] Under the assumption that believing is a weaker attitude than knowing, Wiltschko proposes a finer interpretation of the contents of the ground. If a proposition in the ground is without an explicit propositional attitude like 'belief', then it indicates that the person is completely certain about that proposition i.e., they 'know' it. In contrast, a propositional attitude like belief is correlated with uncertainty (or a gradable notion of certainty) or subjectivity on behalf of the ground-holder.

(8)




### 5.3. Hindi Particle *-to* within the ISH Framework

In this section, I adopt the ISH Framework to propose an analysis of Hindi enclitic particle *-to*. The reasons for adopting this framework as a research program to couch particle *-to* in are multifold. Firstly, it integrates UoLs that have both a truth conditional meaning as well as a non-truth conditional meaning into the syntactic spine (Wiltschko 2021:204). It proposes that UoLs that have a truth conditional meaning associate with the spine within the propositional structure (structure that can be evaluated for a true or false value) and a UoL with a non-truth conditional meaning associates with the interactional structure (structure beyond the scope of truth conditional evaluation). Secondly, ISH provides a syntactic representation of speaker-oriented ground and addressee-oriented ground. It expands the standard assumption about the contents of the ground that includes, besides the proposition itself, the ground-holder's publicly displayed subjective attitudes (like beliefs, questions, desires etc) about the proposition. Thus, it represents a person's mental state more accurately and according to the ISH, grammar configures ways in which the interlocutors synchronize and get to know each other's mental states in interaction.[17] Thirdly, this model does not consider an illocutionary act to be a monolith and decomposes it further — into an illocutionary force (the speaker's intention) and an interactive force (regulation of interaction with another interlocutor). The illocutionary point of an utterance is encoded in the propositional structure whereas the different felicity conditions associated with each are manifested via the different functions in the interactional structure. Fourthly, an intonational tune can be a UoL in this framework,

[17] Wiltschko maintains that synchronization of the mental states of the interlocutors does not entail an agreement about all the contents of the ground. It just means that the interlocutors are aware about each other's knowledge or belief states in an interaction.

which can associate with a spinal functional head in the interactional structure. The importance of these factors will become apparent in the discussion below.

### *5.3.1 The Analysis*

Adopting the ISH framework, I propose that the Hindi enclitic particle *-to* is a UoL that syntactically associates with the interactional structure and not the propositional structure.[18] Evidence for no syntactic relevance of this particle for the propositional structure is discussed in section 5.4. Within the ISH framework, this prediction is borne out as a corollary of the assumption that this particle does not contribute to the truth conditional meaning and hence, it can syntactically associate with a functional head (to receive interpretation) only in the interactional spine and not in the propositional spine. Underlying this assumption is the observation that UoLs that regulate interaction are not sensitive to the truth-value of the utterance involved.

The proposed non-truth conditional meaning of enclitic particle *-to* is that of conveying 'speaker salience' on the denotation of the constituent it marks (chapter 4). Since this is a type of use conditional (UC) meaning, particle *-to* can be assumed to be inherently specified for a [$UC_{spkr\text{-}sal}$] semantic feature.[19] Assuming an extended interactional spine structure for Hindi, I propose that the enclitic particle *-to* can be analysed as a UoL that is integrated within the propositional structure but it functionally associates with the speaker-oriented grounding layer in the interactional structure.

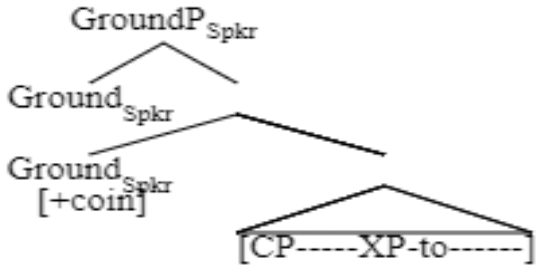
[18] Any utterance containing a *to*-marked XP has a speaker or a contextual 'judge' (as argued for in section 4.4.2). The addressee is not directly relevant for this particle's interpretation but there is always an implicitly assumed addressee (if not an explicit one in the context) for any utterance containing a *to*-marked XP.

[19] The idea of incorporating [UC] as a feature is credited to Ayesha Kidwai (pc). The discussion of [UC] as a feature is taken up in section 5.4. For the current purpose it is sufficient to note that our system needs to distinguish UC meanings of different kinds. So I propose that the particle *-to* is specified for a [$UC_{spkr\text{-}sal}$] feature whereas an ethnic slur like *chinki* in Hindi (cf. section 4.3.2), that has a pejorative evaluative meaning, is specified for a [$UC_{pejorative}$] feature.

Wiltschko (2021:209) proposes the availability of two options to account for those UoLs that are distributionally available at one place but are functionally interpreted relative to another place. First, an abstract Agree relation is motivated to account for the fact that UoLs like German discourse particles are spelled out in one position (adjoined to IP) and interpreted relative to the grounding layer (Thoma 2016). A second option is a covert movement operation of the particle to associate with GroundP at the level of interpretation. Wiltschko leaves this option open-ended. I propose an Agree-type long-distance feature valuation for the Hindi case since particle *-to* is not a freely occurring lexical item in the language but is syntactically associated and semantically dependent on another lexical category XP. A covert movement operation would have to target and move the full *XP-to* phrasal unit into the $\text{Ground}_{\text{Spkr}}$ head position. Phrasal units, under standard generative assumptions, cannot move to head positions. Therefore, this option is not syntactically feasible for Hindi.

The $[\text{UC}_{\text{spkr-sal}}]$ feature, by the operation of (abstract) Agree, associates with and values the unvalued co-incidence feature of $\text{Ground}_{\text{Spkr}}$ head as positive. In the ISH model, only substantive content can value the [*u*coin] feature of a functional head. I propose that in the case of particle *-to* with no conceptual substantive content *per se*, the $[\text{UC}_{\text{spkr-sal}}]$ feature, by virtue of being a semantic feature, can value the [*u*coin] feature of $\text{Ground}_{\text{Spkr}}$. This [+coin] feature of $\text{Ground}_{\text{Spkr}}$ head asserts that the proposition (in the complement position) is a part of the speaker's ground. (9) is the syntactic tree for an utterance (CP) containing a *to*-marked XP within the ISH Framework.

(9)



Thus, within the ISH framework, particle *-to* regulates interaction by indicating the mental state (the ground) of the speaker and thereby, facilitating the synchronization of the mental states of the interlocutors. The next section investigates what is asserted to be added to the ground by an utterance containing the particle *-to*.

### *5.3.2. What is added to the Ground$_{Spkr}$?*

A core tenet of the ISH is that the interactional language is partly dependent on grammar and partly on the pragmatic assumptions about normal conversations. Grammar does not decide *when* interactional language will be used (that is regulated by discourse pragmatics) but only how it is configured when it is used. As a null hypothesis, ISH claims that language speakers employ discourse markers to overtly mark a deviation or violation from the felicity conditions that are normally associated with declaratives, interrogatives or imperatives to encode an assertion, a question or a desire respectively.[20]

To illustrate the case of one clause type, a declarative form (let's assume *p*) in combination with some specific felicity conditions functions to assert *p*. An unmarked declarative *p* (like example (10a)) is used as an assertion when two associated felicity conditions hold — the speaker knows *p* and the speaker assumes that the hearer does not know *p*. In this case, a positive valuation of Ground$_{Spkr}$ adds the bare proposition to the ground i.e., *p* itself.[21] In contrast to (10a), a declarative form marked by particle *-to* (like example (10b)) indicates a deviation from the felicity conditions associated with an assertion. This deviation is in terms of the first felicity condition i.e., instead of the bare proposition, what is asserted to be added to the speaker's ground in case of a *to*-marked utterance.

[20]Wiltschko (2021:104) assumes that the complement of the Ground is CP (the 'linking' layer in USH model) whose core function is to encode clause type and/or polarity. It is the clause type and the felicity conditions encoded by the grounding and responding layer that compose the speech act of an utterance.

[21] A bare proposition in the ground is implicitly embedded under a 'knowing' propositional attitude verb. *Know(p)* implies that the ground-holder is 100% certain that *p* holds.

(10) a. rahʊl aja
Rahul come.PFV.MSG
'Rahul came.'

b. rahʊl=to aja
Rahul=TO come.PFV.MSG
'Rahul-to came'

A First Pass at Derivation

I propose that by marking a constituent with particle *-to* in *p*, what is added to the speaker-ground is not the proposition *p* itself but the speaker's attitude towards *p*, i.e the speaker asserts that they hold a propositional attitude of belief towards the proposition *p* in their ground.[22]

Motivation for proposing a propositional attitude of 'belief that p' for particle *-to* is sourced from the distribution of this particle in example (1), repeated here as example (11). The utterance in example (11a) is interpreted to be an assertion. The utterance in example (11b) is interpreted to be a question. Crucially, the utterance in example (11b) can be interpreted as a confirmational question that comes with a speaker bias. This question can be paraphrased as the assertion followed by a confirmational marker (given in the English translation for example (11b)). I propose that this speaker bias is the speaker's propositional attitude of belief. This speaker bias is removed if the particle *-to* is removed from example (11b). Thus, the source of speaker's attitude that they believe that the proposition holds is conveyed by particle *-to* marking of a constituent in an utterance.

(11) Context: A student, Rahul, was supposed to sing a song and a ghazal for a school function. After the function was over, someone says:

[22] I assume that the hybrid semantic interpretation of the sentence (the pair of truth conditional (TC) meaning and UC meaning) as well as the discursive interpretation of imposed salience resulting in emphasis or contrast on the *to*-marked XP are C-I interface phenomenon and not regulated by the syntactic spine *per se*. The role of syntax of particle *-to* is limited to asserting the speaker's mental state or ground, which is mediated via the spinal function of grounding layer.

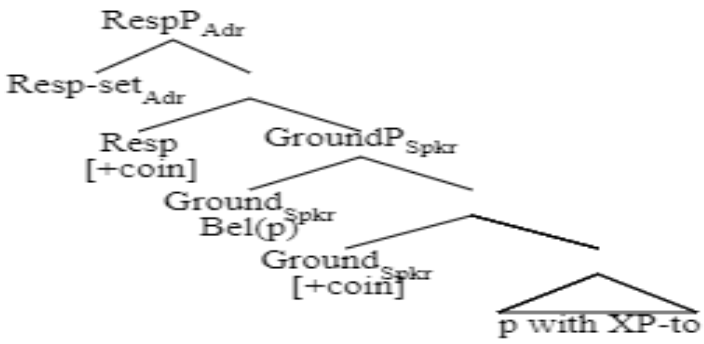
| | | |
|---|---|---|
| rahʊl=ne | gana=to | gaja |
| Rahul=ERG | song=TO | sing.PFV.MSG |

a. ‘Rahul sang the song.’ - with a falling intonation at the end
b. ‘Rahul sang the song, right?’ - with a rising intonation at the end

The derivation of the utterance in example (11a) proceeds by the particle *-to* positively valuing the [ucoin] feature of GroundSpkr. What is asserted to be in the speaker’s ground is that the speaker believes *p* (i.e., Rahul sang the song). Similarly, in example (11b), particle *-to* conveys that the speaker has Bel(*p*) in their ground. The change in speech act from an assertion to a question is effected by felicity condition linked to the responding layer. A rising intonational cue can be analysed as a UoL in this framework. This UoL positively values the [ucoin] feature of the response-set indexed to the addressee in the structure of example (11b). This asserts that the speaker puts GroundP (the belief that *p*) in the stack of items that the addressee is supposed to respond to (his response-set). This tags this utterance as an ‘initiating’ move that asks for a response from the addressee. This is the felicity condition associated with a question. Thus, in the derivation of the utterance in (11b), particle *-to* introduces the speaker’s attitude (i.e., Bel(p)) in the ground and the rising intonation introduces the condition that the speaker wants the hearer’s response on GroundP (i.e., their belief about p). A combination of these two interactional layers changes the speech act of the utterance in (11) from an assertion to a question.

(12) exhibits this syntactic derivation of a confirmation question containing XP-*to* within the ISH framework. The propositional attitude of Bel(p) is added to the $\text{Ground}_{\text{Spkr}}$ by the positive valuation of [ucoin] feature of the Ground head by the particle *–to*, The GroundP is added to the $\text{Resp-set}_{\text{Adr.}}$ via a positive valuation of the [ucoin] feature of the Resp head by the rising intonation cue.

(12)



*Prima facie*, this first pass attempt seems to be making the right empirical predications. For example, it predicts that the particle *-to* can occur in confirmational questions (since they involve a validation or a non-validation of the speaker's belief about a proposition) but it cannot occur in constituent questions or information-seeking questions (since these do not involve a belief state of the speaker towards the proposition). It could also explain why the same utterance (11b) but without particle *-to* is interpreted as a polar question.[23]

However, this belief-based analysis of particle *-to* cannot predict its marking in an imperative clause like example (13).

(13) Context: Speaker A says to speaker B, who has just entered his house to meet him.

| ek | gɪlas | pani=to | pijo |
|---|---|---|---|
| one | glass | water=TO | drink.IMP |

'(You) drink a glass of water.' (request)

The ISH framework assumes that "the *(interactional or speech act)* content of an imperative is the desire for the propositional content to be true (Des(p))" (Wiltschko 2021:111).[24] An unmarked imperative *p* asserts that the speaker has the propositional attitude of desire towards the propositional content in their ground. The speaker puts Des(p) on the Table which is interpreted as issuing a request to the addressee. Under these assumptions, an imperative with or without a discourse particle conveys the speaker's desire and not the belief about the proposition. Since particle *-to* is felicitous in an imperative utterance, this particle cannot be analyzed as grounding only the speaker's attitude of belief.

[23] By logical deduction — this an utterance does not come with the speaker's assertion of their propositional attitude towards the utterance and is thus, free of speaker bias. The rising intonation marks the utterance as an initiating move for which the speaker wants the addressee to respond.

[24] The italicized words in the quotation are not in the original text and have been added by me for expository purpose.

The Proposed Analysis

In view of the issue raised by imperatives, I modify my previous analysis and propose that instead of the speaker grounding their belief state regarding the proposition *p* (i.e., Bel(p)) by using particle *-to* in an utterance, what gets grounded is the speaker's attitude about proposition *p* specifically relative to the marked salient constituent. This is represented in (14).

(14) $\text{Ground}_{\text{Spkr}}$: Propositional Attitude (p(salient constituent))

I propose that the realisation of the speaker's propositional attitude in the speaker's ground is dependent on how the coincidence feature of $\text{Ground}_{\text{Spkr}}$ gets valued. Since it is the [UC] feature that values [ucoin] in our system, I propose that the type of [UC] feature that enters into Agree relation modifies the propositional attitude of the speaker that gets added to the speaker's ground. Additionally, the semantic argument (if any) of the lexical item with the [UC] feature is also visible to the grounding layer. When the $[\text{UC}_{\text{spkr-sal}}]$ feature of particle *-to* values [ucoin] of $\text{Ground}_{\text{Spkr}}$, what is asserted is that the speaker holds the relevant propositional attitude towards the proposition specifically relative to the marked salient constituent. In contrast, when the $[\text{UC}_{\text{pejorative}}]$ feature (see footnote 19) does the feature valuation of [ucoin] of $\text{Ground}_{\text{Spkr}}$, the speaker's relevant propositional attitude along with an added pejorative attitude towards the marked constituent is added to the speaker's ground.

The permissible propositional attitude depends on the type of content encoded in the propositional structure (i.e., declarative propositions are objects of 'belief', interrogative propositions are objects of 'ask' and imperative propositions are objects of 'desire'). Example (15a) exhibits a *to*-marked constituent in a declarative; example (15b) exhibits a *to*-marked constituent in a declarative form but with a rising intonation (the case of confirmational question) and example (15c) the imperative counterpart.[25]

[25] Under this proposed analysis, ISH does not make any predictions as to why particle *-to* cannot occur in constituent questions (interrogative forms). The grounding layer would assert that the speaker's ground contains his propositional attitude of asking about P specifically relative to the marked constituent. I propose that this derivation crashes not at the interactional structure but at C-I interface

(15) a. ram=ne=to gana gaja
Ram=ERG=TO song sing.PFV.MSG
'Ram sang the song.'

b. ram=ne gana=to gaja
Ram=ERG song=TO sing.PFV.MSG
'Ram sang the song, right?'

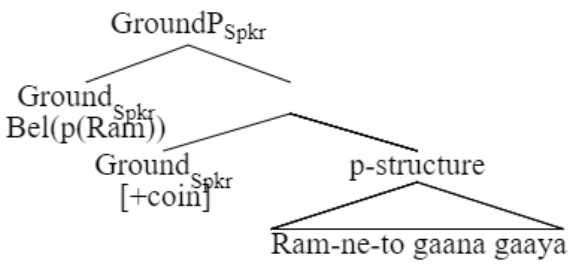
c. ek gılas panı=to pijo
one glass water=TO drink.IMP
'Please drink a glass of water.'

Example (15a) projects only a grounding layer. By the positive valuation of [ucoin] of $\text{Ground}_{\text{Spkr}}$, what is asserted is that the speaker believes the proposition *p* specifically relative to the marked salient constituent— 'Ram'. This is represented in (16) and the ISH tree for this utterance is given in (17).

(16) Assertion : ram=ne=to gana gaja
'Ram sang the song.'

$\text{Ground}_{\text{Spkr}}$ : Bel(p(Ram))

(17)




For example (15b), the derivation mechanism is the same as the one proposed in the previous section for the confirmational question except with one difference — the speaker's belief about *p* is specifically relative to 'song' now and is put in the addressee-indexed response-set. The addressee's response can validate or invalidate the speaker's belief. (18) outlines what is added to $\text{Ground}_{\text{Spkr}}$ and $\text{Resp-set}_{\text{Adr}}$ by the

---

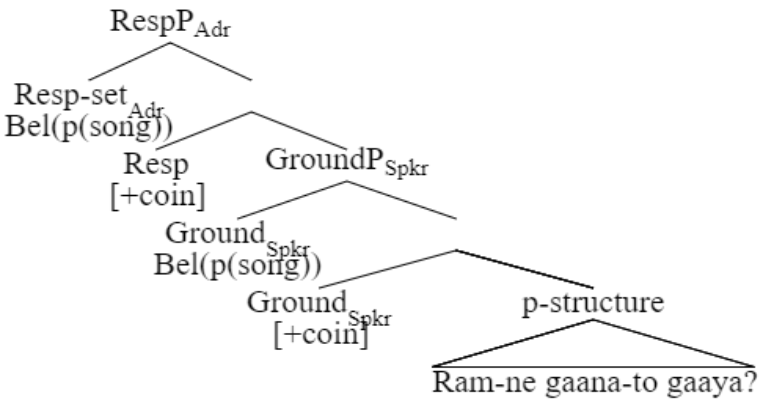
where two elements (the wh-word and the *to*-marked XP) contradict each other in being salient in the sentence and hence, important in the speaker mental model or being newsworthy for the upcoming discourse (cf. section 4.2.3).

positive valuation of [coin] feature of Resp head by uttering (15b). (19) exhibits the ISH tree for the same.

(18) Question: ram=ne gana=to gaja?
'Ram sang the song, right?'

$Ground_{Spkr}$: Bel(p(song))
Resp: [+coin] via rising intonation
$Resp\text{-}set_{Adr}$: $Ground_{Spkr}$P i.e., Bel(p(song))

(19)

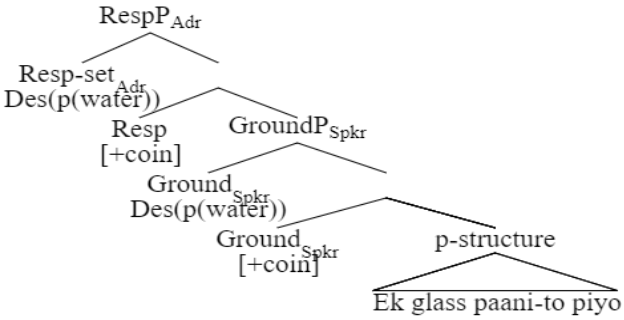



In example (15c) the speaker's desire for the proposition, specifically relative to 'water' is added to the speaker's ground. This propositional attitude of desire is then put in the addressee-indexed response set (i.e., a request is issued) for them to indicate their intention of fulfilling the request or not. This is represented in (20). The derivation of the utterance in (15b) in the ISH framework is given in (21).

(20) Request: ek gılas panı=to pijo
'Drink a glass of water'

$Ground_{Spkr}$: Des(p(water))
$Resp\text{-}set_{Adr}$: $Ground_{Spkr}$P i.e., Des(p(water))

(21)



In conclusion, the ISH framework can account for the empirical distribution of the enclitic particle *-to* in the language without requiring this particle to encode a [topic] feature or move the constituent it attaches to within the clausal spine. The proposed meaning of this particle is directly involved in associating this particle with the relevant spinal functional head — the grounding layer. In this section, I have proposed that this particle is integrated into the propositional structure but is syntactically interpreted relative to the interactional structure. This particle grounds the speakers' modified propositional attitudes about the encoded proposition in their mental ground. Thus, the speaker-oriented ground is linguistically manifested in this framework, which facilitates the synchronization of the interlocutors' mental states in an interaction.

The next section investigates the syntax of particle *-to* within the propositional structure. Contrary to the cartographic approaches like Rizzi (1997), that posit a ForceP layer in the CP domain, ISH maintains that force is not combined with clause type in the CP layer since "force is a function of the interactional structure dominating it" (Wiltschko 2021:117). Propositional structure ends at CP layer that syntactically manifest polarity and/or clause type. Within this model, UoLs can associate with the spine before or after syntactic computation.[26] I assume that this notion of syntactic computation corresponds to the notion of narrow syntax within the minimalist program.[27]

For the next section, I view the propositional structure from a minimalist perspective since ISH leaves the question of syntactic status of UoLs (like discourse markers) inside the propositional domain open for research.

[26] ISH claims that UoLs that associate with the spine early display 'category-neutral properties' because they can potentially associate with the syntactic spine at various positions. In contrast, UoLs that associate late are linked to categorical information and they "spell out syntactic configurations" (Wiltschko 2021:210).

[27] 'Narrow syntax' is the computational procedure of Faculty of Language (FL) that takes items from the lexicon 'Lex' and maps them onto expressions 'Exp' that are a pair of interface-legible sem(antic) and phon(etic) representations i.e. Exp=<Sem,Phon> (Chomsky 2000).

## 5.4. Particle *-to* within the propositional structure

This section is structured as follows: as a starting point, sub-section 5.4.1 proposes a featural composition of this particle; sub-section 5.4.2 addresses the question of merge of this particle in the syntactic spine and proposes an account of it in terms of Problems of Projection approach to minimalist enquiry (Chomsky 2013, 2015); sub-section 5.4.3 checks whether *-to*-marking has an effect on the agreement paradigm in the language or if it plays a role in syntactic movement of the constituents.

### *5.4.1. Featural (De-)Composition of Particle -to*

Within linguistics, the notion of a feature has been defined as "a set of values and the available options for their realization on linguistic elements" (Kibort 2008:6). They are basically the 'properties' that segregate some elements from the others. The features that make up particle *-to* can be analysed from a typological perspective (cf. Kibort 2008, 2010) or from a syntactic approach (cf. Svenonius 2007, Adger and Svenonius 2011). Following the standard assumption that lexical items are represented as feature bundles in the lexicon, I propose that enclitic particle *-to* is composed of a phonological feature (that is relevant for externalization at the sensory-motor interface) and a [UC] feature in the Hindi lexicon.

The preceding section had introduced the [UC] feature as a semantic feature. The intuition underlying it was the fact that linguistic entities not only encode a conceptual meaning but also encode an instructional or a use-conditional meaning (see section 4.3). Therefore, the set of semantic properties that are employed to segregate lexical items should be expanded from the set of conceptual features (like [animate], [human], [plural] etc.) to include a feature that conventionally encodes a use-conditional meaning i.e., a [UC] feature. Furthermore, I propose that enclitic particle *–to* is lexically specified for a [$UC_{spkr\text{-}sal}$] feature since it encodes a UC meaning of conveying speaker salience on a constituent. The [$UC_{spkr\text{-}sal}$] feature is type of [UC]

feature, just like [$UC_{pejorative}$] is a type of [UC] feature.[28] I assume that these are distinct types of the [UC] feature and not different tokens of the same type of feature.

Within the typology of grammatical features proposed by Kibort (2008, 2010), any feature of a language can be classified based on the parameters that fix how a feature value is realized on element. The [$UC_{spkr\text{-}sal}$] feature of particle *-to* is an 'inherent 'feature' (i.e., the feature value arises from within the element itself and is not determined by some other element) and is 'fixed' (since it is lexically supplied and is not selected from a given range of values) based on 'semantic' criteria. Since this feature is "a feature whose values are not involved in agreement or government but are inherent only", it is categorized as a 'morphosemantic' feature within this system (Kibort 2008:7).[29]

Adger and Svenonius, henceforth A&S, (2011) discuss the concept of features as properties of syntactic atoms that enter into derivation within the recent conceptualisations of minimalist syntax. They abandon a privative feature system in favour of a non-privative system as the ideal feature system for modelling syntactic dependencies since it violates the Inclusiveness principle (Chomsky 2000). Within such a system, a feature can be viewed as an attribute-value matrix. Features can be of a first order type (that segregate feature classes e.g., category feature) or can be of second order type (that "syntactically distinguish some instance α of a first-order feature from other instances of α" (A&S 2011:9); strength, interpretability and valuation of a feature are second-order features that establish syntactic dependencies. Following Svenonius (2007), they can be an interface feature (that play a role in both

[28] I propose that particle *-to* encodes a [$UC_{spkr\text{-}sal}$] feature and not a UC feature that encodes salience which selects a value 'speaker-salience' from the range of values {speaker-salience, hearer-salience} i.e. [$UC_{sal}$:spkr-sal] because, to my current understanding, there is no manifestation of the other 'value' i.e. hearer-salience as a feature in the language. Thus, there is no need to posit a feature-value matrix in this case.

[29] Kibort's feature inventory does not maintain a category of (pure) semantic features. Also, the concept of agreement evoked here necessarily requires copying of feature values between two units in the sentence. Thus, the concept of abstract Agree mentioned in ISH is a different phenomenon from this type of feature value sharing agree(ment). Hence, particle *-to* is not involved in agreement in the sense implied here.

syntactic processes as well as phonological/semantic interpretations) or a syntax-internal feature (only has a role in the syntax).

The [UC] feature does not directly fit into the distinctions posited by A&S. This feature is a first order feature since it distinguishes a class of features distinct from other classes. The interpretation of this feature is available at the interface; however, it is not interpretable on the particle itself (rather, this interpretation is accorded to the constituent it marks). Hence, it cannot be labelled as [$iUC_{spkr\text{-}sal}$] on particle *-to*. The type of feature valuation proposed in Wiltschko (2021) is non-canonical in the theoretical apparatus since the unvalued [ucoin] feature of $Ground_{Spkr}$ head is valued as positive or negative and not as the value of the [UC] feature of particle *-to*. If such an abstract type of Agree relation can be accommodated as a syntactic dependency, then the [UC] feature can be analysed as an interface feature (since it is used in a syntactic operation besides yielding a semantic interpretation at the interface).

In their system, "evidence that a given semantically interpretable feature is visible to syntax comes from data that shows that semantically interpretable feature triggers Merge, Agree, or Spell-Out operations" (A&S 2011:19). Within the current understanding of minimalism (Chomsky 2013, 2015, 2019), the syntactic operation of Merge is not feature-driven and is available for free in the computation (discussed in the section 5.4.2). Therefore, this parameter cannot be used as an evaluative rubric to determine the syntactic visibility of a semantically interpretable feature anymore. As for the second point, particle *-to* does not enter into a feature valuation via copying type of Agree relation (discussed in section 5.4.3). It enters into an abstract Agree relation as proposed in the ISH framework (that is distinct from the standard Agree operation). This section also exhibits that this particle does not have an impact on the Spell-Out operation since it does not do overt or covert movement. Thus, I conclude that the particle *-to*, with a semantically interpretable [UC] feature, is not relevant for the syntactic computation.

### *5.4.2. Merge of Particle -to*

Within the recent incarnation of the minimalist approach to language enquiry — Problems of Projection, henceforth POP (Chomsky 2013, 2015), the questions of compositionality, displacement and projection (i.e., labeling) take center-stage with linear order relegated to the externalization mechanism at the Articulatory-Perceptual (A-P) interface. The syntactic operations of Merge and a Labeling Algorithm (LA) are responsible for the compositionality-displacement issues and the projection issue respectively.

Merge is the simplest combinatorial operation made available by the Universal Grammar. It takes two syntactic objects X and Y to form a new syntactic object—the set {X, Y}. This is represented in (22). The label of this newly formed syntactic object is determined not by merge but by another operation — the LA. The label is a requirement of the Conceptual-Intensional (C-I) interface to semantically interpret a syntactic object. Crucially, merge is not triggered by the requirement to check some uninterpretable feature of a lexical item (like assumed in the previous theoretical frameworks) and is available for free in the computation. Simple merge is of two types — an External Merge operation (that takes two distinct syntactic objects to combine) and an Internal Merge operation (in this, one of the syntactic objects to be merged is a part of the other syntactic object i.e., X is a part of Y).

(22) Merge (X, Y) = {X, Y}

The Labeling Algorithm (LA) is a search operation within the minimal search domain for the possible label of a newly formed syntactic object. The mechanisms involved in this operation remove the requirement for labels to be an endocentric notion since now the label of an SO (syntactic object) {X,Y} is not restricted to the labels of X or Y. The algorithm proceeds as: if a strong head H merges with an XP i.e., {H, XP}, then the strong head projects. If an XP merges with another phrasal category YP i.e., {XP, YP}, then this syntactic object can be labeled via two methods. Either the two phrases share (through agreement) some prominent syntactic features (like phi features or Q–feature) in which case the prominent shared feature labels the SO. A second option is if one of the phrases moves out of the SO (via internal merge), the

remaining phrase projects its label. The LA, cited from Ginsburg (2016), is summarized in (23).

(23) a. When a strong head X is merged, the label is X.

b. If {XP,YP} share prominent features that are capable of labelling, the shared features label.

c. If YP moves out of {XP,YP}, the XP labels. If XP moves out of {XP,YP}, YP labels.

The POP model assumes that lexical roots and functional head T are weak elements that are incapable of labeling an SO. They can be strengthened (to be able to label the SO) via a mechanism of 'feature inheritance'. Feature Inheritance is a process by which the features of a phase head are passed onto the head of their complement projection. Thus, within a clausal syntactic derivation, the phase heads C and v pass on their phi features (the prominent syntactic features) to the T and lexical V heads they are merged to, respectively. These weak heads become strengthened after inheriting phi features and are able to label the SO they are a part of it.

Chomsky (2015), adopting a syntactic analysis of the notion of categories developed by Borer (2005, 2013) and Marantz (2008), assumes that the substantive lexical items are (unspecified) roots, whose categorical status is dependent on the subsequent merge of a functional category marker (like 'n', 'v' etc.). The merge of a root R with a category marker K results in the SO — {K, R}. The label of this SO is the label that the category marker K projects.

Deriving *to*-marked XPs in Hindi [30]

Particle -*to* is a lexical item with no syntactically prominent feature (like phi or Q feature) inherent to it. Thus, within the apparatus of the POP model, it is as a weak element that is incapable of labeling. Based on its empirical distribution (chapter 1), it is observed that this particle is a category-neutral LI that can mark nouns, adjectives,

[30] The analysis proposed here is the result of a series of discussions with Ayesha Kidwai (pc) and I gratefully acknowledge her contribution in formulating this analysis.

adverbs, verbs, postpositions etc.[31] Semantic interpretation-wise, a *to* marked noun is interpreted to be a noun, a *to*-marked adjective is interpreted to be an adjective, a *to*-marked adverb is interpreted to be an adverb and so on. Building on this insight, I propose that particle *-to* externally merges to a lexical root R before its category marker is merged to it. The so formed syntactic object is the set {R, *to*}. Both the members of this set are weak elements that cannot label the SO formed. This unlabeled SO is merged with a category marker K.

Following Ginsburg (2016) who, in his modeling of the POP framework, predicts that features other than phi or Q features are capable of labelling, I propose that the categorical feature of the category marker K is a feature that is capable of labelling a syntactic object. In the unmarked case where K merges with R i.e., {K, R}, the category feature of K labels this SO. In the marked case of a *-to* marked lexical root, the weak lexical root R inherits the category feature from K via the process of feature inheritance.[32] The strengthened root R now projects the label of the previously unlabeled syntactic object {R, *to*}. Assuming Ginsburg (2016), the label of an SO containing a strengthened element could be either the inherited feature or the label of the strengthened element itself. For expository ease, let us assume this SO to be labelled by the strengthened root. The next SO is {K, R}, which is then labeled as K. This derivation of a *-to* marked lexical root R (which is categorized as K) is sequentially given in (24a)-(24e). (24a) represents the External Merge of lexical root R and particle *–to*. (24b) exhibits the merge of category marker K to the unlabeled {R, *to*}. The feature inheritance from K to R is shown in (24c). The post strengthening

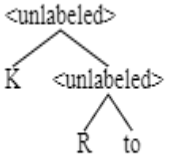
[31] Particle *-to* can mark all those LIs that have a conceptual or substantive content. Within the verbal domain, this particle can mark both the main lexical verb or the light verb in a V-V complex predicate construction (that have base lexical content but may be partially bleached in the light verb) but it cannot mark the auxiliary or the copula (no lexical content).

[32] Since v* (a verbal category marker) can pass its features to an embedded lexical verb in its domain (see Ginsburg (2016) for modeling of the derivation in the POP framework), I extend the phenomenon of feature inheritance as being permissible for all lexical category markers (‘a’ for adjective, ‘n’ for nominal, ‘adv’ for adverb, ‘p’ for preposition, ‘v’ for verb). While v* head passes its phi features (and not its categorical feature) to its complement, I assume that ‘a’ head, ‘n’ head, ‘p’ head and ‘adv’ head pass their categorical features to their complement lexical roots.

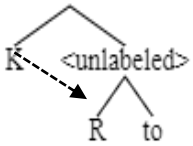
labeling of the unlabeled SO as R is given in (24d). (24e) shows the labeling of the second unlabeled SO as K.

(24a)

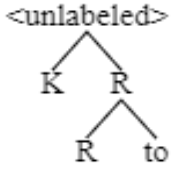

(24b)



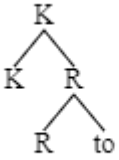

(24c)




(24d)




(24e)




On the basis of the proposed derivation, this analysis claims that the locus of attachment of particle *-to* in the syntactic spine is not the full phrase XP but rather locally to the X head. The phrasal structure of this X head starts building by the external merge of the internal argument (if available) after the labeled {X, *to*} structure has been merged with a category marker. The subsequent derivation after this point then proceeds as proposed in the POP framework.

Any lexical root marked by particle *-to* can be syntactically derived by employing the template outlined above. One such derivation ('nominal root-*to'*) is expanded upon in the discussion below. Other 'lexical root-*to'* combinations can be derived by

implementing the same procedure and I do not discuss each case separately to avoid redundancy.

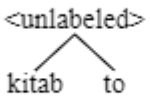
Derivation of N-*to*

This section discusses the derivation of a *to*-marked NP in the POP framework. The object NP *kitab* ('book') is marked by particle *-to* in example (25).

(25) ram=ne kitab=to pəɽʰɪ
Ram=ERG book=TO read.PFV.FSG
'Ram read the book.'

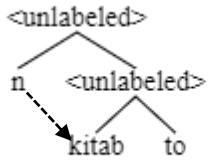
The lexical root *kitab* 'book' is merged with particle *-to* to form an unlabeled syntactic object {*kitab, to*} as shown in (26a).

(26a)



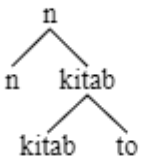

The nominal category marker 'n' is merged next. This categorizer passes its categorial [n] feature to the lexical root *kitab* via the process of feature inheritance (26(b)).

(26b)



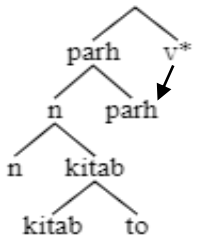

The strengthened root *kitab* projects its label for {*kitab, to*}and the nominal categorizer projects its label for the syntactic object {*n, kitab*} in (26c).

(26c)



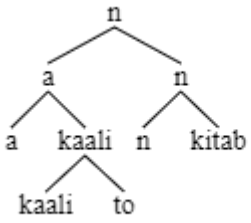
This SO is then merged with the lexical verb *parh* 'read'. This SO remains unlabeled until the merge of v* head that can pass on its phi-features to this weak root. (26d) exhibits the strengthened verbal root projecting its label for the syntactic object {n, *parh*} after feature inheritance.

(26d)




In the case of adjuncts like adjectives or adverbs being *to*-marked in Hindi, I adopt Chomsky (2004) in assuming that adjuncts are formed on a separate computational plane in the workspace and are merged to the syntactic spine via an asymmetric 'Pair Merge' operation. Pair Merge of two objects X and Y results in a partially ordered set <X,Y> where X is merged to Y and not vice-versa. Following Chomsky (2015), I assume that the label of a pair-merged syntactic objects <X, Y> is the label of its head Y and not its adjunct X. Thus, for a *to*-marked adjective modifying a nominal in Hindi (for example: *kaali-to kitab* 'black book'), I assume that particle -*to* merges with the root *kaala* to form an unlabelled SO, which is then merged with adjective category marker 'a'. After feature inheritance and labelling, the adjunct structure so formed is pair merged to the nominal category marker n (that contains the lexical root as complement). The nominal categorizer projects the label of this <a, n> structure. The final structure in (27)

(27)



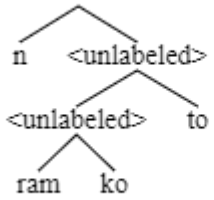
Derivation of N-case clitic-to

Example (28) exhibits a case marked object nominal being marked by particle *-to* in Hindi. In this case, the accusative (or differentially object marked 'DOM') case clitic *-ko* structurally intervenes between the lexical root *ram* and the particle *–to.*

(28) mɛ ram=ko=to mılı
I Ram=ACC=TO meet.PFV.FSG
'I met Ram.'

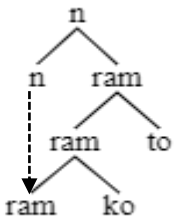
I propose l,m;kvjkgvmc ,*ram, ko*} is merged next with particle *-to*. The new syntactic object {{*ram, ko*}, *to*} is also unlabeled. The nominal category marker 'n' is merged next (29a).

(29a)




The lexical root inherits the categorical features from the nominal category marker and projects its label. The subsequent labeling of *ram-ko-to* (after feature inheritance) is represented in (29b).

(29b)




One consequence of this analysis is that the syntactic computation will build both 'ram-ko-to' and '*ram-to-ko' structures in parallel (since syntax does not restrict which particle or clitic will be merged to the lexical root first and which later). Since the ungrammatical sentence in example (30) does not crash in the proposed syntactic derivation (i.e., it does not crash in narrow syntax), the locus of this non-convergence is some place else.

(30) *mɛ ram=to=ko mɪlɪ
I Ram=TO=ACC meet.PFV.FSG
'I met Ram.'

As a tentative proposal, I suggest that the derivation in (30) crashes at the C-I interface. For any expression to receive an interpretation at the C-I, it must satisfy the interface conditions imposed by this performance system. I suggest that the example (30) crashes because it does not satisfy some specific interface condition requisite to receive interpretation. I suggest on such conception of a hypothetical interface condition that may explain the (non-)convergence of example (30). The C-I might require that the broader discourse-pragmatic interpretations (like some constituent being salient in the sentence from the speaker's perspective) can be accorded to a constituent only after its semantic interpretation is fixed in a sentence. Some Hindi case clitics have been proposed to be associated with a semantic constraint on the nominal it can mark, besides encoding a case feature (Aissen 2003, Butt and King 2004).[33] Since the particle *-to*, that has been proposed to signal the salience of the constituent it marks, cannot structurally intervene between a semantically relevant case clitic *-ko* and the nominal core in (30), it can be deduced that such discourse-pragmatic interpretation can be accorded to a constituent only after its semantic specification is complete. This suggests that the internal architecture of the C-I interface is granular and different types of interpretations are accorded to an expression in a sequence (with discursive interpretations being assigned after semantic interpretations). This tentative suggestion is theoretically explored further in chapter 6.

As an extension of this suggestion, hypothetically then focus particles, that are assumed to affect the truth conditional meaning of a sentence, should be structurally closer (in case of particle stacking) to the nominal core than the particles that encode

[33] Aissen (2003) assumes Hindi *-ko* to be sensitive to the 'specificity' or 'animacy' of the nominal; Butt and King (2004) propose that the Hindi accusative case clitic *ko* encodes a [+specific] feature; ergative case clitic *ne* encodes a [+volitional] interpretation. Such semantic associations have also been proposed for other languages. For example: Korean case clitics like *–i* and *–ka* encode a semantic restriction of picking out a specific entity in the discourse (see Kim 2013)

discourse-oriented use conditional meaning. This prediction is borne out by the grammaticality of minimal pair of sentences in examples (31) and (32) that involve a focus sensitive particle *-hii*. This particle functions as an exclusivity operator like 'only' (see section 1.4.2.1 in chapter 1). It adds a presupposition that the predicate is true exclusively for the *hi*-marked entity in the context and none other alternative n the discourse. Thus, this particle affects the semantic evaluation of the sentence and, assuming the tentative proposal sketched above to be on the right track, should be structurally closer than a discourse-pragmatic meaning encoder like particle *-to*. This order is manifested in the grammaticality of 'ram-hi-to' in example (31) and the ungrammaticality of 'ram-to-hi' in example (32).

| | | | | |
|---|---|---|---|---|
| (31) | ram=hi=to | kitab | pəɽʰta | hɛ |
| | Ram=HI=TO | book | read.IPFV.MSG | AUX.PRS.3SG |
| | 'Only Ram reads the book.' | | | |

| | | | | |
|---|---|---|---|---|
| (32) | *ram=to=hi | kitab | pəɽʰta | hɛ |
| | Ram=TO=HI | book | read.IPFV.MSG | AUX.PRS.3SG |
| | 'Only Ram reads the book.' | | | |

An interesting case in Hindi is the particle stacking order when a nominal is marked by all three — a case clitic, a focus-sensitive particle and discourse salience marker. The suggestion outlined above positions the discourse salience marker *-to* to be at the outer edge of the structure, with the other particles being associated closer to the nominal core. Example (16) exhibits that out of all the permutations of particle stacking in Hindi, the grammatical ones have particle *-to* as the outer edge of the nominal constituent. Thus, the CI interface might be modelled such that it requires a language to stack its markers or particles to mirror the semantic interpretative notions being closely associated to the lexical core than the discourse-pragmatic interpretative notions.

(33) ram=ne=hɪ=to kɪtab pəɽʰi
ram=hɪ=ne=to kɪtab pəɽʰi
* ram=ne=to=hɪ kɪtab pəɽʰi
* ram=hɪ=to=ne kɪtab pəɽʰi
* ram=to=ne=hɪ kɪtab pəɽʰi
* ram=to=hɪ=ne kɪtab pəɽʰi
'Only Ram read the book.'

To conclude, this section had proposed a POP model-based account for the merge of particle *-to* in the clausal spine. This particle externally merges to a lexical root and the category marker of that lexical root labels the syntactic object so formed via the process of feature inheritance and subsequent strengthening of the lexical root. A second idea developed in this section pertained to the particle stacking order in the nominal domain. It was proposed that the syntactic processes involved in structure building do not constrain the order of these particles. It was suggested that the CI interface imposes interface conditions on how an expression's interpretation is built–up. The ideas regarding the architecture of the C-I interface developed in this section are discussed further in chapter 6. Returning to the case of Hindi enclitic *–to*, the next section investigates the visibility of this particle for syntactic agreement and movement operations.

### *5.4.3. Syntactic Relevance of Particle -to*

This section focuses on those areas where marking of a constituent XP by particle *-to* could hypothetically have an effect on the syntactic computation of the clause containing that *to*-marked XP — agreement relations and movement operations. This section proposes that syntax is not sensitive to the presence or absence of particle *-to* in the clause.

Effect on Verbal and Nominal Agreement

Agreement phenomenon in Hindi has been a topic of extensive discussion in the literature (see Pandharipande and Kachru 1977, Mahajan 1990, Bhatt 2005, Bhatt and Keine 2017, Pareek 2017, Bhatia 2019 amongst others). Agreement, under the standard generative assumptions, is viewed as a syntactic dependency relation between a head (a probe) that has an unvalued phi-feature that requires valuation and another lexical item (a goal) that has the relevant matching feature valued on it and is capable of valuing the unvalued feature of the probe.[34] Following Chomsky (2000), I

[34] Phi features are the feature bundle comprising of [Person], [Num(ber)] and [Gen(der)] features

assume that 'agreement' is a relation (that a phi-probe enters into) and Agree is a computational operation (that implements both agreement and case-checking).[35]

The general facts of verbal agreement in Hindi are: the structurally highest non-overtly case-marked argument is the target of agreement (the 'Hindi-Urdu Agreement Generalization' proposed by Pandharipande and Kachru 1977). Bhatt (2005) proposes that Hindi has two probes — a finite probe (with unvalued [uPerson] and [uNum] features) and a participial probe with (with unvalued [uNum] and [UGen] features). In the clause, the T head has the finite probe with unvalued [uPerson] and [uNum] features and it probes for an argument (the goal) with their valued counterparts in its local (c-command) domain. The Asp(ect) head has the participial probe with unvalued [uNum] and [UGen] features.[36] Crucially, overt case-marking of an argument NP blocks it from being probed for agreement by the finite or participial probes in Hindi. Thus, in the unmarked case, subject NP agrees with and values the relevant phi features of the T head and the Asp head (this is the case in example (34)). On the other hand, an ergative case-marked subject NP with a non-marked object NP results in the verbal complex agreeing with the object NP in (35).[37]

(34) neha rəvi=ko piʈti hɛ
neha.F ravi.M=ACC hit.IPFV.FSG AUX.PRS.3SG
'Neha hits Ravi.'

(35) rəvi=ne billi dekʰɪ
ravi.M=ERG cat.F see.PFV.FSG
'Ravi saw a cat.'

[35] Chomsky (2000:101) conceptualizes Agree to be a language-specific operation of the $C_{HL}$ (the computational procedure of human language) that "establishes a relation (agreement, case-checking) between an LI α and a feature F in some restricted search space (its domain)".

[36] According to Bhatt, the participial probe is also manifested in the nominal domain as derived adjective can be analysed as participial form of the infinitival verb form. For example: *səɾ* 'to rot' (root)→ *səɾna* 'to rot' (infinitival form) → *səɾe hue kele* 'rotten apples' (participial form of adjective+noun).

[37] And in the case where both the subject and the object NP are case-marked, the verb exhibits default agreement pattern i.e., 3rd person, singular, masculine. This is exhibited in example (i).

(i) neha=ne sima=ko piʈa
neha.F=ERG sima.F=ACC hit.PFV.MSG
'Neha hit Seema.'

Example (36) exhibits that enclitic particle *-to,* unlike a case clitic, does not block the phi-probe of the T head or the Asp head on the argument NP it marks. Thus, particle *-to* is not a syntactic barrier to verbal agreement in Hindi. Its presence or absence does not affect the verbal agreement paradigm.

(36) neha=to rəvi=ko piʈti hɛ
neha.F=TO ravi.M=ACC hit.IPFV.FSG AUX.PRS.3SG
'Neha hits Ravi.'

Hindi exhibits agreement between an attributive adjective or a participial adjective and the noun within the nominal domain too. The adjective probes for its unvalued [uNum] and [UGen] features that are valued by the valued phi features of the nominal head. Examples (37a) and (37b) exhibits the nominal agreement paradigm on an attribute adjective that ends in /-aa/ i.e., *kaalaa* ('black').

(37a) mɛ=ne ek kali kɪtab kʰərɪdɪ
I=ERG one black.FSG book.FSG buy.PFV.FSG
'I bought a black book.'

(37b) mɛ=ne kʊʈʃʰ kale ʧəʃme kʰərɪde
I=ERG some black.MPL google.MPL buy.PFV.MPL
'I bought some black googles.'

Example (38) shows that marking of either the goal (38a) or the probe (38b) by particle *-to* has no effect on the nominal agreement relation.

(38a) mɛ=ne ek kali kɪtab=to kʰərɪdɪ
I=ERG one black.FSG book.F.SG=TO buy.PFV.FSG
'I bought a black book.'

(38b) mɛ=ne ek kali=to kɪtab kʰərɪdɪ
I=ERG one black.FSG=TO book.FSG buy.PFV.FSG
'I bought a black book.'

Based on these factors, I conclude that particle *-to* is not a barrier for nominal or verbal agreement relations in the Hindi clause structure.

Effect on movement

An LI that is visible to the syntactic movement operation could:

(i) Either trigger the movement of a constituent to some syntactic location

(ii) Or intervene in the movement of a constituent.

The empirical distribution of particle *-to* in example (39) indicates that this particle does not have a fixed designated position in the clausal spine (with elements moving around it) but rather it can (base) merge to a range of constituents across the spine.[38]

(39) a. ram=ne adʒ ek ʧʰoʈa kam kʰətəm kija hɛ
Ram=ERG today one small work finish do.PFV.3MSG AUX.PRS.MSG
'Ram has finished a small work today.'

b. ram=ne=(to) adʒ=(to) ek=(to) ʧʰoʈa=(to) kam=(to) kʰətəm=(to) kija=(to) hɛ=(*to)

This particle does not fit into the theoretical accounts (like Frascarelli and Hinterhözl 2007, Holmberg 2020 etc.) that assume morphological topic markers to move into a functional Top(ic) head position in the left periphery of the clause and the constituents that receive topic interpretation to move into the [spec,Top] position (see section 2.2.4 for a discussion of these models). The claim, that particle *-to* is not a morphological topic marker, was made based on arguments discussed in detail in chapters 2 and 3. To emphasize the dissociation of the discursive interpretation of topicality from this particle, example (40) revisits a scenario where this particle marks a constituent that does not receive the topic interpretation. Speaker A's sentence is of the form of *Tell me about X Test* (Neeleman and van de Koot 2008) introduced in chapter 2 that pre-designates the topical constituent 'Ravi'. Speaker B's response has a non-topical constituent 'Maths' being *to*-marked. If this particle was a functional Top(ic) head that licensed the topic interpretation in its specifier position in the syntactic spine, then the derivation of the *to*-marked sentence in (40) would not have converged.

(40) Context: A set of parents are checking up on their children's preparation regarding the upcoming school examinations and are talking amongst themselves.

[38] Particle *-to* can mark any one of the sentential units in example (22b) except the sentence-final auxiliary, that I assume has no lexical content that can be marked salient in the discourse.

Speaker A: mʊdʒʰe rəvɪ=ke bare mẽ bətao
I.DAT ravi=GEN about tell.IMP
'Tell me about Ravi.'

Speaker B: rəvɪ=ne mætʰs=to pəɽʰ lɪja
ravi=ERG maths=TO read take.PFV.MSG
'Ravi has studied Maths.'

Thus, based on the above two factors, I conclude that particle *-to* does not trigger movement of constituents in the syntax. Moreover, the POP framework (Chomsky 2013, 2015) discards the idea that movement is triggered by the properties (uninterpretable features) of an LI, as assumed in the previous versions of the minimalist theories (Chomsky 2004, 2008). In the POP system, an XP moves so as to be able to label an unlabeled syntactic object {XP, YP}. Once XP has moved out, this structure can be labeled as YP (see (23c)). Labels are relevant for semantic interpretation of a syntactic object at the interface. Although an unlabeled SO with no uninterpretable feature does not cause the derivation to crash but it is assumed to not receive an interpretation at the C-I interface (Ginsburg 2016). Within this framework, I have proposed that the syntactic object containing the particle *-to* i.e., {R, Prt} is labelled by the category marker that is merged next to this SO. Thus, there is no motivation to move either this particle or the lexical root that it merges to on grounds of requiring a labeled projection. In the following section, I investigate the question of whether this particle has an impact on a type of movement of the constituent that it can merge to.

Scrambling is a type of syntactic movement of an argument inside the clausal structure without having any theta-related motivation for movement.[39] Example (41) shows the leftward scrambling and the rightward scrambling of an object argument in a sentence. Example (41a) is the base SOV order sentence in Hindi. Example (41b) is the leftward scrambled OSV structure (the object moves to a position that is linearized as being leftward from its base position, which now has a co-indexed trace $t_i$ in it).

[39] Within the A versus A-bar movement dichotomy in syntax (where the former is motivated by argument-structure based notions like theta or case while the latter is not), scrambling is analysed to be a type of A-bar movement (see the literature cited in the main text for an expanded discussion on this).

Example (41c) is the rightward scrambled SVO structure (the object moves to a position that is linearized to the right of its base position).

(41) a. rəvɪ=ne seb kʰaja
ravi=ERG apple eat.PFV.MSG
'Ravi ate the apple.'

b. [seb]$_i$ ravɪ=ne $t_i$ kʰaja

c. ravɪ=ne $t_i$ kʰaja [seb]$_i$

Different analyses for the scrambling phenomenon in Hindi have been proposed by Mahajan (1990, 1997), Bhatt and Dayal (2007), Manetta (2014), Kidwai (1999, 2000, 2021) etc. I adopt the analysis proposed by Kidwai (2021) whose proposal accounts for the observed asymmetries in the leftward and rightward scrambled structures based on a slightly modified labeling algorithm within a POP framework. Leftward scrambling has been attested to have an effect on strictly C-I relevant semantic interpretations like the binding configurations (condition A, B, C violations), co-reference configurations like weak crossover (WCO) effect (a co-reference relation obtained in the case of crossing of a pronoun over its antecedent possessive pronoun (of some NP) within a clause being unlikely) and quantifier scope relations. These effects are not evidenced in the case of rightward scrambling. For instance, the example (42) from Kidwai (2021:3, ex (1)) exhibits the asymmetric impact of scrambling on WCO effect. The leftward scrambling of the object permits a grammatical co-referential configuration (violating WCO effect) in example (42b) but a rightward scrambled object does not improve the base WCO effect and is thus, ungrammatical with a co-referential relation as indicated in example (42c).

(42) a. uski$_i$ bɛhən [hər ləɽke=se]$_{*i}$ mɪlɪ
his sister each boy=SOC meet.PFV.FSG
'His$_i$ sister met each boy$_{*i}$.'

(42) b. [hər ləɽke=se]$_i$ uski$_i$ bɛhən $t_i$ mɪli
'His$_i$ sister met each boy$_{*i}$.'

(42) c. uski$_i$ bɛhən $t_i$ mɪli [hər ləɽke=se]$_{*i}$
'His$_i$ sister met each boy$_{*i}$.'

To account for this differential C-I impact, Kidwai (2021) re-introduces and operationalizes the Edge Feature (EF) postulated for phase heads within a POP system to syntactically encode a position that is relevant for the C-I interpretation i.e., it has an effect on the outcome at the CI interface. In her system, some v* phase heads have two probes— an EF probe and a u-phi probe—that are both capable of labelling syntactic objects. The u-phi probe is obligatorily transferred via feature inheritance to its complement V head but the EF feature is not transferred but remains on this v* head. A v* head with this optional EF feature has a second [spec,v*] position into which an object moves via internal merge. This position is the object shift position that has an effect on the C-I outcome. The SO created here receives the label of EF feature of v*. The derivation proceeds further with the object moving to the next phasal domain and eventually ending up in an unlabelled structure. Thus, the instances of this object form a chain of positions with the base copy position being the argument structure position (complement of lexical V). This object may optionally move through the labelled edge of v* where it can receive interpretations relevant for the C- interface and ends up at the highest unlabelled position where it is semantically vacuous. The instance of leftward scrambling that exhibit an effect on the C-I outcome necessarily move through this intermediate labelled [spec,v*] position where they are visible to the C-I interface. Rightward scrambling involves a v* head without any EF feature and thus, the object can only be interpreted at its base position (relevant for argument structure) or the highest moved copy position that is unlabelled and so is semantically vacuous. Only the base copy of a rightward scrambled object is visible to the C-I interface and thus, it has no effect on WCO effect or binding configuration relaxations etc.

Within this system, the syntax makes available an optional technical device (the Edge feature) that facilitates creation of a syntactic position with C-I relevance. Since this feature is optional on a verbal phase head, not all instances of leftward scrambling can be analysed to pass through the intermediate position in the object chain, where it can be evaluated for C-I related binding configurations etc. For example, Kidwai (2021:12) analyses example (43) to involve a verbal phase head without an EF feature. As a result, the object can be semantically interpreted only at its base position (where the binding conditions are licit i.e., the reflexive is structurally lower so that it

can be bound by the subject antecedent) and not at its semantically vacuous, highest copy position (where it is externalised at the sensory-motor interface).

(43) əpne ap=ko$_i$ ənu=ne$_i$ t$_i$ dek$^h$a
self=ACC Anu=ERG see.PFV.MSG
'Anu$_i$ saw herself$_i$.'

Kidwai (2021:14) notes that there is no correlation between the "semantically oriented labelling feature" (the EF) and "CG-management related IS-notions such as highlighting, correction, emphasis, contrast, frame-setting etc.". This argument gets support from particle *-to* marking of leftward scrambled object argument, such as in example (44). The object argument does not move through an EF designated [spec, v* position] in this sentence but it can still be marked by particle *-to* and thereby, convey being speaker salient in the discourse. The IS interpretation of topic is accorded to '*apne aap ko to'* as a combined result of three factors: the imposed salience; the clause initial position; the *to*-marked constituent having the denotation relevant for topic type interpretation (in this case, an entity in CG content). Depending on the availability of relevant alternatives in the discourse context, the *to*-marked scrambled object argument could receive a contrastive topic interpretation too.

(44) əpne ap=ko=to$_i$ ənu=ne$_i$ t$_i$ dek$^h$a
self=ACC=TO Anu=ERG see.PFV.MSG
'Anu$_i$ saw herself$_i$.'

Thus, leftward scrambled arguments, with or without moving through a syntactic position (that can have an effect on the CI-related semantic interpretation), can be marked by particle *-to*. However, rightward scrambled arguments cannot be marked by this particle in Hindi. Example (45) forms a minimal pair with example (41), the only difference being — the object argument *seb* 'apple' is marked by particle *-to* here. It is observed that, as expected, there is no intervening effect of particle *-to* on the leftward scrambling of the object in example (45b) but a *to*-marked object yields an ungrammatical sentence when rightward scrambled (example (45c)).

(45) a. rəvɪ=ne seb=to k$^h$aja
ravi=ERG apple=TO eat.PFV.MSG
'Ravi ate the apple.'

b. [seb=to]$_i$ ravɪ=ne $t_i$ k$^h$aja

c. *ravɪ=ne $t_i$ k$^h$aja [seb=to]$_i$

To account for this ungrammaticality, I suggest that there is another interface condition that could be stipulated that licenses the discourse salience interpretation to a constituent in a sentence. An entity is marked salient to indicate its relevance for the current utterance by attracting the interlocutor's attention to it (see section 4.2.3). From the perspective of parsing, the attention accorded to sentential constituents decreases from the sentence initial to the sentence final position. Therefore, it seems logically contradictory to mark something as salient and then move it to a sentence position that is marked to not be salient in the sentence. Thus, it is possible that the C-I interface imposes an interface condition that prohibits a marked salient entity (such as the *to*-marked entity) from occupying the sentence final position. Then, the ungrammaticality of example (45c) could be accounted for in terms of a possible violation of an interface condition stipulated for C-I relevant discourse salience interpretation.

Extending the ideas outlined above, I suggest that the interface requirements highlighted in this section are indicative of a possible internal architecture of the C-I interface. Since a non-*to*-marked rightward scrambled structure (41c) can receive both an argument structure interpretation, any permissible co-referential or binding-theoretic interpretation, this suggests that its *to*-marked counterpart (45a) would receive these interpretations too. The ungrammaticality of (45c) suggests that there is a separate module within the C-I interface where a sentence derivation could possibly crash because of some idiosyncratic requirement of this module (in this case, the requirement of salient marked entities not being permitted in non-salient positions). From this, it can be conjectured that the C-I interface is composed of distinct modules, each with its own prerequisite interface condition that evaluates the expression received and results in specific interpretative effects. One module could yield the argument structure related semantic interpretation for an input expression. Hypothetically then, arguments in their base position are visible to this argument-structure specific C-I module. Another articulated C-I module could be tentatively suggested that interprets the co-reference between entities, the scopal relations and the

binding of anaphora. The semantically relevant EF feature in the Kidwai system would then be relevant for this part of the C-I interface. A third tentative module regulates the discourse-related interpretations like discourse participants' attitudes or beliefs, discourse salience or the IS-relevant notions like topic, contrast or focus. This idea of a modular architecture of the C-I interface is very tentative in nature and is discussed again in chapter 6.

To summarize this section, I have proposed that particle *-to* is a lexical item with no syntactic relevance. It is composed of a phonological feature (that gives instruction for its externalization at the sensorimotor interface) and a semantically relevant interface feature — [$UC_{spkr\text{-}sal}$]. It has no syntactically relevant feature. Assuming a POP framework, this particle is externally merged to a lexical root, which is then labeled by the category marker selecting it. This particle is like a weak head that cannot label a syntactic object. This particle is invisible to the syntactic Agree operations and does not interfere with the nominal or the verbal agreement paradigm in Hindi. This particle does not trigger movement nor does it have an impact on the syntax of movement of constituents like in scrambling operation. The derivation of an expression cannot crash because of this particle within the narrow syntax but it can possibly crash at the C-I interface. This crash could be associated with a failure to satisfy some interface conditions that mediate the interpretation of such discourse particles. One interface condition was suggested which requires that the particles with a truth-conditionally relevant semantic interpretation be structurally closer to the lexical root core than the particles that manifest discourse-related interpretation. A tentative second interface condition was also suggested that requires for an entity to be overtly salient marked in the sentence, it should occur in a sentence final position. The sentence final position is, on grounds of parsing efficiency, structurally least prominent.

## 5.5. Conclusion

This chapter has adopted an atypical approach to analysing the syntax of particle *-to*. Since the empirical distribution of this particle is spread across category types and clause types, any proposed analysis should adequately account for both of them. I

have adopted two frameworks to incorporate these facts — the ISH and the POP. The ISH framework gives the freedom to generalize the syntactic association of this particle across declarative, interrogative and imperative clause types. The POP framework, as a research enterprise, has reduced the technology involved in a computational procedure to a minimal set of operations operating on restricted set of syntactic atoms.

To analyse the syntax of particle *-to* within the propositional structure, I adopted the POP framework. In this minimalist approach to the derivation of propositional expressions, I claim that the particle *-to* once merged in the derivation, has no syntactic role to play in the computational procedure. It is merged to any root that has a lexical content and that is selected by a category marker. It does not interfere with agreement relations nor does it affect movement operation.

For the syntax of particle *-to* with the interactional structure, I adopt the ISH framework to claim that the particle *-to* is configured by Universal Grammar too. This particle associates with a $\text{Grounds}_{\text{peaker}}$ functional head in the interactional spine and signals to assert that the proposition is available in the speaker's ground. The syntactic function of grounding is contingent on the presence of a relevant substantive LI in the propositional domain that can value its co-incidence feature. I have proposed that particle *-to* encodes a semantically relevant interface feature — the $[\text{UC}_{\text{spkr-sal}}]$ feature— whose semantic content is sufficient to value the unvalued feature of $\text{Ground}_{\text{speaker}}$ head.

This chapter has claimed that particle *-to* has no syntactic relevance within the propositional structure (which I model based on the POP framework) but it becomes syntactically relevant once the interactional structure of a grounding layer and a responding layer is syntactically manifested (according to the ISH framework). The syntax of this particle is derivationally simple, it is the C-I interface that introduces conditions on the semantic-pragmatic interpretations and has an effect on the convergence of a *to*-marked sentence in Hindi. A tentative proposal to posit multiple domains or modules in the CI interface, each with its own interface condition that

regulates the type of interpretation accorded to an expression, has been floated in this chapter. If this suggestion is on the right track, then such an enquiry into the internal structure of the CI interface pushes the linguist theory closer to discovering the ‘third factor’ conditions that shape the human language faculty.

# CHAPTER 6

# CONCLUDING REMARKS

In this chapter, section 6.1 tracks the development of the thesis proposal through the chapters 1 to 5. By investigating the enclitic particle *-to* in Hindi at the syntax-semantics-pragmatics interface, this thesis has raised theoretical questions about the direct correspondence between grammar and the information structure on one hand and the architecture of the interface of the faculty of language with the cognitive systems of thought on the other hand. Section 6.2 discusses these architectural design questions in light of the evidence adduced by this thesis.

## 6.1. Summary of the thesis

The literature review of Hindi enclitic particle *–to* in Chapter 1 had introduced it as an emphatic marker or a morphological topic marker for the language. Taking Kidwai (2000) as the starting point, this chapter had explored the range of empirical distribution of this particle beyond it marking the maximal projections of NP, VP and PP. It was exhibited that this particle can mark all lexical categories like adjectives, numerals, determiners, nouns, verbs, adverbs, postpositions etc. depending on the discourse context. Additionally, the non-occurrence of this particle with negation marker was taken as an evidence to propose that this particle can mark only those lexical categories that have some conceptual meaning. The distribution of this particle had exhibited that different types of pragmatic inferences (a contrastive or an 'at least' type inference) can be evoked by marking of an entity by this particle. The ordering relation of this particle with semantic operators like *–hii* and *–bhii* had exhibited that this particle interacts with the semantic interpretation accorded to a sentence and has an effect on the convergence of an expression.

Chapter 2 enquired into the topic marker status of this particle by unpacking what the information structure notion of topic means. After reviewing both the literature on information structure and the grammatical reflex of topic in a language, a conception of topics (an entity in CG, Krifka 2008) and a methodological tool (the diagnostic tests for topichood) were adopted to diagnose the topical constituents in a sentence.

Evidence against the topic marker analysis of particle *–to* was provided in the form of the different syntactic behavior of this particle from canonically assumed topic markers — Japanese *–wa* and Korean *–nun* particle. Secondly, it was exhibited that both unmarked and *to*-marked constituents can receive a topic interpretation at the interface, which raised a design issue of economy in the language apparatus. A pragmatic requirement unique to the *to*-marked constituents was adduced in the form of the discourse-old or given status necessary for a constituent before it can be marked by particle *–to*. Since topics can be discourse-new too, this pragmatic licensing constraint cannot be accounted for within a topic marker analyses of particle *–to*.

Chapter 3 strengthened the argument for dissociating topicality as an inherent property of this particle by exhibiting that this particle can mark both — non-topical constituents (like adjective, numerals etc.) as well as topical candidates that do not receive a topic interpretation. To claim this, the biased position for a topic interpretation was diagnosed to be the sentence initial position. Testing the semantic-pragmatic import of this particle when it marks a non-topical, in situ object NP, it was exhibited that certain implicatures are evoked in these configurations. Since these inferences were contingent on the availability of an alternatives set and they expressed a juxtaposition of discourse status of the entities, it was concluded that these are contrastive implicatures that facilitate a contrastive interpretation for these *to*-marked entities. These uncertainty and scalar implicatures were analysed to be conversational implicatures that are calculable and cancellable. Crucially, this contrastive interpretation is not directly linked to particle *–to* since it depends on the pragmatic factors as well as syntactic configuration. Thus, this chapter had concluded that the enclitic particle *–to* is not a topic marker *per se* since it does not directly correspond to the IS notion of topicality. The topic interpretation, the contrastive interpretation or the compositionally built-up contrastive topic interpretations are derived interpretative effects for this particle.

Chapter 4 delved into the meaning core of this particle and exhibited that this particle does not affect the truth conditions of the sentence nor does it contribute to the at-issue content encoded by a sentence. It was claimed that this particle signals discourse salience or prominence on the constituent it marks and, assuming multidimensional

models of meaning, this was formalised as a type of procedural or instructional or use conditional meaning that is lexically encoded by this particle. Since emphasis and contrast are the two motivations because of which a speaker can mark an entity as salient, the interpretative effects of topicality and contrast were proposed to be the interpretative effects of increasing the salience of a constituent. This particle was claimed to be a use conditional item that exhibits the properties typical to UCIs, such as independence, descriptive ineffability and perspective dependence. This chapter adopted the Hybrid Semantics framework to account for the semantic contribution of a to-marked sentence as a tuple of truth conditional content (formally modeled as a set of worlds) and use conditional content (modeled as a set of contexts).

For the syntax of particle *–to*, chapter 5 had claimed that this particle is not relevant for the syntactic derivation within the propositional domain (no intervention in verbal or nominal agreement or movement operations) but it is relevant for the extended interactional domain in the Interactional Spine Hypothesis framework (Wiltschko 2021). The meaning of speaker salience conveyed by this particle was operationalized as a [UC] feature that can value a functional $\text{Ground}_{\text{spkr}}$ head (that is a syntactic representation of the speaker's mental ground). By uttering a *to*-marked sentence, it is claimed that the speaker asserts that they hold a propositional attitude towards the proposition specifically relative to the salient marked entity. This accounted for the syntactic distribution of particle *–to* in declarative and imperative clauses as well as confirmational questions. This chapter had also proposed a Problems of Projection (Chomsky 2013, 2015) based analysis for the merge of particle *–to* in the syntactic spine. This particle was claimed to be a weak head that merges to a lexical root, which gets strengthened by the next merged category marker so that it can project the label of the category marker. The order of merge of particles was proposed to not be relevant for the narrow syntax within this framework but was relevant for receiving semantic interpretation at the C-I interface. Similarly, a sentence-final constituent being *to*-marked was proposed to not be problematic for the syntax *per se* but that lead to a crash of the derivation at the interface.

To conclude, this thesis had argued for the particle *-to* to be a marker of discourse salience (and not a topic marker) that procedurally or use-conditionally encodes the

meaning of conveying speaker salience on the *-to* marked entity. This increased salience of an entity facilitated its interpretation as being either the topical constituent of the sentence (if it was a topical candidate and in the clause initial position) or being in a contrastive relation with the other alternatives (subject to the availability of this set of alternatives in the discourse). Thus, the IS notions of topic and contrast are pragmatic interpretative effects that were mediated via particle *-to*-induced increase in salience of a constituent in the speaker's discourse model. This particle, by functioning as a morphological prominence-lending cue, was involved in the function of common ground management. Expanding the common ground to include not only the discourse referents and the propositions but also the individual interlocutor's propositional attitude, the ISH framework had provided a syntactic manifestation of the speaker's ground or mental state, where the speaker salience marked by the particle *-to* could be incorporated. This particle had been argued to not be relevant for the narrow syntactic derivation but to have an impact on the interpretation at the C-I interface.

### 6.2. Design Implications of the proposal for the Faculty of Language (FL)

The proposal outlined in section 6.1 has direct implications for the architecture of the Faculty of Language as assumed in Chomsky (2000, 2001, 2004, 2005, 2008, 2013, 2015, 2019).[1]

Any current research with minimalist goals aims to prove the Strong Minimalist Thesis (SMT) — that "language is an optimal solution to interface conditions that FL must satisfy" (Chomsky 2005:3) — or at least a weaker version of it. Under the assumption that SMT holds, any property of a language must have a 'principled' explanation in terms of either the interface conditions imposed by the external systems interacting with the FL or in terms of the 'third factor' conditions (general

[1] Within the minimalist tradition, FL consists of a language-internal module (that comprises of a lexicon and a generative procedure that operates on finite syntactic atoms from the lexicon to derive infinite expressions that map onto a sound-meaning pairing <PHON,SEM>) that interfaces with cognitive interpretive systems (the sensorimotor (SM) system that interprets <PHON> and the conceptual-intensional (C-I) system that interprets <SEM>).

properties like computational efficiency) that hold for acquisition of a language. The interface condition operative at C-I interface requires an expression to have a multiplicity of semantic properties — the argument structure related properties ("theta-theoretic" properties), the scopal properties and the discourse-related properties (for new v/s old information distinction or specificity etc.). Chomsky (2004:110) reduces this multiplicity to a duality i.e., the "argument structure and everything else". The expression derived by narrow syntax has to provide basis for this duality of interpretation at the C-I interface. This was theoretically implemented by positing a duality of syntactic operations (external and internal merge) that yield this duality of semantic interpretations (Chomsky 2008). The interface interpreted the theta-theoretic property of an expression based on the external merge positions of the constituents in the narrow syntax. The 'other' semantic properties of an expression were contingent on the internal merge position of a syntactic object.

Having minimalist aspirations at the core, this thesis also assumes the SMT to be on the right track and conjectures that *-to* marking of a constituent in Hindi can be analysed as attributing a semantic property to the expression containing the *-to*-marked constituent. This semantic property has a 'principled' explanation in terms of the interface conditions imposed by the systems of thought. This thesis has claimed that particle *-to* yields observable discourse-related effects of topicality and contrast (derived via imposed discourse salience on an entity). Assuming the duality of interpretation to correspond to duality of syntactic operations, these discourse related interpretations should have been licensed via the internal merge of the *-to*-marked constituents to some higher clausal position. However, there is no theory-independent motivation or evidence for obligatory movement of *a -to*-marked entity to a clause internal position and it can remain in-situ in the structure. Thus, *prima facie*, this linguistic phenomenon of *-to*-marking is not tenable within a Chomsky-style FL architecture that strictly maps a duality of syntactic operations onto a duality of semantic interpretations.

Tentatively, the theoretical goal of SMT can be maintained if the collapse of multiplicity of semantic properties into a duality (cf. Chomsky 2004) is revoked and the discourse-related semantic properties are treated as separate interface conditions

from the non-theta-theoretic scopal properties of an expression. Within this reformulated categorization of semantic interface conditions, scopal properties can be interpreted on the basis of internal merge configuration (as originally proposed in the theory) but discourse-related properties can be fed via two structural configurations — either by internal merge operation or by the presence of a discourse particle (with an interpretable interface feature) that are base merged at some position in the hierarchical structure. A constituent that has moved to clause-initial position and that may receive a sentence topic interpretation (if feasible by discourse context considerations) exhibits the first type of interpretive mapping of an expression by the interface. Particle *-to* marking of a constituent can be analysed as the second type of mapping of an expression to its discourse-related semantic property, that is usable for the C-I interface.

The distinct trinity of semantic properties (theta-related, scopal and discourse-related) lies at the base of the tentative modular architecture of the C-I interface suggested in sections 5.4.2 and 5.4.3., where each type of semantic property of an expression is interpreted by a distinct component of the C-I interface.[2] Since the early formulations

[2] A similar but conceptually different tripartite division of the interface of the computation with the systems of thought has been postulated in Platzack (2000) and Grohmann (2004). Platzack (2000) proposes that instead of one interface level LF of the computation with the cognitive faculties (Chomsky 1995), the computation interacts dynamically with the interpretive systems at three interface points in derivation. The theta-structure related information encoded by the VP domain is handed over at a Thematic Form (TF) interface. The grammatical aspects of the clause like tense and aspect specification etc. present in the IP domain is handed over for interpretation at a Grammatical Form (GF) interface. The discourse related information that links the proposition to the discourse is encoded by the CP domain and a third interface level of Discourse Form (DF) feeds into the interpretive system. TF, GF and DF are syntactic structures that interface with the cognitive system of thought. Grohmann (2003a: 75) introduces the concept of prolific domains to refer to a "contextually defined part of the computational system…that provides the interfaces relevant to the context and… consists of internal structure, interacting with derivational operations." He posits vP to form a Θ-domain (locus of thematic relations), TP to correspond to a Φ-domain (locus of agreement properties) and CP to form a Ω-domain (locus of discourse information). In this system, as soon as a domain is completed, the spell-out operation applies to feed the encoded interpretations to SM and CI interface. The common denominator in these two proposals is that the three types of semantic information are ascribed to three chunks of the syntactic derivation with the internal structure of the CI interface left unexplored, assuming the systems of thought to have access to the syntactic derivation dynamically. Discourse related information is

of the minimalist enquiries into language, it has been clear that "the external systems are not well understood" and that "progress in understanding them goes hand in hand with progress in discovering the language systems that interact with them" (Chomsky 2000:98). This thesis can be read as an exercise in understanding the internal structure of the CI interface better. If the proposed semantic and syntactic analysis of particle -*to* is on the right track, then the source of problem in the cases pertaining to particle stacking on an entity (see section 5.4.2) or -*to* marking of a rightward scrambled entity (see section 5.4.3) is neither a theta-theoretic nor a scopal interpretation-related issue but only a discourse interpretation-related issue.

The nature of two possible interface conditions of the discourse module of the CI interface was suggested in chapter 5. If a tentative interface condition can be postulated to be operative at the C-I interface that specifically regulates that the interpretation of a discourse-related semantic property of an expression can only be assigned after its other semantic properties (theta-theoretic or scopal) are assigned, then that could explain why an entity cannot be marked by a discourse particle before its semantic specification is complete.[3] An expression within which an entity is marked by the discourse particle -*to* before it can be marked by other semantically relevant particles (like a case clitic or focus particles) is not 'usable' at the interface since it potentially violates the hypothetical interface condition suggested previously. This can explain the deviance or the crash of the derivation of such an expression with a particle stacking order not reflecting the order of the interface modules.

---

provided by the CP linked DF or Ω-domain. In contrast to these analyses, I am suggesting an 'internal structure' to the interfacing systems of thought without relegating discourse information to the CP domain. The modules of the interface may be relativized to attributing distinct types of semantic interpretation. For the part of interface that accords discourse-related information, discourse particles like Hindi enclitic -*to* need not obligatorily move to the CP domain. In my conception, this particle can receive discourse-related interpretations at its point of merge because the interface module specific to discourse properties can accord an interpretation like discourse salience locally to a constituent

[3] Such a specific interface condition is in contrast to a general interface condition like that of 'Full Interpretation' that requires that "all terms of a syntactic object must be interpreted, none can be ignored" (Chomsky et al. 2019:14).

A second tentative interface condition, that was suggested in section 5.4.3, dealt with the specific nature of the semantic interpretation of the concept of discourse salience. A rightward scrambled constituent cannot be marked as salient in the sentence. A possible explanation for this could be that left and right peripheries have been evidenced to play different roles in the perceptual and parsing considerations of the language design (Chomsky 2013:41). An increase in salience can be conceptually linked to an ease in perception and parsing since the salient entities are (psycho-cognitive) attention-attractors. Thus, an overtly marked salient entity can logically receive its discourse-salience interpretation either in the left periphery or in its base merge position. Thus, a possible interface condition that licenses a discourse-salience interpretation to a marked entity in an expression requires that this entity should not surface in the logically least salient position i.e., in the right periphery of the clause. This can possibly explain the deviance or crash of an expression containing a *-to-* marked rightward scrambled constituent.

To conclude, by making explicit the design implications of the proposal, this thesis has attempted to engage in a theoretically 'interactive' research task, one that tries "to clarify the nature of the interfaces and optimal computational principles through investigation of how language satisfies the conditions they impose — optionally, insofar as SMT holds" (Chomsky 2005:3). The proposed interface conditions in this thesis are preliminary in nature and have a very tentative character, but are nevertheless useful starting points in a finer-grained modeling of the performance systems and the bare output conditions they impose on FL.

## BIBLIOGRAPHY


Abraham, W. (2007). Topic, focus and default vs. contrastive accent. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 183–206). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Adger, D., & Svenonius, P. (2011). Features in minimalist syntax. *The Oxford handbook of linguistic minimalism*, *1*, 27-51.

Agnihotri, R. K. (2007). *Hindi: An essential grammar*. Routledge.

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, *21*(3), 435-483.

Aissen, J. L. (1992). Topic and focus in Mayan. *Language*, 43–80.

Aoun, J., & Benmamoun, E. (1998). Minimality, reconstruction, and PF movement. *Linguistic Inquiry*, *29*(4), 569-597.

Ariel, M. (1990). Accessing antecedents. London: Routledge.

Atlas, J. D., & Levinson, S. C. (1981). It-clefts, informativeness and logical form. In *Radical pragmatics* (pp. 1-62). Academic Press.

Austin, J. L. (1962). Speech acts.

Bajaj, V. (2014). Hii at the semantics-pragmatics interface. In *Proceedings of the 31st West Coast Conference on Formal Linguistics, Cascadilla Press* (Vol. 31).

Bajaj, V., & Syrett, K. (2016). -hii: Modality Meets Exclusivity. *University of Pennsylvania Working Papers in Linguistics*, *22*(1), 3.

Bak, S. Y. (1977). Topicalization in Korean. *University of Hawaii Working Papers in Linguistics*, *9*(2), 63-88.

Bak, S. Y. (1984). Why not a topic in a relative clause?

Baltin, M. R. (1982). A landing site theory of movement rules. *Linguistic inquiry*, *13*(1), 1-38.

Bayer, J. (1996). Directionality and Logical Form. On the scope of focussing particles and Wh-in situ.

Bayer, J. (2008). Clause structure in South Asian languages (review). *Language*, *84*(3), 597-612.

Bayer, J. (2018). *A Comparative Note on the Bangla Particle" to" and the German Particle" doch"*.Konstanz-Online Publications.

Bayer, J. (2018). *Criterial freezing in the syntax of particles* (pp. 225-263). De Gruyter Mouton.

Bayer, J. (2020). *A Comparative Note on the Bangla Particle" to" and the German Particle" doch"* (pp. 17-31).

Bayer, J., & Struckmeier, V. (Eds.). (2017). *Discourse particles: Formal approaches to their syntax and semantics* (Vol. 564). Walter de Gruyter GmbH & Co KG.

Bayer, J., Dasgupta, P., Mukhopadhyay, S., & Ghosh, R. (2014). Functional Structure and the Bangla Discourse Particle to. *Retrieved July*, *19*, 2015.

Beaver, D. (1994). Accomodating Topics. Abraham, W. (2007). Topic, focus and default vs. contrastive accent. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 183–206). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bezuidenhout, A. (2004). Procedural meaning and the semantics/pragmatics interface. *The semantics/pragmatics distinction*, 101-131.

Bhatia, S. (2019). Computing agreement in a mixed system.

Bhatt, R. (2005). Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory*, *23*(4), 757-807.

Bhatt, R., & Dayal, V. (2007). Rightward scrambling as rightward remnant movement. *Linguistic Inquiry*, *38*(2), 287-301.

Bhatt, R., & Dayal, V. (2020). Polar question particles: Hindi-Urdu kya. *Natural Language & Linguistic Theory*, *38*(4), 1115-1144.

Bhatt, R., & Keine, S. (2017). Long-Distance Agreement. *The Wiley Blackwell Companion to Syntax, Second Edition*, 1-30.

Bianchi, V., & Frascarelli, M. (2010). Is topic a root phenomenon?

Biezma, M., & Rawlins, K. (2017, December). Rhetorical questions: Severing asking from questioning. In *Semantics and linguistic theory* (Vol. 27, pp. 302-322).

Binding. *Linguistic Inquiry* 24:461-507.

Birner, B. J., & Ward, G. (2009). Information Structure and Syntactic Structure. *Language and Linguistics Compass*, *3*(4), 1167–1187.

Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers* (Vol. 99). Cambridge university press.

Blome-Tillmann, M. (2013). Conversational implicatures (and how to spot them). *Philosophy Compass*, *8*(2), 170-185.

Bocci, G. (2007). Criterial positions and left periphery in Italian: Evidence for the syntactic encoding of contrastive focus. *Nanzan linguistics*, *1*(3), 35-70.

Bolinger, D. (1965). *Forms of English: Accent, morpheme, order*. Harvard University Press.

Borer, H. (2005). *Structuring sense: Volume 1: In name only* (Vol. 1). Oxford University Press.

Borer, H. (2013). *Structuring sense: volume III: taking form* (Vol. 3). Oxford University Press.

Bošković, Ž. (2004). Topicalization, focalization, lexical insertion, and scrambling. *Linguistic Inquiry*, *35*(4), 613-638.

Boswijk, V., & Coler, M. (2020). What is Salience?. *Open Linguistics*, *6*(1), 713-722.

Bresnan, J., & Kanerva, J. M. (1989). Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic inquiry*, 1-50.

Bresnan, J., & McHombo, S. A. (1987). Topic, Pronoun, and Agreement in Chichewa. *Language*, *63*(4), 741–782. https://doi.org/10.2307/415717

Brown, P., & Levinson, S. C. (1978). Universals in language usage: Politeness phenomena. In *Questions and politeness: Strategies in social interaction* (pp. 56-311). Cambridge University Press.

Büring, D. (1997). *The meaning of topic and focus: the 59th Street Bridge accent* (Vol. 3). Psychology Press.

Buring, D. (1999). Topic. In *FOCUS Linguistic, Cognitive, and Computational Perspectives* (pp. 142–165). Cambridge University Press.

Büring, D. (2016). *Intonation and meaning*. Oxford University Press.

Büring, Daniel (1997). 'The Great Scope Inversion Conspiracy', *Linguistics & Philosophy* 20: 175–194.

Butt, M. (2010). The light verb jungle: Still hacking away. *Complex predicates in cross-linguistic perspective*, 48-78.

Butt, M., & King, T. H. (1996). Structural topic and focus without movement. *Online Proceedings of LFG*.

Butt, M., & King, T. H. (2003). Case systems: Beyond structural distinctions.

Butt, M., & King, T. H. (2004). The status of case. In *Clause structure in South Asian languages* (pp. 153-198). Springer, Dordrecht.

Butt, M., & King, T. H. (2014, July). Questions and information structure in Urdu/Hindi. In *Proceedings of the LFG14 Conference*. Stanford: CSLI Publications.

Butt, M., Bögel, T., & Jabeen, F. (2017). Polar kya and the prosody-syntax-pragmatics interface. In *Proceedings of the LFG'17 Conference* (pp. 125-145).

Chafe, W. L., & Li, C. N. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View in Subject and Topic.

Chiarcos, C. (2009). *Mental salience and grammatical form: toward a framework for salience metrics in natural language generation* (Doctoral dissertation).

Chiarcos, C. (2011). On the dimensions of discourse salience. *Bochumer Linguistische Arbeitsberichte*, *3*, 31-44.

Chiarcos, C., Claus, B., & Grabski, M. (Eds.). (2011). *Salience: Multidisciplinary perspectives on its function in discourse* (Vol. 227). Walter de Gruyter.

Chierchia, G. (1993). *Extension of Dynamic Binding: Reconstruction, Topicalization and Crossover*. University of Milan.

Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and grammar: An introduction to semantics*. MIT press.

Choe, H. S. (1995). Focus and topic movement in Korean and licensing. *Discourse configurational languages*, *269*, 334.

Chomsky, N. (1965). Persistent topics in linguistic theory. *Diogenes*, *13*(51), 13-20.

Chomsky, N. (1986). *Barriers* (Vol. 13). MIT press.

Chomsky, N. (1993). *Lectures on government and binding: The Pisa lectures*. Walter de Gruyter.

Chomsky, N. (1995). *The minimalist program* (Vol. 28). Cambridge University Press.

Chomsky, N. (1999). *Derivation by phase* (Vol. 18, pp. 1-52). MIT, Department of Linguistics.

Chomsky, N. (2000). Minimalist inquiries: The framework (MITOPL 15). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-155.

Chomsky, N. (2001). 'Derivation by phase'. In M. Kenstowicz (ed.) Ken Hale: A Life in Language. Cambridge, Mass: The MIT Press.

Chomsky, N. (2007). Of minds and language. *Biolinguistics*, *1*, 009-027.

Chomsky, N. (2008). On phases. *Current Studies in Linguistics Series*, *45*, 133.

Chomsky, N. (2013). Problems of projection. *Lingua*, *130*, 33-49.

Chomsky, N. (2015). Problems of projection: Extensions. *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, *223*, 1-16.

Chomsky, N. (2019) "UCLA Lectures on Linguistic Theory," delivered at UCLA from April 29 to May 2, 2019.

Chomsky, N., Gallego, Á. J., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*, 229-261.

Cinque, G. (1997). "Topic" Constructions in some European Languages and "Connectedness." In *Materials on Left Dislocation* (Vol. 14, pp. 93–118). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press.

Clamons, C. R., Mulkern, A. E., & Sanders, G. (1993). Salience signaling in Oromo. *Journal of Pragmatics*, *19*(6), 519-536.

Cohen, A. (2007). Bare plurals as topics in English and Italian. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 509–522). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Cohen, A., & Erteschik-Shir, N. (2002). Topic, focus, and the interpretation of bare plurals. *Natural Language Semantics*, *10*(2), 125–165.

Coppock, E., & Champollion, L. (2019). *Invitation to formal semantics*. Manuscript, Boston University and New York University.

Cresti, D. (1995, July). *Indefinite Topics* (Ph.D.). Massachussets Institute of Technology.

Cruse, A. (2004). Meaning in language: An introduction to semantics and pragmatics.

Dahl, O. (1974). Topic-comment structure revisited. In *Topic and Comment, Contextual Boundedness and Focus*. Hamburg: Buske Verlag.

Dahl, Ö. (1975). *Journal of Linguistics*, *11*(2), 347–354.

Dalrymple, M., & Nikolaeva, I. (2011). *Objects and Information Structure*. New York: Cambridge University Press.

Daneš, F. (Ed.). (1974). *Papers on functional sentence perspective* (Vol. 147). Academia.

Dash, B., Datta, M. & Simpson, A. (2021, March). *Analyzing Hindi & Bangla -to: A case for Anti-Exhaustivity as a Discourse Category* [Conference Presentation]. FASAL-9, University of Minnesota.

Davison, A. (1984). Syntactic Markedness and the Definition of Sentence Topic. *Language*, *60*(4), 797–846. https://doi.org/10.2307/413799

Davison, A. (2004). Non-nominative subjects in Hindi-Urdu VP structure and case parameters. *TYPOLOGICAL STUDIES IN LANGUAGE*, *60*, 141-168.

Davison, A. (2004). Structural case, lexical case and the verbal projection. In *Clause structure in South Asian languages* (pp. 199-225). Springer Netherlands.

Dayal, V. (2011). Hindi pseudo-incorporation. *Natural Language & Linguistic Theory*, *29*(1), 123-167.

Dayal, V. (2019, October 31). *The Multiple Faces of Hindi-Urdi bhii* [Conference Presentation]. SALA-35, INALCO, Paris.

Deo, A. (2021). Could be stronger: Raising and resolving questions with Hindi= to. Manuscript.

Diesing, M. (1992). Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry*, *23*(3), 353-380.

Dik, S. C. (1997a). *The theory of functional grammar: the structure of the clause*. Walter de Gruyter.

Egg, M. (2013). Discourse particles, common ground, and felicity conditions. In *Beyond expressives: Explorations in use-conditional meaning* (pp. 125-149). Brill.

Enrico, J. (1986). Word Order, Focus, and Topic in Haida. *International Journal of American Linguistics*, *52*(2), 91–123.

Erteschik-Shir, N. (1986). Wh-Questions and Focus. *Linguistics and Philosophy*, *9*(2), 117–149.

Erteschik-Shir, N. (1997). *The dynamics of focus structure*. Cambridge University Press.

Erteschik-Shir, N. (2006a). 16.1 What is the status of gradience?. In *Gradience in grammar: Generative perspectives*, 317.

Erteschik-Shir, N. (2006a). On the architecture of topic and focus. In *The architecture of focus*. De Gruyter Mouton, 33-58.

Erteschik-Shir, N. (2007). *Information structure: The syntax-discourse interface* (Vol. 3). Oxford University Press.

Erteschik-Shir, N. (2013). Information Structure and (in) definiteness. In *Crosslinguistic Studies on Noun Phrase Structure and Reference* (pp. 23–51). Brill.

Falk, S. (2014). On the notion of salience in spoken discourse-prominence cues shaping discourse structure and comprehension. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, (30).

Fanselow, G., & Lenertová, D. (2011). Left peripheral focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory*, *29*(1), 169–209.

Farkas, D. F., & Bruce, K. B. (2010). On reacting to assertions and polar questions. *Journal of semantics*, *27*(1), 81-118.

Fery, C. (2007). The prosody of topicalization. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 69–86). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Féry, C. (2016). *Intonation and prosodic structure*. Cambridge University Press.

Féry, C., & Ishihara, S. (Eds.). (2016). *The Oxford handbook of information structure*. Oxford University Press.

Fintel, K. von. (1995). *Presupposition Accomodation and Quantifier Domains : Comment on Beaver's "Accomodating Topics."* MIT.

Fintel, K. von. (2008). What is Presupposition Accommodation, Again?*. *Philosophical Perspectives*, *22*(1), 137–170. https://doi.org/10.1111/j.1520-8583.2008.00144.x

Firbas, J. (1962). Notes on the function of the sentence in the act of communication. *Sbornik praci filosoficke fakulty brnenske university A 10.133—48 (Brno)*.

Firbas, J. (1962). Notes on the function of the sentence in the act of communication. *SFFBU, 10,* 133-148.

Firbas, J. (1964). On defining the theme in functional sentence perspective. *Travaux linguistiques de Prague, 1,* 267-280.

Firbas, J., & Jan, F. (1992). *Functional sentence perspective in written and spoken communication*. Cambridge University Press.

Fortuny, J. (2008). *The emergence of order in syntax*. Benjamins.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics* (pp. 71-120). Palgrave Macmillan, London.

Frascarelli, M. (2007). Subjects, topics and the interpretation of referential pro. *Natural Language & Linguistic Theory*, *25*(4), 691-734.

Frascarelli, M., & Hinterholzl, R. (2007). Types of topics in German and Italian. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 87–116). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Frey, W. (2004). Notes on the syntax and the pragmatics of German Left Dislocation. In *The syntax and semantics of the left periphery* (pp. 203-234). De Gruyter Mouton.

Frey, W. (2010). Ā-Movement and conventional implicatures: About the grammatical encoding of emphasis in German. *Lingua*, *120*(6), 1416-1435.

Gallego, Á. J. (2013a). A configurational approach to the left periphery. In *23rd Colloquium on Generative Grammar, Universidad Complutense de Madrid, May* (pp. 9-11).

Gazdar, G. (1978). Pragmatics, implicature, presupposition and logical form. *Critica*, *12*(35).

Gazdar, G. (1979). A solution to the projection problem. In *Presupposition* (pp. 57-89). Brill.

Gécseg, Z., & Kiefer, F. (2009). A New Look at Information Structure in Hungarian. *Natural Language & Linguistic Theory*, *27*(3), 583–622.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Ginsburg, J. (2016). Modeling of problems of projection: A non-countercyclic approach. *Glossa: a journal of general linguistics*, *1*(1).

Ginzburg, J., & Sag, I. (2000). *Interrogative investigations*. Stanford: CSLI publications.

Givón, T. (1983). Topic continuity in spoken English. *Topic continuity in discourse: A quantitative cross-language study*, *3*, 347-363.

Givón, T. (1992). The grammar of referential coherence as mental processing instructions.

Göbbel, E., & Breul, C. (2010). *Comparative and Contrastive Studies of Information Structure*. Amsterdam: John Benjamins Publishing Co. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=340314&site=eds-live&scope=site

Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41-58). Brill.

Grice, H. P. (1978). Further notes on logic and conversation. In *Pragmatics* (pp. 113-127). Brill.

Grice, H.P. (1989). *Studies in the Way of Words*. Harvard University Press.

Grohmann, K. (2004). Prolific Domains in the Computational System. *Actas de JEL*, 211-216.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse.

Grosz, P. G. (2016). Information structure and discourse particles. In *The Oxford handbook of information structure*.

Gundel, J. K. (1974). *The role of topic and comment in linguistic theory*. University of Texas at Austin.

Gundel, J. K. (1985). "Shared knowledge" and topicality. *Journal of Pragmatics*, *9*, 83–107.

Gundel, J. K. (1988). Universals of topic-comment structure. *Studies in syntactic typology*, *17*, 209-239.

Gundel, J. K., & Fretheim, T. (2004). Topic and focus. *The Handbook of Pragmatics*, *175*, 196.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274-307.

Gundel, J. K., Hegarty, M., & Borthen, K. (2003). Cognitive status, information structure, and pronominal reference to clausally introduced entities. *Journal of Logic, Language and Information*, *12*(3), 281-299.

Guru, K. (1920) *Hindi Vyakaran*. Banaras Nagari Pracharini Sabha.

Gutzmann, D. (2013). Expressives and beyond: An introduction to varieties of use-conditional meaning. In *Beyond expressives: Explorations in use-conditional meaning* (pp. 1-58). Brill.

Gutzmann, D. (2015). *Use-conditional meaning: Studies in multidimensional semantics* (Vol. 6). OUP Oxford.

Gutzmann, D. (2017). Modal particles≠ modal particles (= modal particles). In *Discourse Particles* (pp. 144-172). De Gruyter.

Gutzmann, D. (2019). *The Grammar of expressivity* (Vol. 72). Oxford University Press.

Gutzmann, D., & Turgay, K. (2012). Expressive intensifiers in German: syntax-semantics mismatches. *Empirical issues in syntax and semantics*, *9*, 149-166.

Gutzmann, D., & Turgay, K. (2015). Expressive intensifiers and external degree modification. *The Journal of Comparative Germanic Linguistics*, *17*(3), 185-228.

Haegeman, L. (2006). 'Argument fronting in English, Romance CLLD and the left periphery'. *Cross-Linguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture.*, p-25.

Haegeman, L. (2006). Conditionals, factives and the left periphery. *Lingua*, *116*(10), 1651-1669.

Haiman, J. (1978). Conditionals Are Topics. *Language*, *54*(3), 564–589. https://doi.org/10.2307/412787

Hajicová, E. (1983). Topic and focus. *Theoretical Linguistics*, *10*(2-3), 268–276.

Hajicová, E., Partee, B. B., & Sgall, P. (2013). *Topic-focus articulation, tripartite structures, and semantic content* (Vol. 71). Springer Science & Business Media.

Hajičová, E., Sgall, P., & Skoumalová, H. (1993). Identifying topic and focus by an automatic procedure (pp. 178–182). Presented at the Proceedings of the sixth conference on European chapter of the Association for Computational Linguistics, Association for Computational Linguistics.

Hajičová, E., Sgall, P., & Skoumalová, H. (1995). An automatic procedure for topic-focus identification. *Computational Linguistics*, *21*(1), 81–94.

Hale, K., & Keyser, S. J. (1993). On argument structure and the lexical expression of syntactic relations. *1993*, 53-109.

Halliday, M. A. (1967). Notes on transitivity and theme in English Part I. *Journal of linguistics*, *3*(01), 37-81.

Han, C. H. (1998). Asymmetry in the interpretation of-(n) un in Korean. *Japanese/Korean Linguistics*, *7*, 1-15.

Han, C. H. (2002). Interpreting interrogatives as rhetorical questions. *Lingua*, *112*(3), 201-229.

Hara, Y. (2006a). *Grammar of knowledge representation: Japanese discourse items at interfaces*. University of Delaware.

Hara, Y. (2006b). Implicature unsuspendable: Japanese contrastive wa. *Proceedings of Texas Linguistics Society*, *8*, 35-45.

Harbour, D., Watkins, L. J., & Adger, D. (2012). Information Structure, Discourse Structure, and Noun Phrase Position in Kiowa. *International Journal of American Linguistics*, *78*(1), 97–126. https://doi.org/10.1086/662639

Hartmann, D. (1998). Particles. In *Concise encyclopedia of pragmatics* (p. 520). Oxford: Elsevier.

Hedberg, N., & Zacharski, R. (Eds.). (2007). *The Grammar Pragmatics Interface: Essays in Honor of Jeanette K. Gundel* (Vol. 155). John Benjamins Publishing.

Heim, I. (1982). The semantics of definite and indefinite noun phrases.

Hetland, J. (2007). The Korean particle nun, the English fall-rise accent and thetic/categorial judgements. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 117–128). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Hetland, J. (2007). The Korean particle nun, the English fall-rise accent, and thetic/categorical judgements. In *On information structure, meaning and form* (pp. 117-127). John Benjamins.

Heycock, C. (2008). Japanese-wa,-ga, and information structure. In *The Oxford handbook of Japanese linguistics*.

Hinterwimmer, S. (2011). 71. Information structure and truth-conditional semantics. In *Volume 2* (pp. 1875-1908). De Gruyter Mouton.

Hintikka, K. J. J. (1962). Knowledge and belief: An introduction to the logic of the two notions.

Hiraiwa, K., & Ishihara, S. (2002). Missing links: Cleft, sluicing, and "no da" construction in Japanese. *MIT working papers in linguistics*, *43*, 35-54.

Holmberg, V. (2020, September 29). *Topic Particles, Agreement and Movement* [Invited Talk]. Chinese University of Hong Kong.

Hoof, H. van. (1997). Left Dislocation and Split Topics in Brabant Dutch. In *Materials on Left Dislocation* (Vol. 14, pp. 275–306). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Horn, L. (1989). *A natural history of negation.*

Horn, L. R. (1972). *On the semantic properties of logical operators in English*. University of California, Los Angeles.

Horn, Laurence R. (2004). Implicatures. In Laurence R. Horn and Gregory L. Ward (eds.), *Handbook of pragmatics*, 3–28. Oxford: Blackwell Publishers.

Horn, Lawrence R. 1984. Towards a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Deborah Schiffrin (ed.), *Meaning, form, and use in context: Linguistic applications*, 11–42. Georgetown: Georgetown University Press.

Iatridou, S. (1988). Clitics, anaphors, and a problem of coindexation. *Linguistic Inquiry*, *19*(4), 698-703.

Iatridou, S. (1995). Clitics and island effects. *University of Pennsylvania working papers in linguistics*, *2*(1), 3.

Ilkyu, K. (2019). On the Mechanism of Generating Contrastive Implicature of CT in Korean. *한국어학*, *85*, 31-59.

Jackendoff, R. S. (1972). Semantic interpretation in generative grammar.

Jacobs, J. (1997). I-Topikalisierung. Linguistische Berichte, 168:91–134.

Jacobs, J. (2001). The dimensions of topic-comment. *Linguistics*, *39*(4; ISSU 374), 641-682.

Jacobs, S., & Jackson, S. (1983). Strategy and structure in conversational influence attempts. *Communications Monographs*, *50*(4), 285-304.

Jager, G. (1999). Topic, Focus, and Weak Quantifiers. In *FOCUS Linguistic, Cognitive, and Computational Perspectives* (pp. 187–212). Cambridge University Press.

Jiménez-Fernández, Á. L. (2020). *Syntax-information structure interactions in the sentential, verbal and nominal peripheries*. Cambridge Scholars Publishing.

Johnston, M. (1994). When-clauses, adverbs of quantification, and focus. In *Proceedings of WCCFL* (Vol. 13, pp. 549-564).

Kaplan, D. (1999). The meaning of ouch and oops: Explorations in the theory of meaning as use. *Manuscript, UCLA*.

Katz, J., & Selkirk, E. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, 771-816.

Kayne, R. S. (1994). *The antisymmetry of syntax* (No. 25). Mit Press.

Kellogg, S. H., & Bailey, T. G. (1938). *A Grammar of the Hindi Language: In which are Treated the High Hindi, Braj, and the Eastern Hindi of the Rámáyan of Tulsi Dás; with Notes on Pronunciation by T. Grahame Bailey*. Routledge & Paul.

Kibort, A. (2008). Grammatical Features Inventory.

Kibort, A., & Corbett, G. G. (Eds.). (2010). *Features: Perspectives on a key notion in linguistics*. Oxford University Press.

Kidwai, A. (1999). Word order and focus positions in Universal Grammar. *The grammar of focus*, 213-244.

Kidwai, A. (2000). *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi Urdu*. Oxford University Press.

Kidwai, A. (2004). The topic interpretation in Universal Grammar. In Clause Structure in South Asian Languages (pp. 253-289). Springer Netherlands.

Kim, I. (2013). *Korean-(n) un, salience, and information structure*. Yale University.

Kim, I. (2015). Is Korean-(n) un a topic marker? On the nature of-(n) un and its relation to information structure. *Lingua*, *154*, 87-109.

Kim, I. (2016). Can Korean-(n) un mark (contrastive) focus?. *Language Sciences*, *56*, 105-117.

Kim, Ilkyu. (2018). Is contrastive implicature induced by Korean CT-marking-(n)un conventional or conversational?. In *Korean Journal of Linguistics* 43-2, 223-247.

Kiss, E. (1987).*Configurationality in Hungarian*. Dordrecht: Reidel.

Kiss, K. É. (1977). THE NOTIONS OF TOPIC AND FOCUS. *Angol Filológiai Tanulmányok / Hungarian Studies in English*, *11*, 211–223.

Kiss, K. É. (1998). Identificational focus versus information focus. *Language*, *74*(2), 245-273.

Kramsch, C. J. (1983). Discourse Function of Grammar Rules: Topic Construction in German. *The Modern Language Journal*, *67*(1), 13–22. http://doi.org/10.2307/326688

Kratzer, A. (1999). Beyond ouch and oops: How descriptive and expressive meaning interact. In *Cornell conference on theories of context dependency* (Vol. 26). Ithaca, NY: Cornell University.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3-4), 243–276.

Kruijff-Korbayová, I., & Steedman, M. (2003). Discourse and Information Structure. *Journal of Logic, Language, and Information*, *12*(3), 249–259.

Kuno, S. & Takami, K. I. (1993). *Grammar and discourse principles: Functional syntax and GB theory*. University of Chicago Press.

Kuno, S. (1972). Functional sentence perspective: A case study from Japanese and English. *Linguistic inquiry*, *3*(3), 269-320.

Kuno, S. (1973). Constraints on internal clauses and sentential subjects. *Linguistic Inquiry*, *4*(3), 363-385.

Kuno, S. (1987). *Functional syntax: Anaphora, discourse and empathy*. University of Chicago Press.

Kuroda, S.-Y. (2005). Focusing on the Matter of Topic: A Study of Wa and Ga in Japanese. *Journal of East Asian Linguistics*, *14*(1), 1–58.

Ladusaw, W. A. (1994, November). Thetic and categorical, stage and individual, weak and strong. In *Semantics and Linguistic Theory* (Vol. 4, pp. 220-229).

Lahiri, U. (1991). *Embedded interrogatives and predicates that embed them* (Doctoral dissertation, Massachusetts Institute of Technology).

Lahiri, U. (2017). Generalized Quantifiers in Logic and Natural Language. *The EFL Journal*, *8*(1).

Lakshmi Bai, B. (1977). Syntax and semantics of the particle to in Hindi. *Osmania Papers in Linguistics*, *3*, 64-75.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Great Britain: Cambridge University Press.

Lambrecht, K. (2000). When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, *24*(3), 611-682.

Law, P. (2007). Topicalization in Malagasy, Tagalog and Tsou. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 129–154). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Lee, C. (2000). Contrastive predicates and conventional scales. *CLS*, *36*(1), 243-257.

Lee, C. (2006). Contrastive topic/focus and polarity in discourse. *Where semantics meets pragmatics*, *16*, 381-420.

Lee, C., & Gordon, M. (Eds.). (2006). *Topic and focus: Cross-linguistic perspectives on meaning and intonation* (Vol. 82). Springer Science & Business Media.

Lee, H. (1989). Toward a formal characterization of topic construction with special reference to Korean (Vol. 2, pp. 197–210). Presented at the WECOL.

Levinson, S. C. (1983). Pragmatics.

Li, Charles N., & Thompson, S. (1976) Subject and topic: a new typology of language. In Charles N. Li (ed.) *Subject and Topic*. Academic Press: New York.

Lohndal, T. (2013). Sentential subjects: topics or real subjects. In *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL 31)*.

López, L. (2007). *Locality and the architecture of syntactic dependencies*. Springer.

López, L. (2009). *A derivational syntax for information structure*. Oxford University Press, USA.

López, L. (2014). *A derivational syntax for information structure* (Vol. 23). Oxford University Press, USA.

Mahajan, A. (1992). The specificity condition and the CED. *Linguistic Inquiry*, 510-516.

Mahajan, A. (2017). Quantification in Hindi. In *Handbook of Quantifiers in Natural Language: Volume II* (pp. 383-430). Springer, Cham.

Mahajan, A. K. (1990). *The A/A-bar distinction and movement theory* (Doctoral dissertation, Massachusetts Institute of Technology).

Maienborn, C., von Heusinger, K., & Portner, P. (Eds.). (2011). *Semantics: An international handbook of natural language meaning* (Vol. 2). Walter de Gruyter.

Manetta, E. (2014). The shape of the causative verbal domain: Evidence from Kashmiri. *Syntax*, *17*(3), 235-268.

Marantz, A. (2008). Phases and words. In *Phases in the theory of grammar* (pp. 191-222). Dong-In Publishing Co.

Masica, C. P. (1993). *The indo-aryan languages*. Cambridge University Press.

Matić, D., & Wedgwood, D. (2013). The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis1. *Journal of Linguistics*, *49*(1), 127-163.

McCready, E. (2010). Varieties of conventional implicature. *Semantics and Pragmatics*, *3*, 8-1.

McNally, L. (1998). On recent formal analyses of topic. In *The Tbilisi symposium on language, logic, and computation: Selected papers* (Vol. 14, pp. 147-160). Stanford, CA: CSLI Publications.

Meinunger, A. (2000). *Syntactic Aspects of Topic and Comment*. Amsterdam: John Benjamins Publishing Co. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=253439&site=eds-live&scope=site

Miyagawa, S. (1987). Wa and the Wh Phrase. In *Perspectives on Topicalization: The Case of the Japanese "Wa."*

Miyagawa, S. (2010). *Why agree? Why move?: Unifying agreement-based and discourse-configurational languages*. MIT Press.

Miyagawa, S. (2017). Topicalization. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)*, *152*, 1-29.

Mohanan, T. (1994). *Argument structure in Hindi*. Center for the Study of Language (CSLI).

Molnár, V. (1998). On the Syntax, Phonology, Semantics and Pragmatics of the So-Called" Contrastive Topic" in Hungarian and German. *Acta Linguistica Hungarica*, *45*(1), 89-166.

Molnár, V. (2002). Contrast–from a contrastive perspective. In *Information structure in a cross-linguistic perspective* (pp. 147-161). Brill.

Molnár, V. (2006). 'On Different Kinds of Contrast', in V. Molnár and S. Winkler (eds), *The Architecture of Focus*. Berlin: Mouton de Gruyter, 197–234.

Molnár, V., & Winkler, S. (Eds.). (2006). *The architecture of focus* (Vol. 82). Berlin: Mouton de Gruyter.

Molnár, V., Egerland, V., & Winkler, S. (2019). Exploring the Architecture of Topic at the Interface of Grammar and Discourse. In *Architecture of Topic* (pp. 1-44). De Gruyter Mouton.

Molnarfi, L. (2007). On the discourse configurationality of West Germanic. In *On Information Structure, Meaning and Form* (Vol. 100, pp. 155–182). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Montaut, A. (2015). 12 The discourse particle to and word ordering in Hindi: From grammar to discourse. *Information structuring of spoken language from a cross-linguistic perspective*, *283*, 263.

Mulkern, A. E. (2007). Knowing who's important. *The Grammar Pragmatics Interface: Essays in honor of Jeanette K. Gundel*, *155*, 113.

Müller, G. (1998). Incomplete category fronting. In *Incomplete Category Fronting* (pp. 1–28). Springer.

Muller, G. and W. Sternefeld 1993. Improper Movement and Unambiguous

Musan, R., & Krifka, M. (2012). *The Expression of Information Structure*. Berlin: De Gruyter Mouton. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=494256&site=eds-live&scope=site

Nair, R. B. (1991). Expressing doubt and certainty: The tag question and the 'to'particle in some Indian languages. *Language Sciences*, *13*(2), 207-227.

Neeleman, A., & Van de Koot, H. (2008). Dutch scrambling and the nature of discourse templates. *The Journal of Comparative Germanic Linguistics*, *11*(2), 137-189.

Neeleman, A., & Vermeulen, R. (2012). *The Syntax of Topic, Focus, and Contrast : An Interface-based Approach*. Berlin: De Gruyter Mouton. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=530575&site=eds-live&scope=site

Neeleman, A., & Vermeulen, R. (Eds.). (2012). *The syntax of topic, focus, and contrast: An interface-based approach* (Vol. 113). Walter de Gruyter.

Oshima, S. (2001). Phase and Phase Collapse: a Study of Topicalization and Focusing (Part I). *Journal of Inquiry and Research*, *74*, 1-18.

Pandharipande, R., & Kachru, Y. (1977). Relational grammar, ergativity, and Hindi-Urdu. *Lingua*, *41*(3-4), 217-238.

Paoli, S. (2007). The fine structure of the left periphery: COMPs and subjects: Evidence from Romance. *Lingua*, *117*(6), 1057-1079.

Pareek, B. (2017). Nominal Features in Hindi Language Acquisition: A study of agreement and modification. PhD Thesis. JNU.

Partee, B. (1991). Topic, Focus and Quantification. In *Proceedings of SALT I*. Cornell.

Partee, B. (1991, April). Topic, focus and quantification. In *Semantics and Linguistic Theory* (Vol. 1, pp. 159-188).

Partee, B. H., & Sgall, P. (Eds.). (1996). *Discourse and Meaning: Papers in honor of Eva Haji? ová*. John Benjamins Publishing.

Paul, I. (2002). The syntactic encoding of topic and focus.

Paul, W., & Pan, V. J. (2017). What you see is what you get: Chinese sentence-final particles as head-final complementizers. In *Discourse Particles* (pp. 49-77). De Gruyter.

Platzack, C. (2000). Multiple interfaces. *Cognitive Interfaces. Constraints on Linking Cognitive Information, Oxford University Press, Oxford*, 21-53.

Polinsky, M. (1999). *Language*, *75*(3), 567–582. http://doi.org/10.2307/417062

Pollard, C., & Sag, I. A. (1987). Information-based Syntax and Semantics, Vol. 1. Number 13 in Lecture Notes. *CSLI Lecture Notes, Stanford University, Stanford, USA*.

Portner, P. (2005). What is meaning: Fundamentals of formal semantics.

Portner, P. (2007). Instructions for interpretation as separate performatives. *On information structure, meaning and form*, 407-426.

Portner, P., & Yabushita, K. (1998). The Semantics and Pragmatics of Topic Phrases. *Linguistics and Philosophy*, *21*(2), 117–157.

Potts, C. (2003). Expressive content as conventional implicature. In *Proceedings-Nels* (Vol. 33, pp. 303-322).

Potts, C. (2005). *The logic of conventional implicatures* (Vol. 7). OUP Oxford.

Potts, C. (2007b). 'The expressive dimension', *Theoretical Linguistics*33.2: 165–97.

Potts, C. (2012). 94. Conventional implicature and expressive content. In *Semantics: Volume 3* (pp. 2516-2536). De Gruyter Mouton.

Potts, C. (2012). 94. Conventional implicature and expressive content. In *Semantics: Volume 3* (pp. 2516-2536). De Gruyter Mouton.

Potts, C., & Kawahara, S. (2004). Japanese honorifics as emotive definite descriptions. In *Semantics and Linguistic Theory* (Vol. 14, pp. 253-270).

Prince, E. F. (1981, June). Topicalization, focus-movement, and Yiddish-movement: A pragmatic differentiation. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 7, pp. 249-264).

Prince, E. F. (1984). Topicalization and Left-Dislocation: A Functional Analysisa. *Annals of the New York Academy of Sciences*, *433*(1), 213–225.

Prince, E. F. (1986). On the syntactic marking of presupposed open propositions. In *Proceedings of the 22nd Annual Meeting of the Chicago Linguistic Society*.

Prince, Ward, G., & Ellen. (1991). On the Topicalization of Indefinite NPs. *Journal of Pragmatics*, *16*, 167–177.

Recanati, F.(2004b). 'Pragmatics and semantics', in Laurence R. Horn and Gregory Ward (eds), *The Handbook of Pragmatics*. Oxford: Blackwell, 442–62.

Reinhart, T. (1981). Pragmatics and Linguistics: An Analysis of Sentence Topics in Pragmatics and Philosophy I. *Philosophica anc Studia Philosophica Gandensia Gent*, *27*(1), 53-94.

Reinhart, T. (2004). Topics and the conceptual interface. *Context dependence in the analysis of linguistic meaning*, 275-305.

Repp, S. (2010). Defining 'contrast' as an information-structural notion in grammar. *Lingua*, *120*(6), 1333-1345.

Repp, S. (2016). Contrast: Dissecting an elusive information-structural notion and its role in grammar. *Oxford handbook of information structure.* Oxford: Oxford University Press.

Richards, N. (2010). *Uttering trees* (Vol. 56). MIT Press.

Rivero, M.-L. (1980). On Left-Dislocation and Topicalization in Spanish. *Linguistic Inquiry*, *11*(2), 363–393.

Rizzi, L. (1990). *Relativized minimality.* The MIT Press.

Rizzi, L. (1997). The fine structure of the left periphery. In *Elements of grammar* (pp. 281-337). Springer, Dordrecht.

Rizzi, L. (2004). *The structure of CP and IP: The cartography of syntactic structures* (Vol. 2). Oxford University Press.

Roberts, C. (1995). Domain restriction in dynamic semantics. In *Quantification in natural languages* (pp. 661-700). Springer, Dordrecht.

Roberts, C. (1998). The place of centering in a general theory of anaphora resolution. *Centering theory in discourse*, 359-400.

Roberts, C. (2011). 72. Topics. In *Semantics: Volume 2* (pp. 1908-1934). De Gruyter Mouton.

Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and pragmatics*, *5*, 6-1.

Rochemont, M. (1986). *Focus in generative grammar* (Vol. 4). John Benjamins Publishing.

Rochemont, M. (1989). Topic islands and the subjacency parameter. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *34*(2), 145-170.

Rochemont, M. (2014). *A Theory of Stylistic Rules in English (RLE Linguistics A: General Linguistics)*. Routledge.

Rochemont, M. (2019). Topics and givenness. In *Architecture of Topic* (pp. 47-66). De Gruyter Mouton.

Rooth, M. (1985). *Association with Focus (Montague Grammar, Semantics, Only, Even)* (Doctoral dissertation, University of Massachusetts Amherst).

Rooth, M. (1992). A theory of focus interpretation. Natural Language Semantics, 75-116.

Rooth, M. (1996). Focus. Lappin, Shalom (ed.), *The Handbook of Contemporary Semantic Theory*, Blackwell, Oxford, 271-297.

Ross, J. R. (1970). On declarative sentences. *Readings in English transformational grammar*, *222*, 272.

Rothstein, S. D. (1991). Syntactic licensing and subcategorization. In *Perspectives on Phrase Structure: Heads and Licensing* (pp. 137-157). Brill.

Sadock, J. M. (1969a). Hypersentences. *Research on Language & Social Interaction*, *1*(2), 283-370.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and philosophy*, *27*(3), 367-391.

Sawada, O.(2013).The Japanese contrastive wa: a mirror image of EVEN In Thera Crane, Oana David, Donna Fenton, Hannah J. Haynie, Shira Katseff, Russell Lee-Goldman, Ruth Rouvier, and Dominic Yu, eds., Proceedings of the 33rd Annual Meeting of the Berkeley Linguistics Society, Berkeley, CA: BLS, pp. 374–387

Schwarz, B., & Shimoyama, J. (2011). Negative Islands and obviation by'wa'in Japanese degree questions. In *Semantics and Linguistic Theory* (Vol. 20, pp. 702-719).

Schwarzschild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural language semantics*, *7*(2), 141-177.

Searle, J. R. (1976). A classification of illocutionary acts1. *Language in society*, *5*(1), 1-23.

Selkirk, E. (1984). Phonology and Syntax. Cambridge, Mass. MIT Press.

Şener, S. (2010). Peripheral matters in Turkish syntax. *Storrs, CT: University of Connecticut Dissertation*.

Sgall, P., Hajicová, E. & Panevova, J. (1986). *The meaning of the sentence in its semantic and pragmatic aspects*. Springer Science & Business Media.

Sgall, P., Hajicova, E., & Benesova, E. (1973). *Topic, Focus and Generative Semantics*. Germany.

Sgall, P., Hajicová, E., & Panevová, J. (1986). *The meaning of the sentence in its semantic and pragmatic aspects*. Springer Science & Business Media.

Sharma, D. (2003). Discourse clitics and constructive morphology in Hindi. In *Nominals inside and out* (pp. 59-85). CSLI Publications.

Shi, D. (2000). Topic and topic-comment constructions in Mandarin Chinese. *Language*, 383-408.

Simons, M. (2012). 92. Implicature. In *Semantics: Volume 3* (pp. 2460-2486). De Gruyter Mouton.

Speas, P., & Tenny, C. (2003). Configurational properties of point of view roles. *Asymmetry in Grammar1*, 315-345.

Spencer, A. (2005). Case in Hindi. In *Proceedings of the LFG05 Conference* (pp. 429-446). Stanford, CA: CSLI Publications.

Sperber, D., & Wilson, D. (1997). Remarks on relevance theory and the social sciences.

Sperber, D., & Wilson, D. (2004). Experimental pragmatics.

Stalnaker, R. (1974). Pragmatic presupposition. In M. Munitz & P. Unger (eds.), *Semantics and philosophy*. New York: New York University Press, 197-214.

Stalnaker, R. (2002). Common ground. *Linguistics and philosophy*, *25*(5/6), 701-721.

Steedman, M. (1996). Surface structure and interpretation Cambridge.

Steedman, M. (2000). *The syntactic process* (Vol. 24). Cambridge, MA: MIT press.

Strawson, P. F. (1964). Identifying reference and truth-values. *Theoria*, *30*(2), 96-118.

Svenonius, P. (2007). arguments they introduce. *Argument structure*, *108*, 63.

Szabolcsi, A. (1981a). The possessive construction in Hungarian: A configurational category in a non-configurational language. *Acta Linguistica Academiae Scientiarum Hungaricae*, *31*(1/4), 261-289.

Szabolcsi, A. (1981b). The semantics of topic-focus articulation. *3rd Amsterdam Colloquium Proceedings*.

Szabolcsi, A. (1997). Strategies for scope taking. In *Ways of scope taking* (pp. 109-154). Springer, Dordrecht.

Taguchi, S. (2008). Against the null complementizer analysis of Japanese relative clauses. *Nanzan Linguistics Special Issue*, *3*(2), 185-198.

Theodora Alexopoulou, & Dimitra Kolliakou. (2002). On Linkhood, Topicalization and Clitic Left Dislocation. *Journal of Linguistics*, *38*(2), 193–245.

Tomioka, S. (2010). Contrastive topics operate on speech acts. *Information structure: Theoretical, typological, and experimental perspectives*, *115138*.

Tomioka, S. (2016). Information structure in Japanese. *The Oxford handbook of information structure*, 753-773.

Tomioka, S. (2020). 17 Conversational implicature. In *Handbook of Japanese Semantics and Pragmatics* (pp. 773-798). De Gruyter Mouton.

Uriagereka, J. (1995). An F position in western romance. *Discourse configurational languages*, *26*, 79-123.

Uriagereka, J. (2003). Pure Adjuncts. Manuscript, UMD.

Vallduvi, E. (1990). *The Informational Component* (Doctoral dissertation, Ph. D. thesis, University of Pennsylvania, Philadelphia, PA).

Vallduví, E. (1992). A preverbal landing site for quantificational operators. *Catalan working papers in linguistics*, *2*, 319-343.

Vallduví, E. (1995). Structural properties of information packaging in Catalan. *Discourse configurational languages*, 122-152.

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, *34*(3), 459–520.

Vallduví, E., & Vilkuna, M. (1998). On rheme and kontrast. In *The limits of syntax* (pp. 79-108). Brill.

Van Dijk, T. A. (1977). Sentence topic and discourse topic. *Papers in Slavic philology*, *1*(1977), 49-61.

Van Kuppevelt, J. (1996). Inferring from Topics: Scalar Implicatures as Topic-Dependent Inferences. *Linguistics and Philosophy*, *19*(4), 393–443.

Van Rooij, R., & Schulz, K. (2004). Exhaustive interpretation of complex sentences. *Journal of logic, language and information*, *13*(4), 491-519.

Van Valin, R. D. & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

Van Valin, R. D. (1993b). A synopsis of Role and Reference Grammar. *Advances in role and reference grammar*, *1*, 164.

Van Valin, R. D. (1999b). Generalized semantic roles and the syntax-semantics interface. *Empirical issues in formal syntax and semantics*, *2*, 373-389.

Verma, M. K. (1971). *The structure of the noun phrase in English and Hindi*. Motilal Banarsidass Publ.

Villalba, X. (1998). Right dislocation is not right dislocation. *Studies on the syntax of central Romance languages. Girona: Universitat de Girona*, 227-241.

von der Gabelentz, H. (1869). "Ideen zu einer vergleichenden Syntax. Wort- und Satzstellung", Zeitschrift für Völkerpsychologie und Sprachwissenschaft 6, 376-84.

Von Fintel, K. U. (1994). *Restrictions on quantifier domains* (Doctoral dissertation, University of Massachusetts Amherst).

Von Heusinger, K., & Schumacher, P. B. (2019). Discourse prominence: Definition and application. *Journal of Pragmatics*, *154*, 117-127.

Wagner, M. (2008). A compositional analysis of contrastive topics. In *Proceedings of NELS* (Vol. 38, No. 2007, pp. 415-428).

Wagner, M. (2009). Focus, topic, and word order: A compositional view. *Alternatives to cartography*, 53-86.

Wagner, M. (2012). Contrastive topics decomposed. *Semantics and Pragmatics*, *5*(8), 1-54.

Wagner, M. (2012). Contrastive topics decomposed. *Semantics and Pragmatics*, *5*, 8-1.

Ward, G., & Birner, B. (2004). Information structure and non-canonical syntax. *The Handbook of Pragmatics*, 153–174.

Ward, G., & Birner, B. J. (2001). Discourse and information structure. *The handbook of discourse analysis*, 119-37.

Ward, G., & Birner, B. J. (2003). 6 Discourse and Information Structure. *The handbook of discourse analysis*, *18*, 119.

Ward, G., & Hirschberg, J. (1985). Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, 747-776.

Ward, Gregory L. (1985) *The Semantics and Pragmatics of Preposing*. Ph.D. dissertation, University of Pennsylvania.

Wedgwood, D. (2008). Nomi Erteschik-Shir, Information structure: The syntax–discourse interface (Oxford Surveys in Syntax and Morphology 3). Oxford: Oxford University Press, 2007. Pp. vii+ 246. *Journal of Linguistics*, *44*(01), 233-238.

Wilson, D., & Sperber, D. (2002). Relevance theory.

Wiltschko, M. & Heim, J. (2016).The Syntax of Confirmationals: A Neo-Performative Analysis. In Outside the Clause: Form and Function of Extra-Clausal Constituents, edited by Gunther Kaltenböck, Evelien Keizer, and Arne Lohmann, 305–340. London: John Benjamins.

Wiltschko, M. (2014). *The universal structure of categories: Towards a formal typology* (Vol. 142). Cambridge University Press.

Wiltschko, M. (2020, September 115). *What Language in Interaction can tell us about the relation between Language, Thought and Communication* [Invited Talk]. Chinese University of Hong Kong.

Wiltschko, M. (2021). *The Grammar of Interactional Language*. Cambridge University Press.

Winkler, S. (2005). Ellipsis and focus in generative grammar. Vol. 81. *Studies in generative grammar. Walter de Gruyter*.

Zimmermann, M. (2011).Discourse Particles. In Portner, Paul & Claudia Maienborn & Klaus von Heusinger, eds. *Handbook of Semantics*. Berlin and New York: de Gruyter.

Zimmermann, M., & Féry, C. (Eds.). (2009). *Information structure: Theoretical, typological, and experimental perspectives*. Oxford University Press.

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural language semantics*, *8*(4), 255-290.

Zubizarreta, M. L. (1998). *Prosody, focus, and word order*. MIT Press.